Scheme for Promotion of Academic and Research Collaboration

Author

SPARC Monograph

Optional Subtitle

Hema A. Murthy

Optional Affiliation
Optional Specification

2022

Article submitted to SPARC Monograph
Hema A. Murthy, Indian Institute of Technology, Madras
No. 1234

ISSN 1234-5678

Cover: SPARC Monograph.
Print production: Indian Institute of Technology, Madras.

To my ghostwriter

Abstract

From our SPARC proposal (it looks like we had specified potential content)
Book/Monograph Outline
PROPOSED TITLE: Digital Audio Processing Tools for Empirical Musicology Research (add: and Musical instrument research)
AUTHOR(S): All contributing project investigators: Rao, Serra, Clayton, Ramamoorthy, Prasanna, Scavone, Murthy, Patil
BRIEF DESCRIPTION (5-10 sentences)

The monograph focuses on the music and its embodiment as an audio recording. A review of the types of musical instruments (with special focus on traditional Indian instruments) and singing voice will cover their structures, acoustic principles and the sound palette description. Qualitative and quantitative methods for sound analyses with signal representations will be presented. Music is a highly structured signal in any genre and typical structural principles will be linked to corresponding signal representations. This will facilitate discussions of melody and rhythm as well as timbre from the points of view of production and perception.

To help initiate research in music, corpora design principles will be reviewed. Add: measurements for musical instrument studies will be reviewed. Case studies will be presented to illustrate research methodologies.

PURPOSE AND NEED
The monograph will be targeted towards (i) musicians and musicologists looking to expand their understanding and scope of research via analyses of performance audio with digital tools, (ii) engineers and computer scientists seeking applications of acoustics, signal processing and machine learning in music. While ethnomusicology is the prime beneficiary of digital processing for audio recordings analyses, much of the material will serve to increase the appreciation of the universal connections between musical attributes and signal/physical measurements. This can lead to offerings that enhance music pedagogy and musical instrument building. The exposition will be truly multidisciplinary in contrast to the current literature that is clearly segregated by specific approaches coming from a single discipline.

Table of Contents
Chapters:
1. Types of Musical Sound Sources, their structure, acoustic principles, sound palette
2. Acoustic measurements and perceptual descriptions
3. Use of time-frequency representations for audio description
4. Music structure detection
5. Melodic analyses
6. Rhythmic analyses
7. Instrument timbre analyses (including vocal)
8. Music corpora design and annotation
9. Case studies of analyses of performance audios

Report on Scheme for Promotion of Academic and Research Collaboration (SPARC) sponsored

Invited Talks with Panel Discussion on "Music Technology and Music Instrument Modeling" and A course on "Modeling and measurements of musical instruments" Instructor: Prof. Gary P. Scavone, McGill

University, Canada from 9th - 13th December 2019 at Indian Institute of Technology Dharwad, 580011, Dharwad, Karnataka, India.

The invited talks, panel discussion and a course on the theme of "Music Instrument Modeling" were conducted at IIT Dharwad from 9th to 13th December 2019. The course was conducted under the sponsorship of SPARC project (ID:147, administered by IIT Bombay, Principal Investigator: Prof. Preeti Rao, Department of Electrical Engineering, IIT Bombay).

The resource persons delivered invited talks on various topics related to the theme on 9th December, 2019. The resource persons were Prof. Gary Scavone (Schulich School of Music, McGill University, Canada), Prof. Xavier Serra (Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain), Mr. G Raj Narayan (Chief Innovator and Managing Director, Radel), Prof. Preeti Rao (IIT Bombay), and Prof. Hema Murthy (IIT Madras). Prof. Gary Scavone gave directions in musical acoustics research. He detailed the objectives of the research in his lab as to investigate possible design or material improvements for musical instruments (e.g. air column or mouthpiece variations to improve intonation or response), to create sound synthesis models for use in music making, to understand how the musical instruments work and to develop physics-based modeling with computers. Prof. Xavier Serra gave an overview of Music Technology Research in his research group which spans into various research topics as Audio signal processing, Machine learning, Human computer interaction, Music information retrieval, Computational musicology. He explained various music applications developed in the laboratory such as Dunya desktop, Saraga and Riyaz. Mr. G. Raj Narayan delivered a talk on "Art and Science of Musical Instruments". He detailed various music instrument issues faced by the artists and gave directions for scientific interventions. Prof. Preeti Rao summarized various high-level music descriptors for retrieval such as melody, rhythm, harmony and timbre. She detailed the "structure" in Hindustani concerts, structural segmentation, Musicological cues to structural segments and Structural segmentation for sitar/sarod concerts. Prof. Hema Murthy delivered an invited talk on "Analysis of carnatic music". She also detailed music information retrieval process. She emphasized on the elaboration of one or more pieces in a concert, such as, Kriti, Raagam Taanam Pallavi and Other types: padham, jaavali, viruttam, slokam, varnam.

These invited talks were then followed by a panel discussion at the end of the first day. The discussion was anchored by Dr. Ajay Srinivasamurthy, Applied Scientist at Amazon.com, Bangalore, India. Few topics for the discussion were (a) Building a community in Music Technology, (b) Machine learning in music technology and musical instrument modeling, (c) Industry's role in the growth of music technology, and (d) Current initiatives and future directions in music technology.

The invited talks on panel discussion on 9th December 2019 were public events and music enthusiastic from Hubli-Dharwad were invited. Further, the IIT Dharwad faculty members (including Co-PIs at IIT Dharwad, Prof. SRM Prasanna and Dr. Dhiraj V. Patil), MS, PhD research scholars and few BTech students attended the first day events. The course participants from other institutions have also participated in this program.

A course on "Modeling and measurements of musical instruments" was delivered by Prof. Gary Scavone (Schulich School of Music, McGill University, Canada), from 10th - 13th December 2019 at IIT Dharwad. The course was hands-on with programming using Matlab in the computer lab and focused on methods for discrete-time modeling and sound synthesis of musical acoustic systems, as well as common techniques for measuring the characteristics of musical instruments. The course will be oriented toward students with an existing background in digital signal processing (DSP). The initial emphasis will be on signal processing methods for efficient synthesis of simple string and wind instruments. Topics covered will include the use of delay lines to simulate wave propagation and digital waveguide techniques to properly account for associated discontinuities or boundary conditions. Measurement techniques to be covered will include approaches to determine admittances or impedances of string and wind instrument bodies, as well as experimental modal analysis and other imaging techniques to visualize the vibrations of systems. The course was attended by around 15 registered participants, majority of them were research scholars. Overall, the course reception was good among the young research scholars.

Preface

Music pervades life across geographies and cultures, probably more so compared to any of the other performing arts. Over the years, music has been discussed and written about, just as much as it is performed, across its many aspects including the aesthetic, socio-cultural and cognitive. In parallel with this, technology has made inroads to directly impact music practice by way of synthesized instruments including automatic accompaniment aids. Computational musicology refers to research on music that employs the tools and techniques of science, mathematics, computer science and technology. With the explosive growth of digital archives, it is only natural that musicology has attracted computational interest that is well suited to the analyses of large corpora.

Indian art music has become quite structured over the last 500 years, and is well understood by practitioners and trained listeners even as it remains in a state of continuous evolution. However, being an oral tradition, it has hardly been archived or studied from a computational perspective. While the classical music of today originates in the chanting of the Sama Veda of 1000 B.C., it has the unique distinction today of being a living tradition with a vital role in the religious, social and artistic lives of people. The number of annual festivals that feature performances by artists across generations as well as events where musicians and scholars interact attest to the highly visible musical activity. Central to the theory and practice is the distinctive musical concept of raga that seeks to capture the relationship between music and meaning. True to the characteristic of an oral tradition, the music pedagogy is not regimented but depends to a great extent on the process of transmission of knowledge and skill through prolonged exposure to the master. Analytical studies on Indian music can therefore have a particularly significant impact in the larger dissemination of knowledge and understanding. Arguably, the most important component in the appreciation of a new type of music is its structure and how it is built up from the more basic musical elements of melody and rhythm including the high-level relationships between these. Further, any study of music is incomplete without also studying aspects of music making. Indian musical instruments have some fascinating differences from Western instruments of the corresponding category. Coupled with the fact that Indian musical instruments are manufactured by local artisans using expertise that is passed down, an understanding of the construction and relation to sound quality can help us appreciate the nuances and perhaps place generational knowledge on more firm footing.

Recognizing the benefits of using computers in musicology, global efforts early on were directed to the task of music encoding or achieving machine-readable representations of written Western music compositions. This facilitated interesting studies on sizable corpora, such as those linking the origin of musical modes to specific geographical regions via the analyses of folk songs. More recently, research interest has also been directed beyond musical scores, to the analysis of actual performance recordings. Exploiting the outcomes, is the development of intelligent user interfaces for music browsing and navigation. We also see several successful software applications of performance evaluation in music learning.

This book is an outcome of a collaborative project involving researchers from institutes in India and well-known global experts as reflected in the list of authors. The project, sponsored by the MHRD, Government of India, funded SPARC scheme, has investigators who are primarily from the field of Music Information Retrieval, dedicated to developing tools for accessing music and computational models to explain music concepts. Based on their previous experience with speech signal processing, the work started with building melody extraction software for the ornamented style of vocal music that dominates Indian classical and popular genres. The effort received a significant impetus through CompMusic, an ambitious European Research Council project initiated by the Music Technology Group at UPF Barcelona to develop a platform comprising multicultural corpora accompanied by the appropriate software tools based on audio signal processing, machine learning, and semantic web.

This monograph is an attempt to bring together various aspects of studies on Indian art music that can benefit from audio processing and computing, ranging from musicology to retrieval and instrument modeling. To make an obvious point, music concepts and processes need to be understood in order to apply computational methods effectively. This makes the topic necessarily highly interdisciplinary and an ideal testing ground for collaboration involving the Humanities and technology. Exploiting the potential of MIR methods for the extraction of high-level information in the context of Indian classical music can also help us eventually progress towards the development of tools to enhance musical creativity. We are particularly interested in audio based analyses given that empirical research in Indian classical music must necessarily engage with recordings of performances, rather than with written scores. The tools of music information retrieval, of which audio source separation and multipitch tracking are well-known examples, can be useful in obtaining high-level musical descriptions that can serve to answer musicologically interesting questions around contemporary practices in raga music. There is also a growing emphasis on studying Western music forms as performed rather than as written.

What the chapters cover and what binds the different chapters together:

The book begins with a broad overview in the context of musicology research about music of the world and Indian music in perspective (Serra and Srinivasamurthy). Components critical to such research such as the availability of music corpora and the culture-specific audio computing tools are reviewed here. We next look more closely at Indian music in terms of its two equally sophisticated defining traditions, the Hindustani (North Indian) and Carnatic (South Indian) traditions with a historical perspective (Murthy). The scope of computational methods in adding to musicological knowledge is discussed for Hindustani music with illustrative examples drawn from studies of rhythm (Clayton). The chapter on Carnatic music (Venkataraghavan and Sridharan) paves the way for computational studies by laying out the basic framework of the genre in terms of the rhythm, melody and lyrics. We have next a fine example of the potential of computed singing voice frequency spectra to compare the signature singing styles of two legendary vocalists (Deshpande). Computational methods for the discovery of music structure are introduced next (Rao and Rohit). The application of rhythm-based structure analyses to musicological studies is demonstrated through the analysis of Dhrupad bandish sections (Clayton, Rao and Rohit). A comprehensive review of MIR work to date for Carnatic music is presented next (Ranjani et al.). The subsequent chapters are devoted to musical instruments. Starting with an introduction to the unique properties of Indian stringed instruments (Scavone, Patil and Prasanna), we move to a comparative study of Hindustani and Carnatic percussion instruments (Gowriprasad). This is followed by the presentation of an analytical account of the vibro-acoustics of the mridangam (Mohandas and Chandramouli). A pioneer in the field of Indian musical instrument synthesis reflects on his impressive journey from electronic synthesis of tanpura and taala to digital machines that have made their way to nearly every musician's practice room (Raj Narayan).

We hope this book serves to stir the interest of music scholars in corpus studies that can benefit from the use of the revolutionary digital tools. And that the examples presented here help them to frame their own musicological research problems that can benefit from collaboration with the computing community. Further, it has been encouraging to see the growth of research papers on raga music across fora inhabited by the signal processing and machine learning communities. It is hoped that the trend continues and grows in its sophistication with more vigorous collaboration between musicians, musicologists and engineers to enhance and deepen our understanding of the unique musical traditions of India.

Preeti Rao
Hema A. Murthy
March 2022

Contents

List of Figures

List of Tables

Part I

The First Part

Chapter 1

Getting Started on Computational Musicology and Music Information Research: An Indian Art Music Perspective

1

Ajay Srinivasamurthy[a,1], Sankalp Gulati[a,2], Rafael Caro Repetto[b], and Xavier Serra[a]

[a]Music Technology Group, Universitat Pompeu Fabra (UPF), Barcelona
[b]Institute of Ethnomusicology, Kunstuniversität Graz, Austria

Abstract

Culture-aware and culture-specific approaches to computational musicology and music information research (MIR) have been shown to be effective for analysis of a music culture. Recent efforts as a part of the CompMusic project argued that it is essential to consider the sociocultural specifics of a music tradition to effectively define research problems, collect data and propose methods for analysis. The project also demonstrated the use of such approaches, leading to a collective body of work for MIR in Indian art music. However, it is a considerable effort to define relevant research tasks, collect data and develop specific methods for analysis for each music culture, which often poses a significant entry barrier to start work in the field. One approach in such a scenario is to seek and identify parallel tasks, data and methods from the current state of the art in other music cultures and use them for a preliminary and basic analysis of culture-specific tasks, extending them with culture-specific methods to be more effective and relevant. While it is a sub-optimal compromise, such a perspective will enable preliminary analysis of a music culture using existing methods, and integrate it as a use-case with existing common frameworks and approaches in MIR and data-driven computational musicology. In this chapter, we aid such an approach by describing common concepts, frameworks, approaches, resources, data and methods for computational analysis of music from a perspective that could be useful for the analysis of Indian Art music. With this perspective, it is hypothesized that the currently established methods and tools, along with the datasets and corpora built within the CompMusic project will encourage accelerated research into different research problems relevant to Indian art music. The content of the chapter is targeted at musicians, music students, technology enthusiasts, engineers and researchers to provide them a context of the current state of the art that could help them start their work with computational analysis of Indian art music.

1.1 Introduction

, Recent advances in signal processing and machine learning have provided us with computational tools and methods for complex automatic analysis. In the process however, recent machine learning methods and efforts have shifted their focus on applications and tools, rather than the underlying data. Recent approaches consider data as one of the inputs and aim to learn any complex relationship as long as data is annotated and labeled with

[1]Ajay Srinivasamurthy is currently with Amazon Alexa, India and contributed to the work described in the chapter during his PhD at Music Technology Group, UPF, Barcelona before joining Amazon.

[2]Sankalp Gulati is currently with Eka.care, India and contributed to the work described in the chapter during his PhD at Music Technology Group, UPF, Barcelona.

useful metadata. Such an approach assumes that the research problems being addressed are well-defined, the data labels are accurate and objective, and the users (both producers and consumers) of such data and systems agree on the outputs of the system.

The availability of music in the form of audio enables us to take up a data-driven approach to Music Information Research and computational musicology. MIR aims to build automatic analysis methods to extract musically relevant events, metadata, labels and tags from audio. The applications include music discovery and enriched listening of music, complemented by metadata derived from the content of the music. Data-driven computational musicology aims to analyse large corpora of music by means of newly built computational tools. These tools for musicologists aim to complement manual analysis of carefully curated music pieces to larger corpora to yield valid statistical analyses that support and extend musicological hypotheses with quantitative data.

Music is a sociocultural phenomenon that poses limitations to a tool-driven approach by breaking some of the assumptions the existing modeling tools make. Any research problem posed on music needs to consider the sociocultural aspects of music to arrive at a well-defined problem that is musically relevant, grounded in theory and has a real world application for both producers and consumers of music. It is also necessary to understand the nuances of the problem and hence provide suitable approaches to solve them. Each music culture poses some unique problems that are not shared by others due to differences and nuances particular to the music culture. These differences are not deviations from a standard, but integral part of the music performance practice. A comprehensive analysis of music needs to be aware of such differences and nuances for a complete description of the music culture. Hence research problems need to be culture-specific or at least be culture-aware for a musically well-defined analysis of music. Even when the music concepts are well-defined in theory and well-agreed by experts, converting musical concepts to well-formed engineering definitions is a complex task.

While there is growing availability of music to be consumed on demand, most of that data is poorly labeled, or labeled only with editorial metadata, which is not useful for content based analysis. Manually annotating music audio with useful content metadata is an intensive process that needs music experts and involves a fair amount of subjectivity depending on the semantic level of music concept being annotated. Some basic music concepts are labels attached to a piece of music and hence clearly defined in the music theory of a music culture (at least in music traditions that have significant music theory literature), e.g. a song is in C Major or in rāg bhairav , while some concepts are complex descriptions and interpretations of music performers/listeners, which are hard to define and harder to codify into a common annotation schema, e.g., "a song is happy", "this song is the best example of rag bhairav". Automatic analysis methods hence tend to build approaches to analyze, extract and describe well-defined music concepts from audio. The lack of annotated data due to the effort involved, and a lack of agreement on certain subjective concepts in music make it especially hard to obtain well-labeled audio data for analysis. Beyond the approaches that only need some basic metadata, any meaningful analysis in MIR will need to be considered as a small-data supervised learning problem at the moment.

It is with this context that culture-aware and culture-specific approaches to building data, tools and methods have been proposed in the past to work on musically relevant research tasks with the involvement of the community that produces and consumes the specific music culture. Such approaches to computational musicology and music information research (MIR) have also been shown to be effective for analysis of a music culture. A clear case for culture-specific methods is the CompMusic (Computational Models for the Discovery of the World's Music) project [Ser11], which was a research project funded by the European Research Council from 2011 to 2017 and coordinated by Xavier Serra from the Music Technology Group of the Universitat Pompeu Fabra in Barcelona (Spain). It aimed to advance in the automatic description of music by emphasizing cultural specificity, carrying research within the field of music information processing with a domain knowledge approach. The project focused on five music traditions of the world: Hindustani (North India), Carnatic (South

India), Turkish makam (Turkey), Arab-Andalusian (Maghreb, though the project focused on the Moroccan tradition), and jingju (China).

CompMusic aimed to contribute to our multi-cultural society and focused on the advancement in the field of Music Information Research by approaching a number of current research challenges from a culture-specific perspective. CompMusic focused on the extraction of features from audio music recordings related to melody and rhythm, and on the semantic analysis of the contextual information of those recordings. The goal was to characterize culture-specific musical facets of each repertoire and to develop musically meaningful similarity measures with them.

Taking a data-driven approach to research, a major effort of CompMusic was to build a research corpus for each music tradition. The types of data gathered are mainly audio recordings and editorial metadata, which are then complemented with descriptive information about the items we have, and in some cases with music scores and/or lyrics. CompMusic focused on open research, with publications, code, and data available under open licenses. The project clearly demonstrated the use of culture-specific methods leading to a collective body of work for MIR in the music traditions considered within the project.

However, it is a considerable effort to define relevant research tasks, collect data and develop specific methods for analysis for each music culture. This often poses a significant entry barrier to start work on a music culture, especially in the absence of dedicated efforts like CompMusic. Even within the music cultures studied under CompMusic, only a handful of research problems could be solved within the duration of the project, while many identified research problems could not be addressed due to lack of resources.

A second approach in such a scenario for under-resourced music cultures is to seek and identify parallel concepts, tasks, data and methods from the current state of the art in other music cultures. Such parallels can be used for a preliminary and basic analysis of culture-specific tasks, which could then be extended to culture-specific methods to be more effective and relevant. While it seems like a sub-optimal compromise, such a perspective will reduce the entry barrier to define and take up research tasks in music cultures that have not been studied due to lack of resources. It will enable a preliminary analysis of a music culture using existing methods, identifying new challenges and limitations of current methods. It also puts the music culture into perspective within a larger body of MIR work, integrating it further as a use-case with existing common frameworks and approaches in MIR and data-driven computational musicology.

In this chapter, we take the second approach as applied to MIR and data-driven computational musicology tasks for Indian Art Music (IAM). We rely on our experience from the CompMusic project, where we build several culture-specific methods for the analysis of Indian art music. We identify parallel concepts and describe the available tools, techniques, frameworks and methods from the current state of the art for analysis of Indian art music. We also describe how in this process, Indian art music could be a use-case to test common frameworks and approaches in MIR. We aim to provide a series of perspectives on current approaches and methods that can be useful for analysis of Indian art music, while utilizing corpora and datasets that have been built for different MIR tasks in Indian art music.

Through this chapter, we wish to encourage MIR in Indian art music by providing parallel tasks and methods that could be used to bootstrap a new research task. The content of the chapter is targeted at musicians, music students, technology enthusiasts, engineers and researchers to provide them a context of the current state of the art that could help them start their work with computational analysis of Indian art music.

It is important to note that culture-aware approaches are necessary to consider the sociocultural context for any music analysis task, as argued within the CompMusic project. The approach we take in this chapter seems contrary to that argument, but we introduce parallel tasks and generic tools and methods only for basic analyses. The parallels we draw and general approaches we present are limited in their scope, only provide limited equivalency, and don't capture the nuances necessary for in-depth analysis. These approaches are mainly intended for new researchers to start with existing tools to arrive at a basic approach that can then be built upon to solve a musically relevant problem.

We start with a brief introduction to Indian art music, followed by describing some parallels in music concepts between Indian art music and other music cultures that can help to identify and define parallel research tasks. We then describe the audio signal characteristics of Indian art music and list multiple research tasks in MIR and data-driven computational musicology that can benefit from existing methods. We then introduce the corpora and datasets available for analysis of Indian art music, followed by different tools and frameworks that can be used for analysis.

1.2 Parallels in Musical Concepts and Dimensions

While the issue of the existence of universal concepts in all of the world's music traditions is a matter of strong academic discussion and a common agreement is still far to be reached, we focus in this chapter on three musical dimensions, such as melody (including by extension harmony), rhythm and timbre. We further note that these dimensions are not independent and orthogonal, but a significant interplay exists across the dimensions that come together in music performance. Another exercise of identifying parallels is based on functional aspects of music, looking at how music is composed, learnt, performed and produced, and the tools/methods used therein. We will apply both these approaches in the chapter.

Since the focus of the chapter is to identify parallel tasks and tools useful for IAM analysis, it is necessary to find parallels with concepts that have extensive literature in the current state of the art. The current state of the art in MIR focuses mostly on Eurogenetic music, hence we look for parallels to analytical categories developed in European music theory and performance. This would span an entire gamut of musics commonly called western popular and classical music.

We reiterate that the parallel concepts identified here are not interchangeably equivalent due to the differences in nuances. The definitions across music culture extend only to a limited extent. In the process of describing parallels, we also identify some of the limitations of those parallel concepts, which will help us to identify the limitations of the methods that do an automatic analysis of these concepts. We start with a brief introduction to the main music concepts in Indian art music.

1.2.1 Introduction to Indian Art Music

A detailed introduction to Indian art music (IAM) is beyond the scope of the chapter [Bag98a; Sam98b]. However, we provide a gentle introduction to concepts that are necessary for the chapter. In the context of this chapter, IAM refers to two art music traditions of the Indian subcontinent: Carnatic music, widespread in the southern regions of the Indian subcontinent (South India and Sri Lanka) and Hindustani music (also known as North Indian Classical music) prominent in the northern and central regions of India, Pakistan, Nepal, Afghanistan and Bangladesh. IAM today is a confluence resulting from cultural interactions between the Persian, Greek, Arabic, and Indian cultures [Sar11] and is a well studied music tradition with sophisticated and grounded music theory. It has a large audience, continues to evolve in current sociocultural context and has attracted high interest from music scholarship. The presence of a large dedicated audience and research literature forms a solid basis for studying this music culture from both a musicological and computational perspective.

Both Hindustani and Carnatic music are performance oriented, heterophonic, with a main melody being sung or played by the lead artist, and mainly improvisatory in nature. Vocal music is dominant in both traditions but more in Carnatic music. A typical arrangement in a performance of IAM consists of a lead performer, a rhythm accompaniment (typically mridangam and tablā/pakhāvaj in Carnatic and Hindustani music, respectively), a constant sounding drone in the background by the tānpurā and instrumental melodic accompaniment (typically violin and harmonium or sāraṅgī in Carnatic and Hindustani music, respectively). In both traditions, rāga is the melodic framework and tāla is the rhythmic framework. Functional harmony, as defined for eurogenetic musics is not used in Indian art music.

Rāga (melodic framework) and tāla (rhythmic framework) concepts have been largely discussed in a multitude of studies and will not be reviewed in detail here [Bag98a; Sam98b], but it is emphasized that these concepts are fundamental in analysis of Indian art music.

IAM is taught orally through a lineage of teachers and students. While a skeletal prescriptive notation based on an Indian solfège (called the *sargam*) exists, it is not standardized and mainly serves as a mnemonic aid for melodies and lyrics, and is not directly used during a concert performance. As a result, there are recognizable stylistic differences between different interpretations of a composition, melody or a musical concept.

Further detailed descriptions of these music concepts are presented in other chapters of the monograph. We focus only on drawing parallels in the rest of the section, from the sole perspective of noting the possible extensions and limitations of these parallels to Indian art music.

1.2.2 Melody

A musical note, or a svara in IAM is a fundamental concept around melody. The idea of a svara is similar to the eurogenetic definition of a note, but differs in some aspects. The svara in IAM do not rely on absolute pitch positions, but are defined relative to a reference pitch frequency (called the ṣaḍja or ādhāra shadja, referred henceforth as tonic frequency) that is arbitrarily (or better described as conveniently) by the artists in a performance. All pitch positions and music scales are defined based on this reference pitch, making it a fundamental quantity to be estimated for any melodic analysis. The tonic is not based on a tuning reference unlike popular eurogenetic music (e.g. A4 = 440Hz), but rather based on the convenience of the lead artist in a performance.

Given a tonic, the pitch position of a note or a svara is defined based on the tuning system used. It's widely accepted that IAM uses just intonation, with all pitch positions defined in relation to the tonic frequency [KI12a]. However, a svara in IAM does not refer only to a discrete pitch position, but also includes its intonation, which broadly describes how the note is to be performed. Loosely, a svara hence could be explained as a pitch position relative to the tonic, along with the transitions coming into and leaving the pitch position, including any additional ornaments while performing the svara. The transitions and the ornaments are not a characteristic of a svara in isolation, but are typically a property of the svara within a specific rāga. To summarize, a svara is not a frequency value () in IAM, but a contour of frequency values.

Ganguli and Rao [GR18] summarize the rāga as being somewhere between a scale and tune, providing a grammar that specifies tonal material, tonal hierarchy and a set of characteristic melodic phrases. Every rāga has a set of characteristic melodic phrases that act as building blocks to construct melodies and provide a base for artists to express their creativity through improvisation within the rāga grammar. Musicians and musicologists often consider that a rāga can only be learned by getting familiar with landmark compositions (hence the typical phrases) in it. A typical concert starts with the rendition of characteristic phrases of the rāga being performed. Characteristic melodic phrases are also the most prominent cues used by human listeners for identifying rāgas [KI12a]. Due to the ornamentations and continuous note transitions, a melody in IAM is better represented as a pitch contour rather than an abstract sequence of notes that it represents. Given the nature of a svara, commonly interpreted as a melodic contour and not a discrete pitch, the segmentation of melodies in IAM into its component svara is not straightforward (even for humans).

1.2.2.1 Melodic Patterns

Melodic patterns, in their simplest form, are sequences of notes along with their ornaments that frequently repeat in performance. These patterns could be characteristic of a rāga or of a composition. If melodies are represented with pitch contours, melodic patterns would be short segments of those contours, along with dynamics information if available.

Based on these melodic parallels, we identify that tonic recognition, predominant melody tracking, extracting and characterizing melodic patterns to be fundamental melodic analysis problems in IAM.

Table 1.1: Musical parallels in rhythm and meter

Eurogenetic music	Carnatic Music	Hindustani Music
tatum	akṣara	mātrā or sub-mātrā
tactus/beat	beat or aṅga	mātrā or vibhāg
measure/bar (downbeat)	āvartana (sama)	āvart (sam)

Transcription of an audio recording into notes hence is not a meaningful task for IAM - note sequences are only skeletal and prescriptive, music itself is highly improvised, and ornaments make it difficult to summarize a melodic phrase just by a sequence of notes.

1.2.3 Rhythm and Meter

Eurogenetic music typically identifies three metrical levels for organization of rhythmic events in music: tatum, tactus/beat, measure/bar. Tatum is the smallest atomic time interval in a music piece or the fastest pulse present in a music piece. Tactus (or beat) is typically referred to as the pulse that listeners entrain to as they tap their foot or dance along with a piece of music [Han89]. A bar/measure is a group of beats to indicate grouping and a periodic accent in the music. The first beat of a bar is called a downbeat and is often associated with significant rhythmic events. A time-signature defines the structure of a bar and the number of beats it comprises. This hierarchy of different metrical levels helps to organize rhythmic patterns within a music piece.

Rhythmic organization in IAM relies on the tāla framework, which comprises fixed-length hierarchical time cycles. The hierarchical time cycles (cycles of a tāla, or tāla cycles in short) provide a structure for rendition and repetitions of melodic and rhythmic phrases and patterns. The tempo of a music piece (measured as the time difference between the start of two successive metrical cycles) is independent of tāla, and is often independently chosen by the musician based on the composition and other considerations.

The hierarchy of events in a cycle is provided through sub-structures defined within a tāla cycle that help to track progression through the cycle. These sub-structures help to identify the different points of time in the cycle and hence aid in multiple tasks, such as tracking musical time intervals in a musical piece, synchronize the performances by different artists in a concert, demonstrate progression through the cycles (either with hand gestures in Carnatic music, or through canonical rhythmic patterns indicative of a position in the cycle). While the sub-structures are well-defined to provide a clear framework, the musicians are free to improvise within that framework.

Typically three different sub-structures are identified: sub-divisions, beats and a tāla cycle, often with parallels drawn to tatum, tactus and measure. However, there are clear differences in both Carnatic and Hindustani music with such parallels, which will be elaborate further. The equivalence of a tatum-tactus-bar metrical hierarchy to that in IAM is not an exact exercise, but is useful since we could build on methods to analyze the rhythmic events associated with those metrical levels and apply them to analyze similar structures in IAM. Broadly speaking, tatum, tactus and bar are often analyzed with onset detectors, beat trackers and downbeat trackers and drawing possible parallels will help bootstrap these algorithms for metrical analysis of IAM.

For drawing parallels, we use the definitions that tatum is the shortest pulse tracked in the meter, tactus is the "foot-tapping" pulse, and bar as a group of beats identified with a downbeat that is the first beat of the bar. This definition has an implicit assumption that a tactus is a pulse that listeners entrain to, and hence the span of such a pulse is often associated with working memory [HS11]. The definition of a tactus also assumes isochronicity, which is not a necessity in IAM [SHS14].

The metrical hierarchy in IAM is identified through the sub-structures of a tāla in the form of mātrā/akṣara, beats/sections and āvartana (cycle). As we see from Table 1.1, the parallels in IAM to the corresponding eurogenetic entities are not absolute or one-to-one, but depend on the tempo of the music piece and the structure of the tāla.

The parallel to a bar/measure is the tāla cycle (āvartana or āvart). The downbeat, or the first pulse of a bar is known to be significant. Similarly, the instant of beginning of each tāla cycle (also the end of the previous) is referred as sama (or sam in Hindustani music) and is highly significant structurally, marking time boundaries of important melodic and rhythmic events. The sam frequently marks the coming together of the rhythmic streams of soloist and accompanist, and the resolution point for rhythmic tension [Cla00a, p. 81]. A tāla cycle is further divided into sections and beats.

Carnatic music defines the smallest pulse as akṣara (a syllable), which are grouped into beats, which are further grouped into sections (aṅga) that form an āvartana. The number of akṣara in a beat is based on tempo classes and the sub-division structure of the tāla. The sections of a tāla are typically unequal in length and progression through the sections are shown through hand gestures indicative of different types of sections. The beats of a tāla are isochronous, but sections are not. Based on the tāla structure, the tactus pulse is perceived at the beat level (isochronous) or at the aṅga level (typically, non-isochronous). E.g. the tactus of adi tāla is at the beat level, with 8 isochronous beats in a cycle (with 3 sections), while the tactus in mishra chapu tāla is perceived at the section level with beats grouped into 3+2+2 structure in a 7 beat cycle.

Hindustani music also defines hierarchy through mātrā, vibhāg and āvart. A mātrā is the smallest rhythmic sub-unit of a tal. Mātrā are isochronous and are grouped into possibly unequal sections (called vibhāg) that make up the whole āvart (cycle). There are also tempo classes called lay in Hindustani music which can vary between ati-vilaṁbit (very slow), vilaṁbit (slow), madhya (medium), drut (fast) to ati-drut (very fast). Depending on the lay, the mātrā may be further subdivided into shorter time-span pulsations, indicated through additional filler strokes of the tabla. However, since these pulses are not well defined in music theory, we consider mātrā to be the lowest level pulse as a parallel to the tatum. However, depending on lay, the mātrā duration can vary over a wide range of 200ms to 5 seconds. This means that the tactus pulse that listeners are entrained can align with the sub-mātrā fillers (in vilambit lay), or mātrā (in madhya lay), or with even vibhāg (in drut lay).

To summarize, the parallels to tatum-tactus-bar metrical levels using their common definitions does not directly extend to IAM though some parallels exist. Due to the structure of the tāla and tempo classes in IAM, tactus pulse could be non-isochronous and could align with different metrical levels defined in IAM. The common definition of a beat assumes isochronicity and hence does not extend directly to IAM, which can have a non-isochronous pulse to be tracked as a beat. However, isochronous pulse trackers (beat trackers, e.g.) from eurogenetic music can be bootstrapped to track the isochronous Hindustani mātrā or Carnatic beat but in an typically extended range of tempi, and then grouped based on the knowledge of tāla structure to derive relevant non-isochronous tactus pulse.

1.2.3.1 Rhythmic patterns

Rhythmic patterns in IAM are closely aligned with the tāla cycles, and are played on melodic or percussion instruments. The rhythmic patterns serve multiple functions in addition to showcasing the rhythmic structure of the tāla and the composition. Within the scope of this chapter, we wish to draw attention to two specific kinds of rhythmic patterns, canonical rhythmic patterns played to indicate progression through a tāla cycle and different hierarchical metrical events with the cycle, and the improvisatory rhythmic patterns that showcase the variety of rhythmic patterns that could be played within the framework of a specific tāla.

The hierarchical metrical structure of the tāla is indicated through canonical rhythmic patterns that also indicate progression through the cycle. These canonical rhythmic patterns are well-defined in hindustani music (called the theka) and played on the tabla, while they are less well-defined in Carnatic music. The canonical patterns aim to establish the fundamental metrical framework of the tāla through instances of characteristic rhythmic patterns and hence could be used by all musicians to perform together within the tāla framework. The function of these patterns is predominantly to provide the metrical structure for performance.

Improvisatory rhythmic patterns played during improvisatory sections help to showcase the variety of rhythmic patterns that could be played within the framework of a specific tāla. These patterns are typically played during percussion solos (or even the solo performance of a melodic instrument or vocals). The improvisatory patterns focus on variety and involve complex calculations that can be accommodated within the tāla, deviating from canonical structures and provide ample opportunity for demonstrating skill and expertise in rhythm. Improvisatory patterns can often last multiple tāla cycles with long and complex patterns that finally resolve at the specific point in the tāla (most often, the sama).

1.2.3.2 Percussion patterns

A distinctive feature of IAM relates to the use of percussion patterns, which are mostly the patterns intended to be played on percussion instruments of IAM. They are distnct from rhythm patterns, which mainly refer to the temporal arrangement of different events with different accents, while percussion patterns include a temporal arrangement of different percussion timbres.

Percussion patterns in IAM are described and learnt through syllabic onomatopoeic oral mnemonics that relate to the sound of the strokes played on the mridangam/tabla. Most of the time, the percussion syllables also encode the dynamics, intonation and interaction between neighbouring strokes. Hence, as an analogy, the spoken percussion syllables (called the bōl in Hindustani music and solkaṭṭu in Carnatic music) form a language for percussion in IAM and the art of reciting these syllables is often a feature of many solo music and dance performances. Indian art dance forms such as Kathak and Bharatanātyam make extensive use of percussion syllables to define and perform movement patterns. Percussion training extensively uses these syllables to learn and perform percussion patterns on instruments. The analogy of the percussion patterns to speech enables us to utilize a multitude of speech technologies for analysis of percussion patterns in Indian art music.

While percussion instruments used in eurogenetic music could also have their strokes mapped to oral mnemonic syllables, IAM provides an existing system that is familiar to all practitioners. Further, the rich and complex set of timbres played on tabla/mridangam can be mapped to different bōl/solkaṭṭu used in IAM and used for representation, analysis and synthesis of percussion patterns. While the exact set of percussion syllables used in practice is not standardized and varies across music schools and musicians, it is still possible to map timbres to syllables and use them for automatic analysis in transcription problems. These characteristics provide a unique opportunity to study and analyze percussion as a language in IAM.

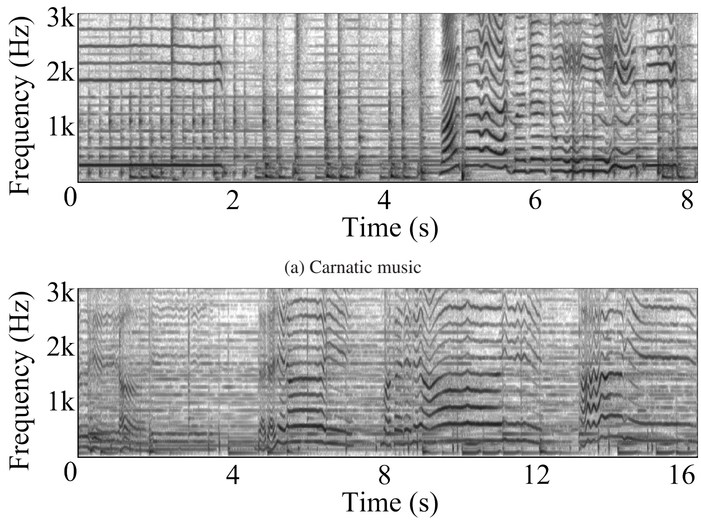
1.2.4 Timbre and Instrumentation

The timbre and instrumentation in IAM is useful to understand and contrast to help with choice of tools for signal processing. We first describe the instrumentation and timbre from a musical perspective and then describe how this manifests in audio signals.

IAM performances typically have limited number of performers, with a lead musician (often vocals), a melodic accompaniment (mostly violin in Carnatic and harmonium/sāraṅgī in Hindustani music) and a (set of) percussion accompaniments (tabla in Hindustani and mridangam in Carnatic music, respectively).

This makes Indian art music predominantly melodic and heterophonic, with usually two (sometimes more) simultaneous melodic voices, leading to two or more overlapping melodic lines. Furthermore, a tānpura (drone) playing in the background reinforces the tonic frequency and often the fifth. All melodic instruments are tuned to the tonic of the lead artist. The primary percussion accompaniments are pitched, and are also tuned to the tonic of the performer. There is no polyphony in the sense of the word defined in eurogenetic music.

The drone in the background could be used for analysis of tonic. The heterophonic nature of music makes analysis of melody relatively simpler, but it is still complicated due to the melodic accompaniment closely following the lead melody with variations with a small delay. The pitched percussion also often overlaps in pitch with the lead instrument, making the task of percussion separation harder.



(a) Carnatic music

(b) Hindustani music

Figure 1.1: Audio signal characteristics of Indian art music signals. The figure shows the spectrogram of the audio excerpt of Carnatic and Hindustani music, showing the frequencies up to 3 kHz.

1.2.4.1 Audio and Signal Characteristics

A piece of music in an audio recording is a primary artifact for analysis of Indian art music. Hence it is useful to identify the characteristics of IAM from a signal processing perspective and describe how different IAM concepts manifest in audio. This will enable us to choose pre-processing steps and feature extraction for many MIR tasks, as well as aid in source separation tasks. Several basic MIR tasks that extract melodic and rhythmic events can benefit from generic signal processing approaches by mapping the music concepts to signal processing terminology.

Figure 1.1 shows an illustrative example of the spectrogram of audio excerpts of both Carnatic music and Hindustani music. To focus on melody and percussion, the spectrogram is plotted only up to a frequency of 3 kHz, though there is information from higher harmonics and percussion strokes at higher frequencies. The time durations of the excerpts are also different, with the Carnatic music example being shorter in duration.

The lead melody and its harmonics can be clearly seen in both the figures, along with the continuously changing fundamental frequency (F0). There is a drone in the background that provides the tonic for the performance, provided by the tānpura (Hindustani music) or tambūra (Carnatic music). The drone can also be seen as the unchanging set of spectral frequencies in both the spectrograms. The melodic accompaniment (violin) in Carnatic music closely follows the lead voice and can be seen with a lower amplitude. Harmonium is the melodic accompaniment in the Hindustani music excerpt, which can also be seen with a lower amplitude in the figure. These observations also have an important bearing on melodic analysis of IAM. Due to the ornamentations and typical intonation of svara, pitch contours are more relevant melodic representations of melody in IAM than a sequence of notes. Even melodic phrases are better encoded as pitch contours rather than a short sequence of svara. Segmentation of a melody into its svaras, understood as melodic contour, could be interesting to observe, for example, how the same svara is performed in relation to its melodic context, or in different rāga.

The percussion instruments tabla and mridangam have a bass drum head, and a treble drum head that is pitched. The pitched drum head is tuned to the tonic of the lead musician. The pitched strokes can be sharp or sustained. Both the figures show the percussion strokes as vertical lines, showing their broad spectrum. We can identify the strokes from the left and right drums distinctly in different frequency ranges in both cases. In addition, we can see some of the harmonics of the pitched percussion strokes, and the quick decay of the percussion strokes. The Hindustani excerpt is taken from a madhya lay piece and we can see the longer notes and sparser tabla strokes, indicating lower rhythmic density due to lower tempo.

1.2.5 Functional parallels

The functional parallels that we discuss pertain to the different functions around Indian art music, mostly discussing the composition, performance, learning, production and consumption. These characteristics primarily help us to put the important facets of IAM into perspective and hence identify and prioritize research problems, choose artifacts, build analysis tools that are musically relevant and meaningful within the IAM community.

1.2.5.1 Music learning

Music learning and teaching is predominantly oral, learnt through repetition and following a teacher. Music scores in sargam notation are used for learning, primarily as a mnemonic aid for learning and and not as an exhaustive representation of the melody to be performed. The oral transmission also leads to stylistic schools based on lineages of masters or geographic regions. IAM percussion has corresponding onomatopoeic oral mnemonic syllables that represent individual strokes/timbres that can be played on a mridangam/tabla. These syllables are used for teaching/learning as well as a form of art in performances. As discussed earlier, availability of syllabic representation means clear parallels to mature speech technologies that can be used for analysis of IAM.

1.2.5.2 Music composition and performance

IAM performances include both compositions and improvised pieces. However, even the performance of a pre-composed music piece includes significant improvisation within the framework of the rāga and tāla of the piece. The music compositions provide the lyrics (for vocal performances) and skeletal score (mostly in the form of sargam) that is mostly prescriptive and improvised upon. This primarily implies that analysis of audio performance is more relevant for IAM than analysis of music scores. While music scores do contain useful information, their use is limited in music learning and performance, and hence in detailed analysis due to extensive improvisation. This also implies that an audio recording along with its metadata is the primary artifact for analysis of IAM.

1.2.5.3 Music consumption, appreciation and critique

Historically, IAM was consumed by listeners through live concerts and more recently through recordings of live concerts. The consumption platforms are changing slowly to digital platforms as is the trend with most music recently. The current practice is that most of the concerts in IAM are organized by music foundations and organizations throughout the year, with specific music festivals with a multi-fold activity in concerts. A few example festivals include the Madras music festival season (Chennai, India) and the Ramanavami music season (Bangalore, India) in Carnatic Music and ITC SRA Sangeet Sammelan, Sawai Gandharva Bhimsen Festival (Pune, India) in Hindustani music. Both these music traditions have a dedicated audience spanning different demographics and expertises. Outside of the concerts, information around IAM is consumed through listener music fora, informal discussions and media critiques/reviews.

The involvement of a large community of people in consumption and appreciation of IAM implies that a large amount of textual information (in the form of news articles, forum discussions, reviews, critiques) is available in addition to formal literature and music metadata. The music concepts are shared across the corpora for analysis, formal musicological and music theory literature, and informal discussion around music. This provides ample opportunities for research into organizing text information around IAM through existing linked data and semantic web approaches.

1.2.5.4 Music production and distribution

A basic unit of distribution of a recorded Carnatic music performance is a concert that consists of several pieces, each in a specific tāla and rāga. The concert has a structure with compositions and improvisations performed in specific forms of Carnatic music in a predictable order. Hindustani music albums comprise one or two longer pieces, in a specific rāga followed by short pieces. A structure hence exists in both Carnatic and Hindustani music, unlike an album of popular music, which mostly has independent tracks. Segmentation of a concert into relevant individual pieces (or into shorter sections within a music piece) is hence a relevant problem for IAM. IAM corpora are hence better organized in such units (concerts) with segmented audio with associated metadata for each music piece. In addition, the corpora could have additional metadata at the release or concert level.

Since most audio recordings available in IAM are recorded from live concerts, the quality of audio can vary considerably, especially with older concert recordings. Given the limited instrumentation in IAM, it is possible to record and store multi-track recordings, which are a very useful source for multiple MIR tasks including melodic and rhythmic analysis, source separation and music transcription. An effort to collect multi-track corpus is described in Section 1.4 of this chapter.

1.2.6 Indian art music: Important metadata

The previous sections drew parallels on multiple aspects of Indian Art Music. We summarize the learnings in this section. The primary object of analysis is an audio recording of a music performance and associated metadata. The rāga and tāla related metadata comprise the most crucial metadata for a piece of music.

Within the frameworks of rāga and tāla, we have analysis of certain concepts that could benefit from existing engineering formulations and hence can be explored as direct extensions to current methods. However, many concepts are socioculturally distinct and pose additional challenges to engineering formulations for their analysis. Key examples include the presence of improvised sections, the absence of a fixed tonic, non-isochronous beats, hierarchical metrical structures with long cycles, melodic representation through pitch contours, absence (or low utility) of written music scores, the presence of syllabic percussion systems, orally transmitted repertoire, and unique aspects of music production consumption and appreciation. With these distinctions, we can now discuss the research problems around MIR and computational musicology from an IAM perspective.

The remainder of the chapter will focus on tasks, data and tools, mostly from the perspective of getting started on MIR and computational musicology for IAM - tasks that could use existing tools for initial analysis and availability of data for analysis.

1.3 Research Problems in Indian Art Music

A detailed list of MIR problems relevant for IAM along with current approaches is presented in other chapters of the monograph. Hence, we won’t formulate the tasks in detail and review existing methods though we provide any broad challenges and directions that have been already explored. The research problems discussed are broadly divided into MIR problems and data-driven musicology problems. The MIR problems aim to analyze and extract musically relevant metadata and descriptors from audio, scores and text. The research problems on data-driven musicology aim to provide tools, methods and musicological insights on large corpora of music data.

1.3.1 Music Information Research (MIR) Problems

The MIR problems consider one of musical artifacts as inputs and analyze it for musically relevant metadata and automatic annotations. Since an audio recording is the primary artifact for analysis of IAM, most relevant MIR problems focus on audio, while we also discuss in brief MIR problems that focus on prescriptive music scores and text. We categorize the problems based on the artifact consumed and the musical concept the problem addresses.

1.3.1.1 Melodic analysis

Considering the rāga as the central element, MIR problems on melodic analysis for IAM include tonic and predominant melody extraction, melodic transcription and the analysis of melodic patterns [Gul16]. Though the concept of a rāga is central to IAM, there is not an equivalent concept that has been explored within previous MIR literature. While modal analysis and melodic analysis of Turkish makams could be considered parallels, most of the melodic analysis work that considers rāga as a central concept are limited to IAM. The melodic analysis problems revolve around the concept of rāga, either contributing to its recognition through other low level melodic descriptors, or its characterization through higher musical abstractions such as melodic scales, phrases and intonation.

Predominant melody extraction is arguably the most useful melodic analysis problem given that the predominant melody is the input to many different melodic analysis tasks. Given the heterophonic nature of IAM, predominant melody estimation is less challenging than polyphonic music, though the possibility of two closely aligned melodic lines (vocal+violin for CM and vocal+harmonium/sarangi) can often pose challenges. A predominant melody in IAM is arguably best represented as a pitch contour (continuous valued) computed at each audio frame (typically a few ms), and then quoted either in absolute frequency (Hz) or normalized by tonic (and represented in cents relative to the tonic). The pitch contour can then be used downstream for all subsequent tasks. Existing methods for predominant melody extraction, e.g. melodia [SG12b] have been experimented to work with acceptable accuracy with some tuning for the task in IAM, and hence can be good initial solution for culture-specific methods if need to be developed. However, predominant melody estimation for IAM is not a completely solved problem and is still a challenging task with scope for improvement and needs further exploration.

Given the primary importance of the shadja (tonic note frequency) in IAM melody, estimation of its frequency in an audio recording of IAM is a fundamental problem in melodic analysis. Since the frequency of the reference pitch (tonic) depends on the specific needs of each leading artist who choose it, the estimation of tonic in IAM is not so much of a tonic note estimation but a tonic frequency estimation. Tonic can be reliably estimated with current methods in clean high quality recordings, with most estimation errors corresponding to estimating a related note frequency (a fifth or an octave, e.g.) as the tonic. Existing methods for tonic estimation broadly use a predominant melody extractor (e.g. melodia) and a representation that summarizes the melodic content (e.g. a pitch histrogram). From such a representation, based on svara relationships, we can extract the tonic frequency [Gul11].

Intonation analysis refers to an analysis of expression of different notes of a music piece. Intonation captures the positions (on a frequency scale) of different notes in the music piece, while also providing additional information on how those notes were performed. Given the improvisatory nature of IAM with a plethora of melodic ornamentation, intonation provides a snapshot of how specific svaras are performed within a rāga. The simplest representation of intonation in IAM is the histogram of all pitch content (predominant melody contour, e.g.) in the music piece. Multiple existing methods can be used to parametrize the histogram to arrive at useful models for intonation analysis [Kod+14a].

Analysis of melodic phrases aim to characterize sequences of svaras (encoded often as short continuous pitch contours) and infer similarity relationships between them. A music piece in IAM typically consists of composition-specific music phrases and rāga-specific music phrases, both performed within the framework of a rāga. Defining similarity metrics between melodic phrases help us to cluster melodic patterns, which can eventually help to compute similarity between two sections of single music piece (intra-pice) or between two music pieces (inter-piece) in a large music collection. Melodic phrases are fundamental hence to organize a music collection based on melodic similarity. Analysis and clustering of melodic phrases also help us to identify the rāga of a music piece and help to provide a musically meaningful summary of the music piece.

When using predominant melody representation using pitch contours, analysis of melodic phrases can be formulated as repeating sequence matching problem in a 1-D data, though time normalization using Dynamic Time Warping might be necessary[Gul16] to compute distance between different pattern instances. Once we have these distance computed, we could use multiple available clustering methods to group phrases and retrievs them.

Automatically identifying the rāga of a music piece is an important MIR task in IAM, helping us to organize large music collections on relevant metadata. Depending on the application, it could be formulated as a recognition task (no prior information or hypotheses on the rāga) or a verification task (where a rāga label is presented with the audio recording and the goal is to verify if the label is correct). rāga related information is present in multiple melodic descriptors we can extract from music recordings, and hence different methods have been employed for rāga recognition. While characteristic melodic phrases are often the best indicators of a rāga, a combination of intonation, melodic phrases and an alignment of these descriptors with rāga grammars from music theory is expected to provide the best performance for rāga recognition.

1.3.1.2 Rhythmic analysis

Automatic rhythm analysis problems in IAM could be broadly categorized into problems of automatic rhythm annotation and analysis of rhythm patterns. Automatic rhythm annotation for eurogenetic music could be defined as estimating different elements of meter from an audio recording - with onset detection [Bel+05], tempo estimation [PP11], beat tracking [DP07] and downbeat tracking [DP06; HDF12; KBW15] as primary explored problems.

Automatic rhythm annotation, in the context of Indian art music can be defined as the estimation of the characteristic components of the tāla from audio [Sri16]. For Carnatic music, the most important rhythm related tags include estimating the median tempo of the piece (in akṣara per minute or beats per minute), the length of the cycle (in number of beats or akṣara), the tāla label (and hence implicitly the underlying metrical structure) and the subdivision structure. For Hindustani music, the most important rhythm related tags include the median tempo of the piece (in mātrā per minute), the lay class, the cycle length (in mātrās) and the taal label. Estimating the time varying tempo curve, the akṣara pulse locations, the beats, the tāla cycle section boundaries and the sama instants are the important time aligned annotation problems in Carnatic music. The most important problems in Hindustani music are the estimation of the mātrā pulsation, the vibhāg boundaries, and the sam instants.

Automatic rhythm annotation is an important rhythm analysis topic, and there are several applications in which these rhythm annotations are useful, such as music autotagging, rhythm based segmentation of audio, beat aligned processing of music, audio summarization, music transcription, and different rhythmic pattern analyses. Tracking the components of the tāla through a music piece is essential for most other rhythm description tasks such as segmentation and extraction of rhythmic patterns to define similarity. It is to be noted that many rhythm annotation problems can be jointly addressed, estimating several components together, e.g. the tāla, tempo, beats and the sama can be jointly estimated, in a task that we could be called as automatic meter analysis.

Automatic rhythm annotation could benefit from existing rhythm MIR tasks, though some modifications might be necessary. The equivalence of tactus to the mātrā in Hindustani music could be problematic since the typical durations of tactus used for beat tracking might not align with the mātrā in vilaṃbit pieces. Beat tracking methods often assume an isochronous beat - and as discussed before, beats in IAM are not isochronous. However, we could use a beat tracking algorithm to estimate an isochronous pulse in an IAM audio recording, and based on the structure of the tāla, we could pick a subset of the estimated isochronous pulse to arrive at the musically relevant beats for the tāla. Downbeat tracking in eurogenetic music has mostly been applied to short measure durations (a few seconds, typically). However, it has been shown that such methods do not extend well [SHS17] when cycle durations could be longer than a minute (e.g. in vilaṃbit pieces in Hindustani music).

While tāla provides a framework and structure, the rhythm and percussion patterns form the content through which the metrical structures and rhythms are illustrated, and hence form the other main component of rhythm analysis. Rhythm patterns mainly refer to the temporal arrangement of different events with different accents, while percussion patterns include a temporal arrangement of different percussion timbres. To contrast, percussion patterns are rhythmic patterns, but rhythmic patterns need not contain only percussion, and can be formed by melodic and/or percussion instruments.

A pattern is defined as a temporal sequence of events and hence it is necessary to estimate onsets of various instruments in music, since that creates the time-aligned sequence of note/stroke events that can be further used to obtain both rhythmic and percussion patterns. Some important sub-problems within pattern analysis are instrument-wise onset detection, pattern transcription, and pattern discovery, each of which is described further. Transcription aims to map an audio recording into a time aligned sequence of symbols (strokes, accents, e.g.). The problem of discovery is more open ended and aims to automatically retrieve interesting patterns and insights about those patterns, in a data-driven way.

The task of instrument-wise onset detection refers to detecting the onsets of specific instruments from an audio signal that is a mixture of many music instruments. For rhythm analysis in Indian art music, instrument-wise onset detection can help to obtain cues for both meter tracking and for analysis of percussion patterns. The onsets of percussion instruments mridangam and tabla provide cues to the tāla and delimit percussion patterns. A differentiation between the left (bass) and right drum onsets in both instruments is additionally insightful and useful. Instrument-specific onsets are often not estimated explicitly, but are estimated as a part of a bigger task, such as percussion transcription.

Rhythm patterns extracted from audio recordings are representative patterns of the tāla, and hence useful for both automatic tāla recognition and meter analysis. The most relevant rhythmic patterns are cycle length rhythmic patterns - patterns that are played in one full cycle of the tāla. Shorter patterns, played within a cycle mostly act as rhythmic atoms to make up the whole cycle and are played more often. However, there are long rhythmic patterns played on mridangam/tabla and accentuated through melody that can last many cycles. Automatic discovery of rhythm patterns can be used to define content based rhythmic similarity between pieces of music, which is expected to be more relevant than metadata based similarity. Automatic extraction of rhythm patterns can also be a tool for musicologists to study various rhythm patterns in larger corpora.

Percussion pattern transcription is mainly applied on audio recordings with percussion solo, and aims to transcribe the audio recording into a time-aligned sequence of drum stroke labels, and in the case of Indian art music, into percussion syllables. Percussion transcription is a sub-problem of the more general music transcription. Transcription of a solo into symbolic syllables is an example of audio segmentation at a fine grain level. Transcription of solo performances are useful for percussion training. Since Indian art music is mostly improvised, the need for such a fine grained transcription system is limited, except for music education and performance analysis applications. However, a transcription can be used to automatically discover percussion patterns and develop rhythm similarity measures using such discovered patterns that could be used for other tasks, e.g. recognition of tabla gharānā [Gow+21].

The benefits of using oral syllabic systems from an MIR perspective are both the cultural specificity of the approach and the accuracy of the representation of timbre, articulation and dynamics. The characterization of these percussion traditions need to consider elements that are essential to them such as the richness of their palettes of timbres, subtleties of articulation, and the different degrees and transitions of dynamics, all of which is accurately transmitted by the oral syllables.

As discussed earlier, the onomatopoeic percussion system in Indian art music provides a language for percussion and hence is the most musically meaningful way to represent percussion patterns of tabla and mridangam. Such a link between drum patterns and natural language has been explored [MD12]. Percussion pattern transcription can be formulated as a supervised learning task, using labeled training data to build syllable stroke models, which can then be used to transcribe a test recording [Ana+14; Gup+15; Kur+15]. Percussion pattern discovery is an unsupervised task, aiming to extract percussion patterns from audio and/or scores in an unsupervised way, though some priors can be used.

Rhythm similarity measures aim to use rhythm descriptors, metadata and segmentation information to provide an objective similarity value between two phrases, two music pieces, or two parts of the same piece. Since rhythmic similarity is not a very concrete notion, we need definitive and objective measures of similarity, especially in a multicultural setting. This would necessitate the use of knowledge based approaches for similarity modeling. The onus of developing new similarity measures clearly lie on the choice of metrics that correspond to rhythm similarity as perceived through musically relevant concepts - based on tāla and the rhythmic patterns.

1.3.1.3 Structural analysis

Structural analysis typically deal with the problem of segmentation, referring to a broad category of problems which involve the labeling segments of audio with a label/tag. Segmentation can be done at several levels, based on different music concepts. Segmentation problems are useful since they provide additional metadata to navigate through music collections (and within a single recording), and to further develop similarity measures.

Segmentation can be done of a full concert into the pieces that were performed in it, which is useful for archival. Segmentation at a structural level within a piece aims to segment the piece into different sections of the piece, and is useful for navigation and similarity. Different music forms in IAM are amenable to different structural segmentation, e.g. an ālāp would benefit from a segmentation based on melodic phrases, while a percussion solo could be segmented based on different sections and percussion patterns played within the solo. A music recording can also be segmented based on the instruments that are playing, a problem that can also be described as instrument tracking in audio. At a fine grained level, a music piece could also be segmented at the level of melodic and rhythmic phrases, which could be used to highlight different parts of a music piece.

A basic idea for segmentation of music pieces has been based on the assumption that change-points in different structural elements (e.g. sections) are accompanied by a change in music characteristics along different musical dimensions (melody, rhythm, timbre). This implies that change points could be detecting changes within and between structural elements (e.g. sections). A well-explored way of doing this is by extracting a novelty function [Foo00] from self-similarity matrices [PMK10] computed from relevant features from an audio recording. The features are chosen and the time scale of similarity computation depend on the segmentation task, e.g. we could use beat similarity matrix for tracking metrical structure of a music piece [Sri+12].

1.3.1.4 Analysis of timbre

The analysis of timbre mainly refers to modeling different instruments present in a recording. A few timbre analysis tasks would include percussion transcription (described above), diarization of instruments played (instrument-based segmentation of an audio music recording) and source separation. Source separation for music is a well-explored topic with both classical methods and deep learning methods applied to music [ABM13; LPS19; RS15; Sma04; Sri+14a] and hence not discussed in depth in this chapter. However, it is to be noted that these methods haven't been extensively applied to IAM, which might pose unknown challenges due to the complex resonances in instruments used in IAM such as the tānpurā, sitār and the vīṇa (due to the bridge) or in mridangam/tabla (due to the loaded membrane).

While we presented problems on each dimension separately, the musical concepts do not exist in isolation and neither do the research problems and approaches. There is an overlap between the melodic, rhythmic, structural and timbral analysis tasks and hence we could devise approaches that aim to analyze and extract different subsets of descriptors across these dimensions. In addition, methods for analysis of one music concept could benefit from available prior annotations (manual or automatically extracted) from other dimensions, e.g. a segmented concert into music pieces with a single tāla will help to track tāla better, and in turn the availability of tāla cycle markers on the music recording will help structural segmentation of a music piece into different sections (since most sections are aligned with tāla cycle boundaries).

1.3.1.5 Analysis of scores

Symbolic music scores in Indian art music are not comprehensive and are only indicative, often only skeletal and prescriptive. They are seldom used in performance, but used to a limited extent in music training and archival through sargam notation for melody and syllabic notation for percussion. There are no standard notation systems for melody or percussion, in both Hindustani and Carnatic music, which are widely accepted and used.

Due to the large deviation of performed music from the indicative scores, score analysis can at best be good starting points towards further analysis. There are no standardized tools (such as humdrum [Hur02], music21[1]) for the analysis of IAM scores. Further, there is no comprehensive collection of machine readable music scores in Indian art music, though there are small collections built for analysis [SC12a].

Automatic score analysis research in the context of melody, rhythm, and percussion for Indian art music is scarce. Symbolic scores have been used for different melodic analyses in Carnatic music [Kod+14a; Ran+19a; RS13] for Carnatic music and for Hindustani music vocal compositions [Cho07; SC12b], creating a machine readable Hindustani melodic music score dataset. Tabla bōl sequences [Cho06] have been used for predictive modeling of solo tabla performances using the multiple viewpoint modeling framework [Cho+10; CSA10].

However, symbolic representations offer multiple advantages. Music scores can be used for symbolic analysis of melodic/percussion patterns, and used to discover patterns from score corpora, a task that much less complex than extracting them from audio. We can then use these patterns and search for them in longer transcribed recordings. Such an approach with pattern discovery from scores followed by pattern search in audio has been explored in the past within a speech recognition + keyword search framework. Since the music scores form the language to disseminate music, we can also utilize the recent advances in language modeling and natural language processing for analysis of music scores.

1.3.1.6 Analysis of text

Textual information for IAM mostly comprises metadata available with audio recordings, discussions about the music on different fora (for audience, musicians, music labels), commentaries, critiques, artist biographies and other information. The text information could be free form text or tagged text (linked text) with marked entities. Both forms of data could be analyzed through semantic web technologies that aim to integrate human knowledge into computer systems to solve complex problems that require human expertise. An ontology specifies concepts, attributes, relations, constraints, and instances in a domain [BHL01; BL04; GFC04]. Since music is a complex and varied phenomenon with many perspectives, a cultural domain specific ontology is needed to define the relationships that pertain to a specific type of music from which we could then derive different relationships between music concepts and community.

Built using the knowledge of music theory and practice, the ontologies would be useful for querying complex musical relationships between the pieces. The ontologies complement the features derived from the data with music knowledge based relationships that can be used for defining similarity, e.g. using a tāla ontology and the knowledge of cycle length, it might be easier to identify the tāla from audio. Further, the ontologies will also be useful to create specific models with priors obtained from the ontology. In summary, ontologies can be built both for a direct use in navigation and inference, and for building domain specific machine learning algorithms.

1.3.2 Data-driven musicology research problems

Recent advances in digital humanities have brought forward aspects of human behavior using large corpora. The focus of current studies in digital humanities lies largely on language and social data corpora while music corpora have received less exploration. Performance analysis of music corpora can provide us with several

[1] https://web.mit.edu/music21/

insights into different aspects of music and show us the contrasts and similarities between music theory and practice. Such analyses on larger corpora can yield us additional insights that are often difficult to obtain with traditional manual analysis.

Corpus studies are in general driven by the common motivation of contributing empirical results that improve the understanding of a specific property of data in the corpus. In music, typically, these properties are melody, harmony, and rhythm. Manual analyses of such properties in music corpora have been performed as long as the related disciplines, such as ethnomusicology or music theory, have existed. However, in the last decades, the availability of computational methods enables the evaluation of larger amounts of data more easily. Data-driven analysis of large corpora is especially amenable to computational methods and can provide additional tools for statistical analysis. Such analyses can provide broad corpus level inferences for a musicologist, complementing a manual detailed analysis of small set of representative pieces.

Data-driven musicology utilizes sizeable corpora (annotated or otherwise) of music - audio, scores, text for statistical analysis to infer musicological insights. For IAM, performance analysis of large set of audio recordings can provide us useful insights to verify common knowledge and in addition gather additional insights into music performances of IAM. The large corpora used for analysis help to provide statistical validity, capturing both generalities through averaging over the whole corpus, but also measure nuances as deviations from the average. There are a multitude of opportunities for data-driven musicology with existing corpora of IAM, but we illustrate specific studies as examples of the scope and utility of the approach.

Srinivasamurthy et al. [Sri+17b] provides insights into aspects of tempo and rhythmic elaboration in Hindustani music, based on a study of a large corpus of recorded performances. Focusing on aspects of tempo and rhythm, the paper demonstrated the value of a computational methodology for the analysis of large music corpora by revealing the range of tempi used in performances, intra-cycle tempo dynamics and percussion accents at different positions of the taal cycle. In the article, typical tempo developments and stress patterns within a metrical cycle were computed, referred to as tempo and rhythm patterns, respectively. Rhythm patterns were obtained by aggregating spectral features over metrical cycles. They reflect percussion patterns that are frequent in the corpus and enable a discussion of the relation between such patterns and the underlying metrical framework, the taal. Tempo patterns, on the other hand, are computed using reference beat annotations. They document the dynamic development of tempo throughout a metrical cycle and revealed insights into the flexibility of time in Hindustani music for the first time using quantitative methods on a large set of performances. This study was further extended to Carnatic music using the same principles and methodology [SIS19].

Similar statistical analysis of melodic concepts is also possible, as evidenced from the work by Ganguli et al. [Gan+16], which aimed to discover implicit patterns or regularities present in the temporal evolution of vocal melodies of Hindustani music. The paper applied existing MIR tools and techniques to a collection of ālāp performance. Using svara-based and svara duration-based melodic features, the paper quantified the manifestation of melodic concepts such as vādi, samvādi, nyās and graha svara in vocal performances.

[Kod+12a] aimed to characterize intonation in Carnatic music by parametrizing pitch histograms. The article obtained a pitch histogram for a recording, and poarametrized the prominent peaks. The parametrization followed by a qualitative assessment on a larger collection of rāga showed discriminative power of the entire representation that is also useful in musically relevant tasks such as performer and instrument characterization.

These examples are only illustrative and there is high potential for data-driven musicological studies that can complement current literature and extend the methods to large corpora of music.

1.4 Corpora, Datasets and Annotations

A significant part of data-driven research in MIR and computational musicology needs good quality data. Data corpora that are representative of the music culture under study are critical for building and testing analysis methods. To develop such MIR approaches and advance knowledge, there is a need for research corpora that can be considered authentic and representative of the real world. The data sources comprise of audio, metadata accompanying audio, music scores, lyrics, manual and automatic annotations, and linked (semantic) data.

Building such data corpora scientifically for MIR itself is a research problem [PF12; Ser14]. Setting up criteria for selection and curation of music, and designing datasets for research are to be done with objective parameters that can then be used to measure the goodness of a corpus for a particular research task. Collection of good quality data and easy access to both audio and metadata is essential for reproducibility of research.

A research corpus is an evolving collection of data that is representative of the domain under study and can be used for relevant research problems. A good data corpus includes data from multiple sources and can even be community driven. In the context of MIR, since it is practically infeasible to work with the whole universe of music, a research corpus acts as a representative subset for research. Hence, algorithms and approaches developed and technologies demonstrated on the research corpus can be assumed to generalize to real world scenarios.

A test corpus or a test dataset is often a subset of the research corpus, possibly with additional metadata for use in a specific research task. In experiments, test datasets are used to develop tools, and to evaluate and improve their performance. Computational approaches are developed using these datasets and then extended to the research corpus. Hence test datasets can even consist of synthetic data that can be used for testing. Unlike a research corpus, a test corpus is fixed for use in a specific experiment. A test corpus can evolve, but each version of the dataset used in a specific experiment is retained for better reproducibility of research results.

Serra [Ser14] has defined the principles to be taken into consideration for the creation of new research corpora, which can also serve for evaluating the goodness of the data collection for a particular research task. According to such principles, the purpose of the corpus has to be explicit, the corpora must have good data coverage for the phenomenon under study, the data must be as complete as possible and of good quality, and finally, corpora should be reusable, implying that the data should be available for other researchers. Wilkinson et al. [Wil+16] recently formulated a set of guidelines to improve the findability, accessibility, interoperability, and reuse (FAIR) of digital assets. While Serra emphasized the quality and coverage of data, the FAIR principles emphasize machine-actionability because humans increasingly rely on computational support to deal with data as a result of the increase in volume, complexity, and creation speed of data. The CompMusic project considered the tenets and guiding principles by Serra to build research corpora with good quality and coverage while emphasizing FAIR principles to store, navigate and reuse data. In addition, data corpora are most useful when there is open access to audio, metadata, and annotations, while allowing for the datasets to grow both in size and quality through community-driven efforts.

1.4.1 Research Corpora and Test datasets for Indian Art Music

All the music cultures under study can be described in terms of musical concepts, music content and the music community. The elements of the corpora can be associated with one or more of these categories and hence useful for computational tasks in these three aspects. Central to each corpus is an audio music recording with its metadata. For both Carnatic and Hindustani music, the CompMusic project aimed to build an annotated audio sub-collection that is representative of the real world performance practices[2] [Sri+14b].

The CompMusic Carnatic music research corpus mainly comprises audio recordings, associated editorial metadata, lyrics, scores, contextual information on music concepts, and community (social) information from online music forums and other sources. Audio recordings, editorial metadata, scores, and lyrics are the content used by signal processing and machine learning approaches. Contextual information and the forum discussions form the music concepts and community information used for semantic analysis. The primary metadata associated with each concert is the name of the release, the lead and the accompanying artists, and the musical instruments in the concert. For each audio recording contained in the release, the relevant metadata are the artists performing on the track, the name of the composition/s and the composer, rāga/s, tāla/s, musical form/s.

[2]https://compmusic.upf.edu/corpora

The CompMusic Hindustani music research corpus comprises commercially available music releases from several music labels. It mainly consists of khayāl and dhrupad vocal music releases, though a significant number of instrumental music releases are present. The metadata associated with each release is the name of the release, the lead and the accompanying artists, and the musical instruments in the concert. For each audio recording in the release, the relevant metadata are the artists performed on the track, the name of the composition/s (bandiś) and the composer/s (if composed), rāg/s, tāl/s, lay/s (tempo class), form/s, and section/s.

The editorial metadata associated with each release in both HM and CM corpora are stored in MusicBrainz[3]. MusicBrainz assigns a unique MBID for each entity in MusicBrainz, such as the artist, composer, instrument, recording, work, and a release. This helps to organize the metadata in an effective way. All the editorial metadata was entered using Latin alphabet and a uniform and consistent Latin transliteration [ISO01] was used when the language of the release was not English.

The test corpora (or test datasets) are designed for specific tasks and contain additional information such as annotations and derived data. They are useful for various melody and rhythm analysis tasks. There are several test datasets for different music cultures built within CompMusic[4].

1.4.2 Saraga - open data collections for Indian Art Music

The CompMusic Carnatic and Hindustani research corpora are representative and vast, but they comprise copyrighted audio. While the audio can be used for personal use, we cannot easily disseminate it further. To avoid this bottleneck, CompMusic project also built Saraga - two large open data collections [Sri+21] of Indian Art Music. The Saraga Hindustani and Carnatic collections comprise audio from vocal concerts, editorial metadata, and time-aligned melody, rhythm, and structure annotations. Shared under Creative Commons licenses, they currently form the largest annotated data collections available for computational analysis of Indian Art Music. The collections are intended to provide audio and ground truth for several music information research tasks and large scale data-driven analysis in musicological studies. A part of the Saraga Carnatic collection also has multi-track recordings, making it a valuable collection for research on melody extraction, source separation, automatic mixing, and performance analysis. An easy access to the audio, metadata, and the annotations in the collections is provided through an API, along with a companion website that has example scripts to facilitate access and use of the data. To sustain and grow the collections, a mechanism is provided for both the research and music community to contribute additional data and annotations to the collections. The paper also presents applications with the collections for music education, understanding, exploration, and discovery.

The data resources described in the section are collectively one of the largest curated and annotated sources of data for MIR and computational musicology research on Indian Art Music. They are carefully built to support multiple research problems and hence are valuable to the community to further research on IAM. They are disseminated through multiple publications [Gul16; Sri+14b; SS14a] that provide details about the content and potential applications of the datasets. The data is also accessible through different tools such mirdata, Dunya and PyCompMusic via easy to use APIs, each of which are described further in the next section.

1.5 Tools and Frameworks

The previous sections introduced relevant research problems in IAM and data corpora available to work on those problems. Though the scope of this chapter did not permit extensive description of the problems or the data available, we hope that the pointers provided and references will help to gather additional information on specific problems readers are interested in. Further, the introductory scope of this chapter also did not let us describe different methods taken up by the community so far for the research problems of interest. Other

[3] https://musicbrainz.org/

[4] http://compmusic.upf.edu/datasets

chapters of the monograph are expected to provide a survey on different problems and approaches explored so far.

In this section, we provide complementary information by introducing tools and frameworks we can use to annotate, store, archive, organize, retrieve, manually analyze and consume the datasets in different MIR experiments on the CompMusic and Saraga datasets. These tools and frameworks are more of enablers for research rather than methods of analysis, but nonetheless important to the access of data and some basic analysis. We do not claim to be comprehensive with the tools and frameworks presented here, but we describe the tools that have been used for the mentioned tasks in the past and have been adequate according to our experience with them.

1.5.1 Tools for Annotations and Analysis

Tools for annotation help us to annotate different tags and time-aligned events on an audio recording. In the process, they also enable us to do manual analysis on the recordings, which could be helpful to build hypotheses, propose analysis methods and design large scale experiments. Sonic Visualiser[5] [CLS10] is a free, open-source application useful for visualization and annotation of audio. It is designed for musicologists, archivists, signal-processing researchers, and anyone who wishes to carefully study an audio music recording. Sonic Visualiser is a program for highly configurable detailed visualisation, analysis, and annotation of audio recordings.

Sonic Visualizer supports the Vamp audio analysis plugin system[6], an audio processing plugin system for plugins that extract descriptive information from audio data - typically referred to as audio analysis plugins or audio feature extraction plugins. Some of the most common feature extraction methods (e.g. melodia for predominant melody extraction) are provided as Vamp plugins that could be used within Sonic Visualizer. For batch audio feature extraction, we can use Sonic Annotator, which is a non-interactive command-line program application that can batch apply the the same feature extraction plugins as Sonic Visualiser.

1.5.2 Tools for archival, retrieval, data access and evaluation

The frameworks for archival and retrieval are large linked data collections that can store different music related metadata and features and provide APIs for accessing the collections. The CompMusic data collections extensively use MusicBrainz, AcousticBrainz and Dunya for the purpose, supported by APIs.

MusicBrainz is an open music encyclopedia that collects music metadata and makes it available to the public. It is slated to the community curated and community maintained source of music information releasing the data under open licenses. It provides a reliable and unambiguous form of music identification through unique IDs called MusicBrainz ID. The AcousticBrainz[7] project aims to crowd source acoustic information for all music in the world and to make it available to the public. This acoustic information describes the acoustic characteristics of music and includes low-level spectral information and information for genres, moods, keys, scales and much more. The goal of AcousticBrainz is to provide music technology researchers with a massive database of information about music. AcousticBrainz is a joint effort between Music Technology Group at Universitat Pompeu Fabra in Barcelona and the MusicBrainz project. At the heart of this project lies the Essentia[8] [Bog+13] toolkit from the MTG - this open source toolkit enables the automatic analysis of music. The output from Essentia is collected by the AcousticBrainz project and made available to the public. AcousticBrainz organizes the data on a recording basis, indexed by the MusicBrainz ID for recordings.

Dunya[9][PSS13a] comprises the music corpora and related software tools that have been developed as part of the CompMusic project. These corpora have been created with the aim of studying particular music traditions

[5] https://www.sonicvisualiser.org/
[6] https://www.vamp-plugins.org/
[7] https://acousticbrainz.org/
[8] https://essentia.upf.edu/
[9] https://dunya.compmusic.upf.edu/

and they include audio recordings plus complementary information that describes the recordings. Each corpus has specific characteristics and the developed software tools allow to process the available information in order to study and explore the characteristics of each musical repertoire. The entire Carnatic and Hindustani research corpora are available through Dunya, which internally uses MusicBrainz IDs for identification and linking entities.

The Saraga[10] data collections distributed under Creative Commons licenses can be accessed in two primary ways. The data in the collections can be accessed through the PyCompMusic API built to access Dunya collections. In addition, the Saraga repository[11] provides a dump of all editorial metadata and some annotations. More recently, the Saraga collections are also available through mirdata[12], which is an open-source Python library that provides tools for working with common MIR datasets, including tools for downloading datasets to a common location and format, validating that the files for a dataset are all present, loading annotation files to a common format, consistent with mir_eval [Raf+14] and parsing track level metadata for detailed evaluations.

1.6 Summary

The chapter, targeted at musicians, music students, technology enthusiasts, engineers and researchers, aimed to provide a context of the current state of the art that could help them start their work with computational analysis of Indian art music. While emphasizing the need for laborious culture-aware and culture-specific approaches to computational musicology and music information research (MIR), we argued that we could alternatively also seek and identify parallel tasks, data and methods from the current state of the art in other music cultures and use them for a preliminary and basic analysis of culture-specific tasks. We noted that such a perspective will enable preliminary analysis of a music culture using existing methods, and integrate it as a use-case with existing common frameworks and approaches in MIR and data-driven computational musicology.

Within this context, the chapter provided a gentle introduction describing common concepts, frameworks, approaches, resources, data and methods for computational analysis of music from a perspective that could be useful for the analysis of Indian Art music. We hope that with the pointers provided in the chapter on the parallels between music concepts, relevant research problems, data corpora available along with tools available to access them would encourage further community-driven efforts and research into various open research problems in Indian Art Music.

[10] https://mtg.github.io/saraga/
[11] https://github.com/MTG/saraga
[12] https://mirdata.readthedocs.io/

Chapter 2

Aspects of Indian Classical Music

Mundur V. N. Murthy

The Institute of Mathematical Sciences, Chennai 600 113 (Retired)

> Music, both vocal and instrumental, undoubtedly played an important part in the cultural life of ancient India. Sanskrit literature, both secular and religious, makes numerous references to instruments of various kinds, and it is, I believe, generally held by archaeologists that some of the earliest mentions of such instruments to be found anywhere are those contained in the ancient Sanskrit works. Certain it is that at a very early period in the history of the country, the Hindus were acquainted with the use of stringed instruments excited by plucking or bowing, with the transverse form of flute, with wind and reed instruments of different types and with percussion instruments. It is by no means improbable that India played an important part in the progressive evolution and improvement of these instruments and might have served as a source from which their knowledge spread both eastwards and westwards. It would form a fascinating chapter of history... But the materials available for the writing of this history seem to be all too meagre.
>
> C V Raman, Palit Professor of Physics, University of Calcutta (1922)

2.1 Introduction

That was C V Raman's introduction to Indian Music in 1922. Much has changed since then, but much has also remained the same. Though there is more material is available now, it is still meagre to understand all stages of development in Indian music. Broadly Indian Music refers to several sub-genres of music which includes folk music and the "classical" music of north and south. Of course here India should really refer to the subcontinent as a whole including Srilanka since there are common threads culturally with regional differences to varying degrees. There is a sacred, musical and in general a cultural geography that unites the people of the Indian subcontinent in spite of its enormous diversity.[1] Folk music is very diverse with each region producing its own version of not only music but also musical instruments. The boundary between sub-genres is fuzzy since each influences several others. More recently, we have to add Indian film music as well as Rabindra Sangeeth to the list. The film music, especially, has made huge strides in creating a new sub-genre of Indian music. In early part of 20th century it was mainly based on the classical music of north and south. But in the fifties the film music came in to its own; even though it retained its base in Indian classical/folk music tradition, layers of jazz and European classical music were added making it a likely new sub-genre of Indian musical experience. There were and still are outstanding music composers who have made a significant mark not only on the Indian film music but also outside.[2] Rabindra Sangeeth is based on poetry written in Bengali and music composed by the poet Laureate Rabindranath Tagore. The rendition is done in a distinctive style though based mainly on the folk music of Bengal. It is also influenced by Indian classical Music as well as European classical music[3]. An

[1] This is a non-expert overview of the development of Indian classical music, by a connoisseur who is neither a professional musician nor a musicologist. Much of what is written here is based on what has been learnt through discussions with musicians and avid listeners over a long period of time.

[2] Often Bollywood or Hindi film music is mistaken for all of Indian film music but there is far more creativity outside of Hindi film music, especially in South Indian films than is often recognised or given credit.

[3] The word European classical music is used here instead of Western classical music since Western music is a much larger set.

endearing feature of the music is its balance between poetry and music which enhances the enjoyment of music for those who know the language.

We will be mostly dealing with the classical music (or some times called art music) which is broadly divided into North Indian (or Hindustani Music) and South Indian (or Carnatic Music). The basis of these two systems is essentially the same- as we shall see later the sequence “Set of Notes (scale)+ Movement+ Intonations(gamakas)= Raga/Melody" is more or less the same in these two systems. Even the rhythmic cycles are similar though they may bear different names. Major differences arise in the structure, presentation and improvisation as well as in the nature of musical instruments employed including drums.

2.1.1 The Oral Tradition

The origins of Indian music[4] system perhaps goes back more than 2500 years to the vedic period as far as we know or as it is generally assumed. Like the origin of a river, beginnings of a musical tradition is hard to pin down; there may be several streams coming together over a long period of time. For example, we do not know much about the music of Harappans or the people of Indus civilisation (3300 BCE- 1300 BCE) who were part of the earliest known civilisation in this part of the world. It is reasonable to speculate that some elements of “Harappan music" must have filtered down to those who came to live in this sub-continent much later. A major difference between Harappans and the vedic people is the presence of writing among Harappans, though we do not yet know what they wrote - the script has not been deciphered yet- and its complete absence among the vedic people in the beginning. That is the crucial point as we shall see in the development of Indian music.

In the absence of writing, the vedic people depended entirely on transmitting their offerings, observations (as in the Vedas) and books of knowledge (the Upanishads) orally. This method of learning and transmitting knowledge persists even today in orthodox circles although writing has been in existence in the subcontinent for more than 2000 years (at least since the time of the edicts of emperor Asoka) in the subcontinent. The prevalence of this oral culture is central to the evolution of many branches of knowledge, especially to the development of music. All evidence points to early vedic hymns, as in Rig-Veda, were composed orally. This is then memorised and whole books passed on orally from one generation to the next. Obviously the oral transmission puts a premium on memory. The most effective way of putting things in memory is to compose every thing in verses on which to impose a melodic and rhythmic structure. Thus music became the backbone to the transmission of knowledge composed in the form of verses. The recitation of vedic hymns, even now, is done mainly varying through three (to maximum seven) pitches in a rhythmic fashion. The prosody and music developed in parallel.

Historically therefore we see that almost all the literary texts, in Sanskrit as well as in even older Dravidian languages are set in verses with a well defined structure [5]. Many of these literary works are long and transmitted orally in most cases. The epic Mahabharata is a great example with nearly 2 million words in its longest version set to verses (with some prose passages). Oral transmission is enabled by setting the verses into music with well defined rhythmic underscore. It not only enables pleasant rendition but more importantly enables the work to be memorised easily. This is generally true of all oral cultures where the literary and religious texts are invariably composed in verses. Surprisingly, this tradition has continued even into recent times long after the arrival of writing.

In fact, even today children learning Carnatic music tend to memorise tens of compositions which would be impossible if they were to be rendered in prose. Therefore music plays an essential role in preserving the ancient texts and forms of knowledge in oral cultures. Naturally the development and evolution of music goes in parallel to the literary tradition. It is not surprising that we find some parallel in the development of Indian

[4]From here on I refer to the Indian Classical Music as Indian Music for ease of presentation. Just being lazy without meaning disrespect to other sub-genres of Indian music.

[5]The earliest prose fiction was probably Bhanabhatta’s work “Kadambari", name of the main character, which has become the name for any literary fiction in some Indian languages like Kannada and Marathi. There are not many such prose texts surviving from ancient times, while there is copious supply of verse compositions in all languages.

mathematical tradition. Like music, elementary arithmetic and geometry formed the basis of many basic vedic rituals. These, the arithmetic and geometry, later developed into sophisticated fields of mathematics in their own right. However, unlike music the Indian mathematical tradition came to an end rather abruptly around 16th century or there about, when the musical tradition was actually expanding and new sub-genres were coming up. This may have some thing to do with the fact that the ruling class or aristocracy (I suppose every where) cared much less for science and mathematics while at the same time required music (and dance), all of arts in general, as show pieces of the courts or simply for entertainment. Well known musicians were acquired as “jewels" of the royal courts. More “royal" the court was, the greater were the number of musicians inhabiting the court.

This trend continued even during the colonial period even when the royal court system was on the decline. Though the formal education system underwent a big change under the colonial system, it did not affect the teaching methods of music in any significant way. The new system brought about by the colonial rulers was more focused on preparing a servile class of loyal citizens of the empire to serve as clerks or in other lower level jobs. This did not require the formal system to include curricula in fine arts and music. Therefore the education in music continued without any serious interruption which explains the continuity in arts. This is unlike in the fields of science and mathematics where there was a clear break. The sciences had to wait another few hundred years for its revival while the musical evolution both in the North and in the South continued without a break to give us some of the most evolved and sophisticated systems.

Furthermore, we may identify two important consequences of music evolving in an oral culture:

- Firstly, the music is predominantly vocal set to a rhythmic pattern. Even when instruments are used they are designed to reproduce or assist the vocal music. This is very much true in Indian music where the instrumental music has not developed independently. Almost all the music is predominantly based on wordy compositions whether it is Dhrupad,Khayal or Carnatic compositions. An important consequence of this is the way the technology of making instruments evolved–the musical instruments of India have some features in the form of curved bridges, use of wooden flute, or drums with membranes wound around dried bamboo shoots.[6]These are the features that bring the sound produced by instruments closer to the variations (glides and oscillations) found in the human voice[7].

- Secondly, because it is an oral tradition,the notes or the pitches involved must be consonant, pleasant sounding, so that the human ear can easily recognise, grasp and reproduce them. This means the set of notes used in the rendition of verses must be in simple integer ratios with respect to a basic pitch or tonic. Since the tonic may vary from person to person (gender dependent too), relative ratios are more suitable than specifying absolute values of the pitches. This is probably the reason for the adoption of the *natural scale* or *just intonation scale* as opposed to equally tempered scale in European music, a later development, which is more convenient for orchestral compositions and for scale transpositions.

2.1.2 Geographical spread of Indian Music

One thing that strikes any observer of Indian music is the geographical separation of the main sub-genres. Until recently, there was a very clear separation of what is called North Indian and South Indian music geographically as indicated by the names. That boundary is getting blurred in recent times. While the fundamental elements of the two sub-genres remain the same, it is also easy to see the differences in the presentation as also types of musical instruments used in a performance.

[6]This is very much unlike the European music where instrumental compositions are predominant in recent times. Often these involve multiple instruments, a few to a few tens of instruments. Instead dominating, the vocal music is just a part of the ensemble if ever it is present.

[7]For more on these unique aspects, see “Unique aspects and sound synthesis of stringed Indian musical instruments" by G P Scavonne, D V Patil and S R M Prasanna in this volume.

Another important factor that one notices is the nature of community of performers. Classical music any where in the world attracts niche performers and audience[8] . This is true of Indian classical music too. In the North Indian or Hindustani music the performers come from a variety of backgrounds. It is a relatively diverse mixture. On the other hand the South Indian or Carnatic music community, till recently, was less diverse (as compared to Hindustani music) restricted mainly to the elite brahmin community and to a lesser extent to few other caste groups. Part of it may be due to the practice of hereditary teaching in the background of a rigid caste system. About a century ago, the women who performed in public came almost exclusively from one community- the Devadasis. However, of late there are many women from other communities though it is dominated by the community of brahmins as in the case of male performers. This is also mirrored to a great extent in the audience diversity (or lack of it). This situation is changing, however slowly, in the last couple of decades where many first generation musicians are coming up, especially in the non-resident Indian communities abroad.[9]

To understand these commonalities as well as differences over a vast region we start from the times, perhaps about two millennia before, when people evolved a musical sense based on melody, rhythm and a structured composition rendered in verses. Over time a system of music which we can call Indian music with all its basic features would have evolved and formalised as outlined in many treatises. The oldest text is *Natya Sastra* attributed to Bharatha composed some time between 200BCE and 200CE. The treatise *Brihaddeshi*, attributed to Matanga, dated between 6th and 8th CE is perhaps the first text to talk of raga and folk music. A definitive text on Indian music is the treatise *Sangita Ratnakara* composed by Sarangadeva in 13th century. It has a detailed discussion of raga (melody) and tala (rhythm). Given the time line of these treatises, a fairly complete system would have developed by the beginning of the second millennium. At this time, the (classical) music was probably closer to the Dhrupad tradition as we know today.

This tradition which flourished in the Indo-Gangetic plains, was scattered around the subcontinent, especially towards South India which had relatively more political stability, during the initial raids of Mahmud of Ghazni. The learned elite who were uprooted from centres of learning, which were either destroyed or in a state of decay, went looking for “greener pastures”. Over time a syncretic music culture evolved in the northern plains as things settled down. The Hindustani Music evolved around this period in the 13th century as a sub-genre, separate from Dhrupad, influenced by the Turkic and Persian musical forms of the people migrating in to the sub-continent. Sitar was invented during this period combining elements of the veena and the Lute. Sarod came to the sub-continent from the North-West, probably the present day Afghanistan. Sarangi which was in existence in folk music for a long time started making its presence in Hindustani music.

While the Dhrupad had evolved into a solemn and spiritual musical tradition, the Hindustani music further incorporated the romantic and the secular elements as well as many differing styles of singing like khayal, thumri in to its repertoire under new systems of patronage. The new patrons were sympathetic to arts and architecture while science and mathematics were an anathema at least in the beginning. Along with music the fusion of cultures gave rise to the creation of remarkable architectural monuments of which the Taj Mahal is probably the most famous though Humayan’s tomb in Delhi is probably the most aesthetic structure from this period.

Due to the turmoil in northern parts and lack of patronage from the new rulers initially, the scholarly centre of gravity shifted to the southern part of India[10]. This is seen in the development mathematics and sciences in South India. It is now well known that Calculus was already developed by the Kerala school of Mathematicians some three centuries before Newton and Liebniz. Some of the musical elite may also have migrated to South

[8]In absolute terms the numbers in the audience in a classical music performance is necessarily small compared to a popular film music or a rock music concert which can easily fill a sports stadium.

[9]The study of diversity (or lack of it) in general in Indian Music is a field which encompasses issues of caste, gender, and religion. A good starting place to go deep may be the books by Saba Dewan-*Tawaifnama*, Context (2019) and V Sriram-*Devadasi and Saint*, East West Books (2007).

[10]For a very detailed account of these developments, see P P Divakaran, *The Mathematics of India, Concepts, Methods and Connections*, Hindustan Book Agency (2017). Even though this book is about the history of Mathematics in India, the initial chapters contain very nice historical introduction to the events from the beginning of the second millennium.

contributing to the evolution of Carnatic music. The south of India was some what insulated from the upheavals occurring in the northern region. The people moving towards south brought with them the rituals and other such practices along with their music. The Bhakti movement in the east and south of India had profound influence on the poetry and music in regional languages. The Tamil and Kannada poetry had a huge repertoire by the 12th century influenced by (as also influencing) movements towards a more egalitarian society. The folk music of the region already had a strong and very well developed rhythmic components involving complex instruments. Coming together of these disparate elements, the music of the travelers, the rich poetry and the existence of a strong sense of rhythm, paved the way for the creation of yet another sub-genre of Indian music–the Carnatic music. Unlike theHindustani music, the Carnatic music is relatively more structured even while improvisation is very much part of it. It also lays greater emphasis on the rhythmic patterns embedded in to the compositions; the origins of this may also go back to the Dhrupadtradition.

While the coming together of these elements created a sophisticated sub-genre of Indian music, the community of practitioners of music was dominated by the elite mainly from the brahmin community. Even when it included other communities like the Devadasis or the Nadaswaram practitioners, they lacked the same kind of patronage and were even considered socially inferior in status. The impact of such an exclusive tendency is visible even today. Through globalisation some of these barriers are slowly softening especially in recent decades. In contrast, on the surface it appears that the Hindustani Musicians come from relatively more diverse backgrounds, partly due to the impact of a variety of influences over a long time which shaped the sub-genre. This is an aspect that is waiting to be studied in depth. As remarked earlier, this also reflects on the audience diversity to some extent; the Hindustani Music has a relatively more diverse audience and hence more numerous compared to the niche audience of Carnatic Music.

2.2 Natural Scale

As discussed in the last section the evolution and use of natural scale, that is when the pitches are in rational ratios, in Indian music may be attributed to the vocalisation (at least partially if not wholly) or more generally to the oral tradition. When we hear the music the ear responds to natural resonances and hence sequences of melodic outlines. The choice of a set of semitones or srutis (or simply notes) starting from a tonic (reference frequency or pitch) in a scale is then dictated by the relative ratios of notes which are concordant (see appendix). For this to happen the physics of sound vibration tells us that the relative magnitudes of the pitches or fundamental frequencies be in integer ratios- simpler the better. There may be many ways of picking these successive ratios. Indeed while the choices are many, over the years we have arrived at an agreed sequence of notes in an octave.

This is not necessarily true in the design of musical instruments like Piano or the classical guitar. The Piano for example, has fixed relative ratio of pitch from one key to the next. The same is true of classical guitar. This aids in the precision construction of the instrument. This also aids in musical performances involving great many instruments. Without this precision tuning, it is impossible to produce a harmonious blend of music as in many symphonies. Furthermore it also helps in pitch transposition of a melody (A-major, C-Major etc). While this is so, the constant pitch ratio, namely $2^{1/12}$, is an irrational hence in principal not consonant. But as we shall see later in this section, it is chosen in such a way that it is very close to the natural scale ratios, and therefore remains musical.

The debate about number of intervals in an octave is also preceded by the assumption that music is vocal music only. In principle in vocal music we may have as many intervals (12 or 22 or what ever) as necessary depending on the underlying assumptions about melody. This arbitrariness in the number of divisions is impossible to achieve in fretted/keyed instruments where the number most often has to be fixed during the construction stage itself. In fact lesser the number of intervals, better it is since the precision of playing also depends on the average size of the finger and the smallest width of the fret or key.[11] Which is why a 12 interval

[11] With the advent of electronic musical instruments, this limitation has been over come- a simple touch is now enough.

sequence of notes in an octave is the most practical system from an instrument makers' point of view and certainly not a necessary condition for vocal music, at least as it seems at the outset.

In this section, we shall discuss some of these issues specific to Indian music in some detail starting with the natural scale, and why the musical instruments are what they are in Indian Music. Before we proceed a clarification- in the context of Indian music, by scale we mean a set of pitches or notes that are used in a raga. The set of notes forms the back bone of a raga over which many other attributes are superposed. These notes are usually defined in an octave. This definition is independent of the tonic, unlike the definition of a scale in modern European music. We shall discuss the transition from scale to raga in the next section.

2.2.1 Natural and Tempered Scales

In Music, Indian or any other system, a scale is defined as a set of pitches or frequencies, $f_1, f_2, \cdots f_{12}$ with $f_{13} = 2f_1$. The interval from f_1 to $2f_1$ is called an octave. With reference to this the interval from $2f_1$ to $4f_1$ will be an octave higher. It is a matter of convention now to choose 12 distinct pitches in an octave. Arguments have been made for 22 frequencies or some other such number while 12 is the smallest number of such a division in an octave.[12] For our purposes here let us keep the number of pitches to 12 in an octave–the simplest. we will comment on the 22 srutis later. These f_i's are not necessarily the same in all systems and even the nomenclature varies. For example in the Indian systems they are denoted variously as

$$S \;\; R_1 \;\; R_2(=G_1) \;\; R_3(=G_2) \;\; G_3 \;\; M_1 \;\; M_2 \;\; P \;\; D_1 \;\; D_2(=N_1) \;\; D_3(=N_2) \;\; N_3 \;\; S' = 2S$$

with 16 possibilities out of 12 note positions. This nomenclature is useful in classifying the 72 basic ragas in Carnatic music. Another simple notation that is also used is

$$S \;\; R_1 \;\; R_2 \;\; G_1 \;\; G_2 \;\; M_1 \;\; M_2 \;\; P \;\; D_1 \;\; D_2 \;\; N_1 \;\; N_2 \;\; S'.$$

This notation is used in Hindustani music and the notion of Thaats or groups of scales uses this nomenclature. The notes are also identified by the names like "komal Rishab (R)" etc (notice the names for a sruti is again a result of the oral culture where the adjective identifies the specific note). The notation with reference to frets on a veena are illustrated in Fig.(2.1). We will stick to the second named sequence, unless otherwise specified, as our preferred notation which is simple to follow without degeneracies in the names or notation.

In order to set up a scale we need to fix the pitch or fundamental frequency of $S = f_1$. This is called the "Tonic" or basic "Sruti". Once this is fixed all other frequencies are fixed relative to this. The immediate question for us is how are the relative frequencies of the notes fixed in a scale? We will do this now.

There are two common systems used in fixing the ratios of notes with respect to the tonic S whose absolute frequency is f_1, say. In Indian systems[13] the tonic can be chosen arbitrarily depending on the performer while in the European music system its value is fixed in a given scale[14]. Let us say this is 1 in some unit.[15]

2.2.2 Ratios in the natural scale

We first discuss the Natural scale (also called the "Just Intonation Scale" or Helmholtz scale) based on the overtone series of a simple vibrating string or air column. Hence it is natural! All the notes in the scale are related by rational numbers.[16] This is most commonly used in vocal music.

[12] Historically there have been even smaller divisions, but since Venkatamakhi classification in 17th century, twelve pitches appears to have become the standard in Indian music. In fact the name octave would indicate a seven note system where the eighth would be a multiple of the first.

[13] This system is also widely used not only in Indian music, but also in the music of many east Asian countries including Japan.

[14] An exception is to be found in musical modes which are tonic dependent. There are seven main modes named after the geographical regions in ancient Greece. For example the Ionian mode is a major scale with starting note C- it is simply C-major scale. It is interesting that modes, which are akin to ragas, have geographical identification like many ragas in the Indian Music; for example the raga names like Multani, Gurjari, Karnata, Kanada are also geographic indicators.

[15] There are also other systems which we will not discuss here: for example the Pythagorean system where a musical interval is given by frequency ratio equal to a power of two divided by a power of three, or vice versa. Some elements of the Pythagorean system is also present in the Natural scale.

[16] A rational number is one which can always be expressed as a ratio of two integers.

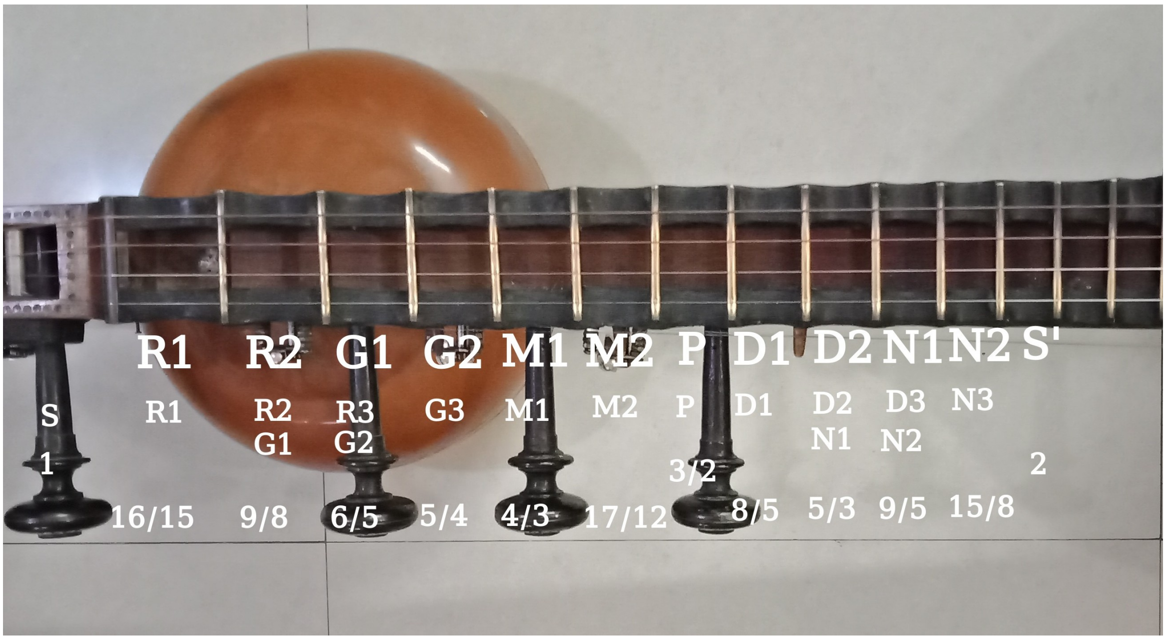


Figure 2.1: The two notations for pitches are illustrated with reference to the frets on a Veena. Open string corresponds to the tonic S. The fractions refer to the ratio of the pitches with respect to S which is taken as 1.

Let us first write down the ratios of all the 12 notes in an octave using our favoured notation (see Fig.2.1):

$$\begin{aligned}
S &= 1 \\
R_1 &= \frac{16}{15},\ R_2 = \frac{9}{8},\ G_1 = \frac{6}{5},\ G_2 = \frac{5}{4}, \\
M_1 &= \frac{4}{3},\ M_2 = \frac{17}{12} \\
P &= \frac{3}{2} \\
D_1 &= \frac{8}{5},\ D_2 = \frac{5}{3},\ N_1 = \frac{9}{5},\ N_2 = \frac{15}{8} \\
S' &= 2
\end{aligned}$$

Notice that the fixed or invariant notes of the scale are S,P,S' in the ratio 1 : 1.5 : 2. Further, $R_2, G_1, G_2, M_1, D_1, D_2, N_1$ are all in simple integer ratios with all the integers in the interval 1 to 9. The others R_1, M_2, N_2 are in ratios where the integers are all in the interval 11 to 19 (except the denominator of N_2). Some times one uses slightly different integer ratios for these three notes which are close to these (for example $R_1 = 25/24, 256/243, M_2 = 45/32$). The most pleasing sounds, a qualitative measure, apparently to the ear are, combinations of notes that are related by ratios of small integers, like (3/2) or (5/4) or (4/3).

There are some nice patterns in these ratios: For example

$$S = 1, \quad G_2 = 5/4, \quad P = 3/2,$$

has the same relative ratios as the sequence

$$M_1 = 4/3, \quad D_2 = 5/3 = G_2 M_1, \quad S' = 2 = P M_1,$$

and

$$P = 3/2, \quad N_2 = 15/8 = G_2 P, \quad R_2' = 9/4 = P^2.$$

In the last sequence we over shoot the octave since $R_2' = 9/4$. We can fold it back into the octave between 1 and 2, after dividing it by 2 and set $R_2 = 9/8$. This method of fixing the ratios is used often in fixing the frets on a Veena.

These three *Triads* then give the following sequence of seven notes in the basic octave (1 to 2):

$$S=1,\ R_2=\frac{9}{8},\ G_2=\frac{5}{4},\ M_1=\frac{4}{3},\ P=\frac{3}{2},\ D_2=\frac{5}{3},\ N_2=\frac{15}{8},S'=2$$

There are some more relations between these ratios: for example we also have $R_2M_1=P,\ G_2M_1=D_2,\ R_2D_2=N_2$. These seven notes in the scale make up a basic scale that is popular and is used in the Raga Sankarabarana in Carnatic Music or Raga Bilawal in the Hindustani Music. Written in terms of intervals between notes this may also be written as 2-2-1-2-2-2-1 which corresponds to a major scale in European music though not identical to it since the relative ratios are different(see next subsection).

To these now we may add the remaining notes R_1,G_1,M_2,D_1,N_1 to make up 12 divisions of an octave. All these twelve notes are expressed as ratios with respect to the tonic S while the tonic itself is arbitrary and is left to the performer to choose.

To summarise, all the twelve notes in the octave are in simple integer ratios or rationals, further they form triads and are also related to each other by further simple ratios. The acoustics, physics of sound, classifies these notes as consonant with respect to each other; hence pleasing to the ear. When ever a string that is fixed at the both ends is plucked, the sound produced is a superposition of the fundamental and its harmonics which are integer multiples of the fundamental. A brief and rather simple discussion of string vibration is given in the appendix-A for the interested reader.

2.2.3 Tempered scale

The Tempered scale or equally tempered scale was developed in the context of keyboard instruments, such as the piano-also used in Guitar, so that they could be played starting from any key since the ratio of one note to the next is always a constant. It is essentially designed for instruments used in large ensembles to provide a precision tuning when playing together. However, as we shall see below it has a draw back-if at all, since the notes in an octave are not consonant unlike the notes in the Natural scale.

Though the notation for the 12 notes in an octave differs from one system to another, let us stick to the notation used for Natural scale so that they may be compared easily. The ratios of all the 12 notes in the harmonic system is rather simple and straightforward:

$$\begin{aligned}
S &= 1=2^{0/12},\ R_1=2^{1/12}\ R_2=2^{2/12},\ G_1=2^{3/12},\ G_2=2^{4/12}\\
M_1 &= 2^{5/12},\ M_2=2^{6/12},P=2^{7/12},\\
D_1 &= 2^{8/12},\ D_2=2^{9/12},\ N_1=2^{10/12},\ N_2=2^{11/12},\ S'=2^{12/12}=2
\end{aligned}$$

Thus the n-th note after S has a frequency ratio with the tonic given by $2^{n/12}, n=0,\cdots,11$. An important fact to be noticed is that the ratios are not rational (unlike in the Natural Scale) hence in principle are not consonant with each other. In fact apart from S,S', every other ratio is an *irrational*. As is known from number theory every irrational is close to a rational in its immediate neighborhood in the interval between 1 and 2. This is a compromise tuning that enables ease of tuning of keyed instruments like Piano and is widely used in European classical music. The relation between any two successive notes remains the same no matter where the scale begins.

2.2.4 Comparison of Natural and Tempered scales

As we have seen the ratios of the pitches are defined differently in these two systems. How close are they to each other. Often this comparison is done in so called *cent scale* where we define the value of the note in an octave as given by

$$x_f=1200\log_2\left[\frac{f}{f_0}\right],$$

where f_0 is the pitch of the tonic (in our case this is simply S) and f is the pitch of any note in the octave represented by x_f. For example the value of S in the cent scale is

$$S = 1200\log_2\left[\frac{f_0}{f_0}\right] = 0.$$

and

$$S' = 1200\log_2\left[\frac{2f_0}{f_0}\right] = 1200,\ \log_2 2 = 1.$$

The $n-$th note in between has a frequency $f_n = 2^{n/12}f_0$ in the tempered scale and therefore

$$x_n = 1200\log_2\left[\frac{2^{n/12}f_0}{f_0}\right] = 100n,$$

which are equally spaced on the log scale. Since it depends only on the ratios of frequencies, the interval remains the same no matter what the frequency range is. This is a great advantage in comparing different scales.

A comparison of natural scale and tempered scale is shown in the table [2.1]. Notice that even though the two systems appear very different, the cent-scale values are close to each other. Even in the natural scale the difference between consecutive notes is approximately 100 cents. This is also shown in figure 2.2 where the ratios are plotted against the cent values of the notes.

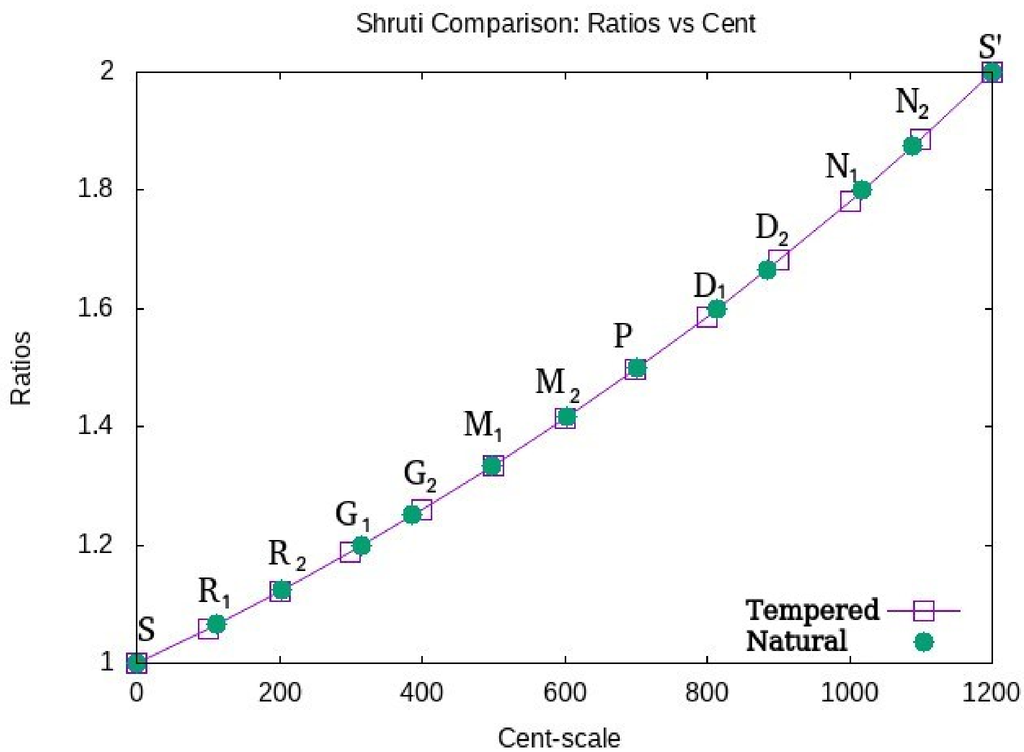


Figure 2.2: The values of the notes/srutis in natural and equi-tempered scales are shown as ratios vs cents. The two are very close in the 12-sruti system and can not be easily distinguished even on cent scale.

As mentioned before since the ratios of successive notes are exactly the same, the tonic in the tempered scale may be shifted with impunity without changing the relation between successive notes. The transition from one note to the next is a half step (or semitone) and is usually denoted by one unit. Starting from the tonic, we may choose seven notes in succession to define a major or a minor scale:

- A major scale is defined by choosing the following steps:

$$Major\ Scale: \quad 2-2-1-2-2-2-1$$

n	Note (sruti)	Natural Ratio	scale Cents	Tempered Ratio	scale Cents	Semitones
0	S	1	0	1	0	Unison
1	R_1	16/15	111.7313	1.0595	100	minor 2nd
2	R_2	9/8	203.9100	1.1225	200	major 2nd
3	G_1	6/5	315.6413	1.1892	300	minor 3rd
4	G_2	5/4	386.3137	1.2599	400	major 3rd
5	M_1	4/3	498.0450	1.3348	500	4th
6	M_2	17/12	603.0004	1.4142	600	diminished 5th
7	P	3/2	701.9550	1.4983	700	5th
8	D_1	8/5	813.6863	1.5874	800	minor 6th
9	D_2	5/3	884.3587	1.6818	900	major 6th
10	N_1	9/5	1017.5963	1.7818	1000	minor 7th
11	N_2	15/8	1088.2687	1.8877	1100	major 7th
12	S'	2	1200	2	1200	octave

Table 2.1: Comparison of Natural and Tempered scale. The values in cents show the differences magnified in the scale. The tempered scale is equidistant in the cent scale. The last column refers to the names of notes as in Tempered sale.

- A natural minor is defined by

$$Minor\ Scale: \quad 2-1-2-2-1-2-2$$

- A Harmonic Minor is defined by

$$Harmonic-Minor\ Scale: \quad 2-1-2-2-1-3-1.$$

Note that all the steps add up to 12 which is how the octave is divided. These are some common musical scales used in European instrumental music.

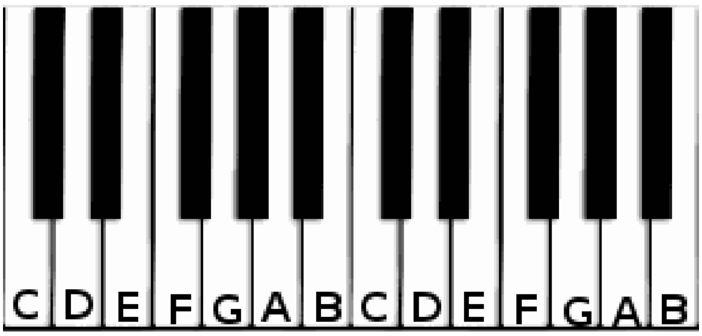


Figure 2.3: The piano keys with the note definition for all the white keys. The black keys correspond to the sharp notes. For example the black key next to F is denoted as F-sharp.

Furthermore if the starting note is C, referring to the Piano Keys shown above in Fig.[2.3], a major scale would be

$$C-D-E-F-G-A-B-C$$

which includes all the white keys in the octave. This is the well known C-major scale which corresponds to in our notation the following set of notes

$$S,\ R_2,\ G_2, M_1, P, D_2, N_2, S'.$$

It involves S,major 2nd, major 3rd, 4th, 5th, major 6th and major 7th within the octave. These are the familiar notes present in the Raga Sankarabarana or Bilawal. A major difference in the Indian classical system is that the tonic S is not fixed. Depending on the pitch of the starting key it will generate other major scales keeping the same set of intervals. Similarly one can construct minor scales too using the steps given above. Depending on the starting pitch or the tonic, the music feels different bringing a wide variety of melodies of compositions in European music. Unlike the Indian music which involves many qualitative elements, the scale based compositions in the European classical music can be rendered into a written form with precise notation which is easily followed. The written form then easily enables multiple instruments coming together in an ensemble as is evident in the performance of many symphonies. The music is truly "classical" in this sense- an established formal and complete system with well developed grammar.

In tempered scale therefore, it is sufficient to define the scales which forms the basis for the rendition of any composition by composers. (An exception may be made in older choir music, and other genres like Blues, Jazz etc where tempered scale is not suitable or not used at all).

2.2.5 Some observations

The 12 note system is influenced by the European tradition which had 12 semitones from which different scales are obtained as seen earlier. Since the ratios are equal from one note to the other, tuning is straightforward and the tonic key may be shifted easily to obtain all the major and minor scales. This has great advantage when a large number of instruments are playing together. On the other hand shifting the tonic in natural scale, changes the relative ratios and hence the scale itself. This is often called *Sruti Bedha* (or tonic shift within the scale) and is used to bring in elements of a different raga. But in reality it may worsen the melodic aspect.

But there is a problem here[17]. In a fretted instrument like the veena the natural scale frequency ratios are fixed by the main string (the first one in front of the performer) which sets the tonic S. The next three strings are set to notes P,S,P an octave lower. In all the range is three and a half octaves. Since the frequency ratios are fixed by the main string for positioning the frets, what happens to the ratios in the second string which is usually set to $'P = 3/4$ an octave lower than the $P = 3/2$ on the main string? The ratios as they occur in the first and second string in a veena are shown in Table (2.2).

Notice that the lower $'D_2$ and the R_1, M_1, M_2, highlighted,have slightly different ratios in the second string as compared to the first string. This is also illustrated in Fig.(2.4). They are close but not equal to the corresponding ratios in the main string. All other ratios are reproduced in the second string which is set to $'P$. Hence the same composition played on the second string in a fixed fret instrument like veena should sound slightly different unless the musician is conscious enough to vary the tension manually while pressing on these frets to correct for the deviation. This may not be a problem for an vocalist or a violin/flute player since the continuous range is available. Corrections may be applied in real time while playing.

Before ending this section, we shall discuss yet another type of division of the octave that is much debated, namely the 22 sruti system[18]. As shown in appendix, the frequency or pitch is inversely proportional to the length of the string. Since every thing is expressed as a ratio, we may choose the length of the open string (S) to be 1. Then the frets of a veena are positioned at

$$\mathbf{1(open)},\ \frac{\mathbf{15}}{\mathbf{16}},\ \frac{\mathbf{8}}{\mathbf{9}},\ \frac{\mathbf{5}}{\mathbf{6}},\ \frac{\mathbf{4}}{\mathbf{5}},\ \frac{\mathbf{3}}{\mathbf{4}},\ \frac{\mathbf{12}}{\mathbf{17}},\ \frac{\mathbf{2}}{\mathbf{3}},\ \frac{\mathbf{5}}{\mathbf{8}},\ \frac{\mathbf{3}}{\mathbf{5}},\ \frac{\mathbf{5}}{\mathbf{9}},\ \frac{\mathbf{8}}{\mathbf{15}},\ \mathbf{1/2}$$

One more octave is covered from 1/2 to 1/4 of the length of the string. Here the length is measured from the bridge on the right side (plucking end). Even though there are only 12 notes from the tonic in an octave, in principle there could be other simple ratios between 0 and 1. In fact there are an infinity of them, 22 srutis being one of them. This is a much discussed set of srutis as it is also mentioned in some ancient texts- the standard 12 srutis as represented in a veena or used as standard now-a-days is a subset of these 22 srutis. We may immediately notice a difficulty here: as we go higher in pitch the width of the frets gets smaller and smaller.

[17] It is here the full power of equally tempered scale is realised. Since the ratio is constant from one note to the next, the note frequencies remain the same no matter in which string/key it is played.

[18] See also the article “Salient features of melody, rhythm and lyrics in Carnatic Music" by Venkata Subramanian Viraraghavan and Sridharan Sankaran in this volume for a detailed discussion. They arrive at a similar conclusion about 22 srutis.

n	Main string	Ratio	Second string	Ratio
0	S	1 (Open)	$'P$	3/4 (Open)
1	R_1	16/15	$'D_1$	8/10
2	R_2	9/8	$'D_2$	27/32 (*)
3	G_1	6/5	$'N_1$	9/10
4	G_2	5/4	$'N_2$	15/16
5	M_1	4/3	S	1
6	M_2	17/12	R_1	17/16(*)
7	P	3/2	R_2	9/8
8	D_1	8/5	G_1	6/5
9	D_2	5/3	G_2	5/4
10	N_1	9/5	M_1	27/20(*)
11	N_2	15/8	M_2	45/32(*)
12	S'	2	P	3/2

Table 2.2: Relative ratios of notes on the first and second string on a veena (only the lower octave is shown -see Fig.2.4). Note that the ratios for $'D_1, R_1, M_1, M_2$ (indicated by *) are not the same on the two strings even though they come in the same order as in the first string.

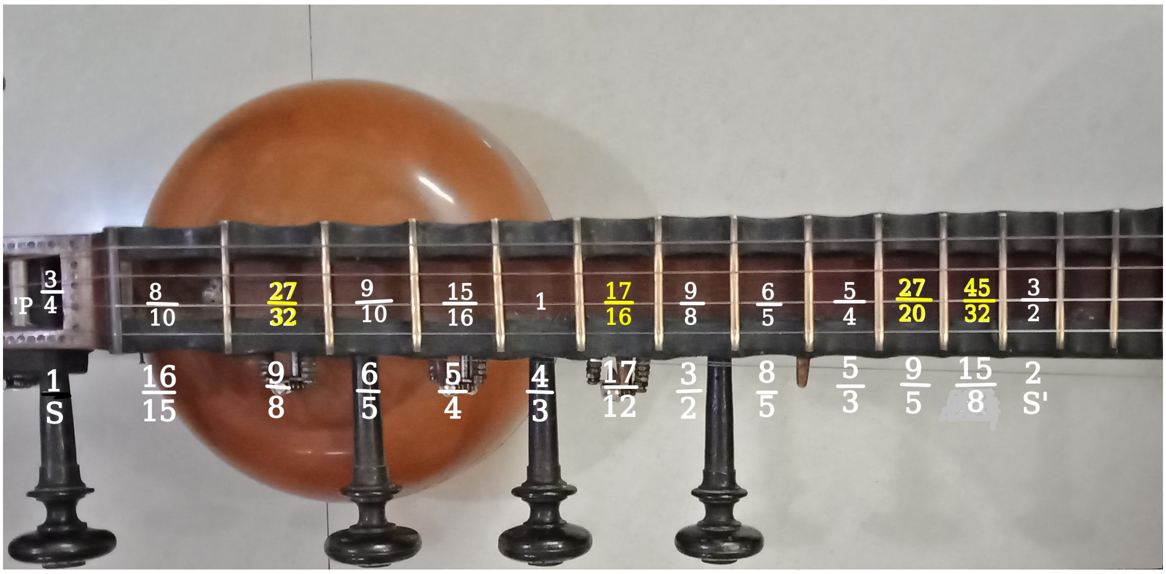


Figure 2.4: The two notations for pitches are illustrated with reference to the frets on a Veena. Open string corresponds to the tonic S. The fractions refer to the ratio of the pitches with respect to S which is taken as 1.

The smallest has the width of an average little finger. Any further division is not possible without increasing the length of the Veena fret board.

The way 22 srutis are organised is rather interesting; all of them are generated using essentially three ratios from the tonic S. Let us denote these three ratios by

$$x_1 = \frac{256}{243},\ x_2 = \frac{81}{80},\ x_3 = \frac{25}{24}.$$

Then the entire set of 22 srutis is reproduced as follows:

$$S = 1,$$

$$R_1 = S \times x_1 = \frac{256}{243},\ R_2 = R_1 \times x_2 = \frac{16}{15},\ R_3 = R_2 \times x_3 = \frac{10}{9},\ R_4 = R_3 \times x_2 = \frac{9}{8},$$

$$G_1 = R_4 \times x_1 = \frac{32}{27},\ G_2 = G_1 \times x_2 = \frac{6}{5},\ G_3 = G_2 \times x_3 = \frac{5}{4},\ G_4 = G_3 \times x_2 = \frac{81}{64},$$

$$M_1 = G_4 \times x_1 = \frac{4}{3},\ M_2 = M_1 \times x_2 = \frac{27}{20},\ M_3 = M_2 \times x_3 = \frac{45}{32},\ M_4 = M_3 \times x_2 = \frac{729}{512},$$

$$P = \frac{3}{2}$$

$$D_1 = P \times x_1 = \frac{128}{81},\ D_2 = D_1 \times x_2 = \frac{8}{5},\ D_3 = D_2 \times x_3 = \frac{5}{33},\ D_4 = D_3 \times x_2 = \frac{27}{16},$$

$$N_1 = D_4 \times x_1 = \frac{16}{9},\ N_2 = N_1 \times x_2 = \frac{9}{5},\ N_3 = D_2 \times x_3 = \frac{15}{8},\ N_4 = N_3 \times x_2 = \frac{243}{128},$$

$$S' = 2$$

Notice that the ratios are obtained cyclically by translating the consecutive notes by x_1, x_2, x_3, x_2, x_1 in each subset denoted by R, G, M, D, N. Essentially there are 4 srutis corresponding to each of R,G,M,D,N along with the fixed notes S,P, making up 22 srutis. The position of these srutis are well defined and contain the 12 srutis as a subset of this 22. If one can sing orally or play these notes precisely using an instrument there should not be any problem. But can you hear the pitch!

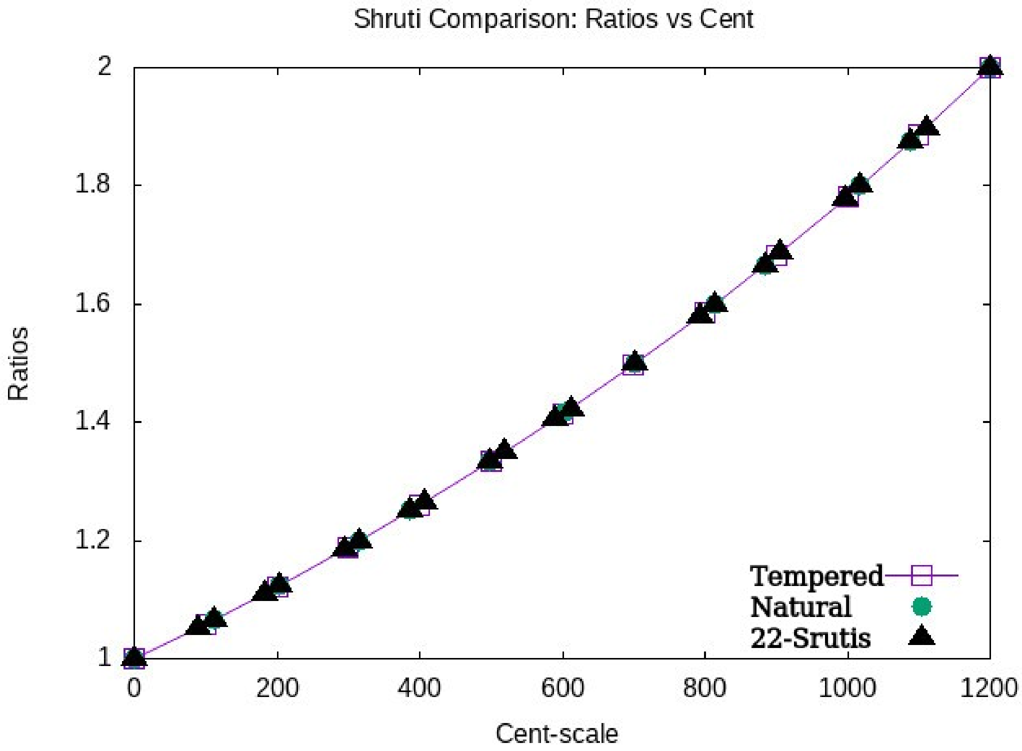


Figure 2.5: The values of the notes/srutis in natural and harmonic scales are shown as ratios vs cents. Superimposed are the 22 srutis which contain the 12 natural ratios.

In fig.[2.5], a comparison of the 22-srutis is shown along with the 12-sruti system (natural or harmonic). It is easy to see that except S,P S' (as it should be) all other notes are doubled giving the 22 srutis. The doubling happens due to the translation using the ratio x_2 which produces the smallest difference.

The translation by the ratio x_2 on a cent scale corresponds to generating a new sruti which differs from the previous one by approximately 20 cents on the average. Experiments on hearing clearly indicate that the ear can not easily distinguish between two pitches which differ by ± 10 cents at best, while on the average it may even by ± 20. In older people the error is even larger. So unless one carefully points out the 22 srutis, as it is often done in demonstrations, in any "blind" performance the listeners will be hard pressed to notice the fine distinction between two srutis differing by 20 cents or so.

It is further complicated for instrumentalists. For example, look at the construction of the veena. The length of the string is actually a variable- the average length of the string is typically about 80cm, or more up to a meter, between two supports. At one end is the fixed sharp bridge, but the plucking end has a bridge that is curved with a string winding length of 2.5-3 cm. Let us denote this by ΔL while L is the length of the string at its maximum. Therefore we have

$$\frac{\Delta L}{L} \approx 0.035$$

The frequency of the string is inversely proportional to the length of the string and hence over one winding of the string around the bridge the frequency can not be defined by a single number. This provides the body (timbre) of the note produced by a veena string (or sitar) which is not really a sharp single wave, but a wave packet. We may assume the width of the wave packet to be approximately given by the variation in length (in linear approximation). Thus we expect a deviation of about 3-4 percent about the mean position of every note.

We have to compare this to the variation produced by the $x_2 = 81/80 = 1 + 1/80$ which is about 1.25 percent over the previous sruti. There is no way this can be separated from the previous sruti. Thus if we collapse all transitions involving x_2, even if this is well defined in principle, we are back to the 12 note system as is commonly used. With the 12 srutis, the differences are on the average 100 cents which is well above the error involved in either hearing or in a wave packet generated by plucking the veena. Another way of looking at this is to observe the histogram of the average of many presentations shown in Fig.[2.6][19]. As can be seen, compared to the European music, the width of the notes in Hindustani and Carnatic music are much wider, in fact close to about 30-40 cents. The difference between R_1 and R_2 is only 22 cents which is an x_2 translation in the 22-sruti system. Therefore it is possible to theoretically define 22 srutis (or for that matter any number of

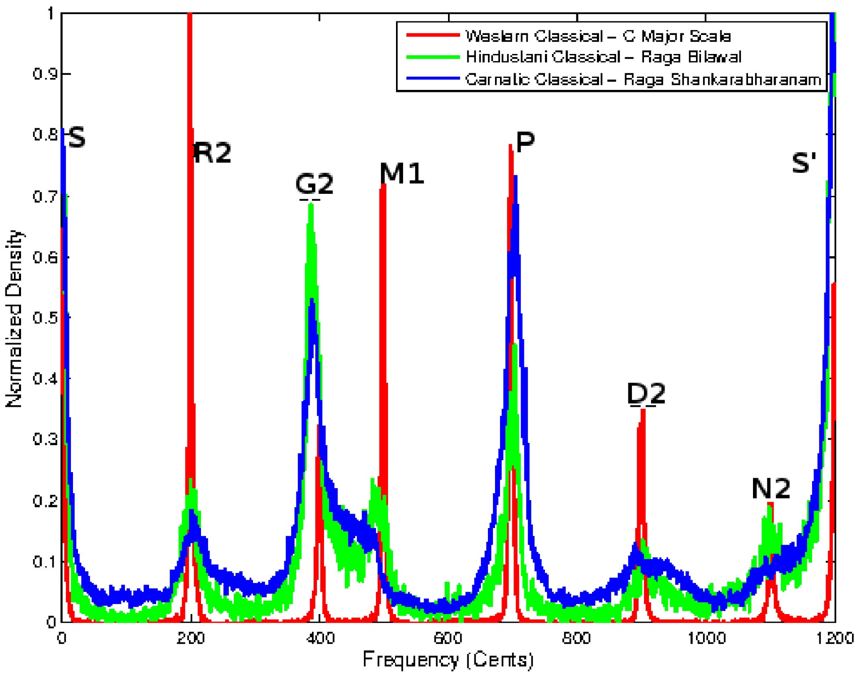


Figure 2.6: Histogram of time integrated performances on the scale denoted above corresponding to European, Hindustani and Karnatic system. For the sake of better comparison the x-axis is given in cent-scale after tonic normalisation of various performances. The shift in peak positions of some notes is due to the differences in natural and tempered scale ratios–Figure courtesy: Hema A Murthy

srutis) in an octave, but in practice some of these srutis even when they are well defined are not distinguished from the previous or the next note. Yet another practical difficulty for instrumentalists is the placement of frets for 22 srutis. Assuming the smallest width to be the average size of the little finger (as it is now), the length

[19]One must add a caveat in the comparison shown in the figure Fig.2.6- while the Hindustani and Carnaticmusic performances are vocal, the European music chosen for comparison is instrumental. The sharpness of the peaks corresponding to European music may be due to the precision of the instruments. However, even if it is vocal music it is reasonable to expect the peaks to be relatively sharper than they are in Indian music.

of the string on a veena has to be almost doubled to accommodate the 22 srutis which makes the instrument unplayable. Therefore the construction of the veena, as it is done now, involves a fantastic optimisation of various parameters making it convenient to play. A precise 22 sruti system is really a vocalists' imagination.

Even in tempered scale it is not necessary to restrict to 12 notes. If we define an arbitrary note as

$$x(n) = 2^{n/q}; n = 0, 1, 2, ..., q$$

we get q subdivisions of an octave, where q can in principle be any integer. By definition the relative ratios are $2^{1/q}$ and hence it is a tempered scale. Once again while such experiments are feasible in principle, in practice the physical size of the instrument has to be kept in mind.

2.3 From Scale to Raga

Some four decades ago a controversy arose in Carnatic music circles. One of the eminent vocalists of the time, M Balamuralikrishna, claimed to have invented many new ragas and composed songs in these ragas. His main critic, veena maestro S Balachander, maintained that ragas exist once the scale is defined through the notes. Therefore you can not invent some thing that already exists. The debate was far more nuanced than this bare summary portrays. Many other senior musicians of the time also got in to the act. The Madras Music Academy referred the dispute to a well known musicologist for resolution. It is not necessary to go in to this controversy here except to say that the opponents appeared to have different definitions of what a raga is!

So is a Raga defined merely through a scale made up of a set of notes, or is there some thing more to it? What kind of classification of ragas can we have? Can it be entirely logical and precise or are there subjective and qualitative elements? These are some questions that we will try to address here even though it is rather hard to find unique answers. One thing is certain, that a mere scale consisting of a set of notes in ascending and descending order alone does not define a raga in general. Scale is a necessary condition but not sufficient to understand a raga in all systems of Indian music.

This statement has to be qualified since in Carnatic music. There is *a precise classification* of certain ragas based on 72 fundamental scales. This classification has its origin in the 17th century treatise by Venkatamakhi called "Chaturdandiprakashika". These scales themselves are called as "Melakartha Ragas" (sequence generating or parent ragas) which are expected to give rise to all other ragas- the mother ragas as it were. All the daughter ragas are expected to arise from these basic scales (ragas). As a necessary condition for a raga this is a precise criterion.

How does one arrive at this number and what are the assumptions? Consider the 12 notes in an octave,

$$S, R_1, R_2(G_1), G_2(R_3), G_3, M_1, M_2, P, D_1, D_2(N_1), N_2(D_3), N_3, S',$$

where we have switched from our favoured notation that was used in the previous section for convenience of classification here. Some of the degeneracies between R,G and D,N are needed for counting purposes here: Let us choose all seven letters appearing in the ordered set, in our preferred notation, as a possible scale

$$S \;\; R \;\; G \;\; M \;\; P \;\; D \;\; N \;\; S'.$$

This has seven notes or srutis with S as the Tonic and the eighth is simply S' which is 2S. Now we ask the question in how many ways such a scale may be constructed? Notice M has two possibilities each while S,P,S' are fixed. The notes R,G are formed out of four note positions in that order. Similarly D,N can be chosen out of 4 notes in that order. In addition M can be chosen in two different ways. Thus the total number of possibilities is

$$^4C_2 \times ^4C_2 \times 2 = 6 \times 6 \times 2 = 72.$$

Suppose a performer wants to sing S,R,G,M,P,D,N,S', the same set of notes in both ascending and descending order, then the number of scale choices is 72[20]. Up to this stage the quantification is rather precise and does not depend on the the actual value of the frequencies.

These 72 distinct scales in *Carnatic Music* are called the *Melakartha* Ragas. This is the basic classification used in Carnatic Music and all other Ragas are supposed to originate in these 72 basic scales or Melakartha Ragas. This may be made even more precise by defining a “distance" between ragas so that the closest Melakartha raga to a chosen raga according to this distance may be considered the mother raga.

The classification system in Hindustani music is more recent, created by Vishnu Narayan Bhatkhande about a hundred years ago. While modeling his system based on the Melakartha system of Carnatic music he created the notion of Thaats. Going back to our original notation

$$S,R_1,R_2,G_1,G_2,M_1,M_2,P,D_1,D_2,N_1,N_2,S',$$

we note that R,G,M,D,N each have two values while S,P,S’ are invariant. Thus if we choose a sequence S,R,G,M,P,D,N,S' in both ascending and descending order, then the number of possibilities is $2^5 = 32$. These 32 scales are called Thaats. Bhatkhande avoided calling them ragas but according to his classification all ragas belong to one of these thaats. Out of these he highlighted ten basic thaats from which almost all traditional ragas in use arise. These are derived from the ragas contained in them namely, Bilawal, Kalyan, Khamaj, Bhairav, Poorvi, Marwa, Kafi, Asavari, Bhairavi and Todi.

Caution must be exercised here since a Raga in general is not merely a scale but combines the scale with many other features which are often not easily quantified. This is especially so in *Hindustani* or North Indian system. We will come back to this question later.

2.3.1 The “grammar" of Raga

A melodic outline played in Carnatic or Hindustani music is based on the concept of a Raga. There may be many points of origin for a melodic outline depending on the emotional content. We shall discuss this later. The scale, however, is essential to a raga– a necessary but not sufficient condition as emphasised. It is the backbone one which the body of the raga is built. It does not entirely define the raga which often includes many qualitative and some times even subjective elements.

Our attempt here is to outline the various elements that go into the make up of a raga starting from the basic elements in the form of notes which together form the scale. To provide an analogy, notes are like atoms while the scale is like a complex molecule. Different molecules may be formed by arranging the same set of atoms in distinct way. These atoms then interact and form the matter as we perceive. Hence even when notes are common, the ragas may sound very different just as matter made of the same set of atoms, may have very different material properties.

To begin with let us restrict these elements to Carnatic music where a relatively more precise definition may be given and many of its features are shared by Hindustani music. We will point out any additional or differing elements as necessary.

- A necessary condition for a raga is the basic scale. Here by scale we simply mean a set of pitches or notes within an octave. A simple raga which contains the same set of notes in ascending and descending order of frequency may have at best 7 notes (septa-tonic) and minimum of 5 notes (penta-tonic). (Some experiments with four notes have been done in Carnatic music).

- The 72 basic scales with all the seven notes, mentioned in the beginning, form the 72 Melakartha Ragas. It is sufficient to define the ascending order of notes (and same set of notes in descending order) for these

[20] If we relax the condition that the notes should be the same in ascending and descending order, then we get a huge number of possibilities. We are not considering these here.

Ragas, nothing else is specified explicitly. All of them are of the form $S,\ R,\ G,\ M,\ P,\ D,\ N,\ S'$ with same notes occurring in the ascending and descending order. So for these we may start by saying

Raga = Scale

- This is too restrictive a definition since it is applicable only to 72 basic ragas. This may now be expanded by allowing different notes in ascending and descending order and as well as movements where the pitches of the notes are not necessarily ordered in increasing or a decreasing order in frequency. Let us call this movement referring only to the ordering of the notes without further ornamentation. For example we may have a raga whose movement ("gathi, nade or chalan") is $S,G,R,G,M,P,D,P,S---S,N,D,P,M,G,R,S$ (say). With this additional criterion we may now define

  Raga = Scale + Movement.

  The movement here simply refers to arrangement of the notes in the ascending and descending order. Up to this, classification is precise and may be sufficient for many ragas–for example raga Garudadwani or Kadanakutuhala which involve complicated arrangement of notes. In any book on Carnatic music a written composition begins with the name of the composer, the name of the raga and the notes involved and their arrangement. Increasingly, even in Hindustani music this method is being followed.

- In Carnatic music scale and movement are not always sufficient to define a raga. A raga like Atana or Deva Gandhari or Arabhi can not be played or sung just with scale even if the complicated movement or arrangement of notes is given. (Similarly Darbari in Hindustani music can not be defined by scale and movement alone.) Based on compositions in such ragas, one arrives at distinct phrases or words made up of a set of notes which may only be derived using the compositions in these ragas. Interactions and transitions between notes or phrases may involve glides and oscillations of many types together called "Gamaka" which literally means ornamentation or embellishment. The composition is often the scaffolding for the raga which, once defined, may be removed leaving the raga to stand on its own. This then gives us the definition

  Raga = Scale + Movement + Gamakas/Phrases

  By and large in Carnatic music this appears to be sufficient; the last element combines many subjective elements of the performer in terms of the various Gamakas (glides, pulls, oscillations etc) even if based on compositions. Gamakas provide various approaches from one note to the next some times even hovering in the continuum between the notes of a scale. While this may be learnt by listening, it is hard to put it precisely in the form of a written notation.

- It gets a little more qualitative when we go over to Hindustani Music. A major difference between Hindustani and Carnatic music in the delineation of a raga is the emphasis on individual notes in Hindustani music. Unlike in Carnatic music where the raga is introduced often using the allowed phrases (based on compositions), in Hindustani music every note of the scale is introduced often individually, phrases are introduced next with additional accents.

  For example even when the scale and movement are the same, the ragas are distinguished based on accents or equivalently emphasis or stress. Ragas Marwa and Puriya have exactly the same set of notes S,R_1,G_2,M_2,D_2,N_2,S' and movement but in Marwa the accent is on R, D while in Puriya it is on M, N. Interestingly this change of accent does make a difference in listening experience. We may thus write

  Raga = Scale + Movement + Gamakas/Phrases+accents(emphases)

  An analogy may be drawn in ordinary language: A given sentence in English (say) spoken by speakers of Indian, Italian, African languages sounds very different even when they are speaking exactly the same set of words. Similarly the same set of notes and movement may sound different with different emphasis on the notes.

> Nevertheless, in most common ragas both in Carnatic and Hindustani music share many common elements except where accents are concerned, which is a severely subjective element. So much so, even when the same raga is rendered by different performers trained in different schools (gharanas), the accents may change. As it turns out, musicians from different schools may even claim their version to be more “authentic" than those of others!

Some of these points are illustrated in the figure [2.6] choosing the simple scale/raga corresponding to Shankarabharanam in Carnatic or Bilawal in Hindustani. As noted before this would correspond to C-major scale if the tonic is chosen as the key C. It is averaged over many performances after tonic normalisation since in Carnatic and Hindustani the tonic is variable. The red line denotes the European music performed in C-major with precise narrow peaks at the note positions. The Hindustani raga, the green line, shows some variability around the note positions but the Carnatic raga, in blue, shows even more variability around the note positions emphasising the fact that scale or set of notes alone is not sufficient to define a raga. The width around the mean position of notes is a sum total of all the other features like pull, glides, oscillations, etc. There is also some secondary peak positions between note positions adding to the subjective, artist dependent, elements in the improvisation of a raga. The shift in some peak positions between the red line on the one hand and the blue/green lines is due to the difference between tempered and natural scales.

2.3.2 Raga Origins

The description of a raga as given above is akin to outlining grammar of a language. However, the grammar becomes evident after the language is fully developed. This applies to ragas also. While the criteria given above for describing a raga is more or less complete, how does a raga come into being? Does one really start with the scale and proceed to introduce additional attributes? It is possible that some ragas may have evolved in that fashion– the popular ragas like Shankarabaranam or Kalyani (Bilawal or Yaman in Hindustani music) and many others may have originated through this route. The attempt at inventing quatro-tonic ragas by Balamuralikrishna also falls in this category. In addition, the Melakartha Ragas by definition are given once the scale is specified even when in actual elaboration gamakas have a role. However, it is unlikely that many of the ragas which have complicated movement and accents could have come about through this microscopic route.

In fact the origin may just be through the opposite, the macroscopic route: The poetry of Bhakti movement period (starting from about 12th CE) in many Indian languages, though simple, is very musical. Even in ordinary recitation there is lilt and music in it. Superposed on this is the emotional content which is high in the Bhakti literature. Thus it is obvious that the melodic outlines involving complicated emotions were born in such a “petri dish". While the vedic origins provided some basic elements through the chants, the Bhakti movement provided the impetus to move towards sophisticated melodic outlines to reflect the many moods of the devotees. Many such melodic outlines may have morphed in to ragas. It hard to separate the origins from the folk music and the bhakti movement. In many cases they coincide while folk music encompasses much larger themes starting from simple daily routines to the deeply spiritual[21].

The earliest poets who influenced the Bhakti movement in South India were the Nayanars (devoted to Shiva) and Alwars (devoted to Vishnu) who produced a rich body of poetry before the 10th century in Tamil. The Shaivaite Bhakti movement spearheaded by the social reformer and saint Basavanna produced a huge body of literature called “Vachana Sahitya" by him and many of his followers like Akka Mahadevi[22]. The language of these verses is very simple, similar to the language spoken by ordinary people and musical. The Bhakti poet Jayadeva composed the Gitagovinda some time in 12th century. The language is Sanskrit with beautiful lyrics describing the romance between Radha and Krishna. The writing is simple, lyrical and rhythmic.

[21] As late as 1950s, we see this interaction. For example, Hindustani musician Kumar Gandharva was inspired by the folk melodies which he used in many of his compositions in ragas that were either unknown or new. His later compositions had a spiritual bent derived from such folk melodies.

[22] A K Ramanujan has translated a selection of these free-verses into English in a book titled *Speaking of Siva*. The book also contains a detailed introduction to the historical setting of the times.

Approximately every 12 stanzas are expected to be sung in a particular raga which is also mentioned in the text [23]. We come across names of ragas like- Malava, Gurjari, Vasantha, Ramakali, Karnata, Desharavya, Bhairavi, etc. Some of these ragas mentioned in the text, Vasantha or Bhiravi, for example, are popular even today both in North and South Indian music genres. There is no notation as such, but the melodic forms in which the poetry is set is given. The devotional poems of Jayadeva may have inspired a class of compositions called padams in South India composed in Kannada, Tamil and Telugu languages. These are forms which include an opening line which becomes a refrain or pallavi, and a following line which is called anupallvi or a sub-refrain and a set of verses called caranams which literally means feat. The pallavi and anupallavi are repeated after every caranam. This tradition which began in the second half of fifteenth century became central to the South Indian classical music from fifteenth to eighteenth centuries[24]. Devotional padams were composed in Telugu by the poet Annamacharya or Annamayya based in Tirupati. Later poets like Kshetrayya go beyond the poems addressed to temple gods in which the god may be a lover articulating a romantic, even erotic, experience devotional character notwithstanding. We know very little about Kshetrayya who may have been associated with Nayaka rulers of Madurai.

Later Bhakti poets were also singers of repute. Most important of them was Purandara Dasa(1484– 1565) who is regarded as a father figure in Carnatic music. He was a Vaishnava saint and a social reformer. He is known to have formalised lessons for beginners in Carnatic music which is followed even today. He wrote the lyrics as wells as composed music for these lessons. However, it is not clear how he rendered his poetry, beyond the basic lessons, into music. Many versions we hear today are less than a century old set to music by musicians like M L Vasanthakumari. Some folk versions do survive, probably in their original form. Theater artist and director, B V Karanth used the folk forms in some of the plays directed by him. Purandaradasa was also a contemporary of another well known Bhakti Poet Kanakadasa. Kanakadasa's poetry is also set to music and is performed in concerts. Another contemporary is Vysaraya Thirtha who composed the famous Krishna Nee Begane Baro which is the bread and butter stuff of South Indian Dancers. Many vaishnava poets of this period collectively produced the Dasa Poetry which is a huge body of literary work in Kannada characterised by lyrical poetry which is also very musical. Again we do not know the original melodic outlines of the Dasa music. Bhadrachala Ramadasa continued the Dasa tradition by composing in Telugu in the 17th century. He may have even influenced the doyen of Carnatic classical music- Saint Thyagaraja.

The process started by the Bhakti movement on large scale was given a formal structure and perfected during the period of Trinity in Carnatic Music- Thyagaraja, Muthuswamy Dikshitar and Shyama Shastry. Of the three, Thyagaraja is the most prolific followed by Muthuswamy Dikshitar. Thyagaraja is said to have composed nearly 700 songs set to ragas[25]. Many of these ragas did not exist before or were not used as much. It is through the compositions of Thyagaraja, especially through the beginning pallavi phrases, that the melodic outline is defined. From the melodic outline, we have to work backwards to get to aspects like scale, movement and accents to analyse the various dimensions of the raga. The number of such examples are huge. In fact in many ragas that were given life by Thyagaraja, like Jayantasena, only his compositions exist. Even today it is a common practice in Carnatic music that the delineation of some of the ragas almost contain the melodic outlines found in Thyagaraja's compositions. Because of the peculiar origin in the Bhakti tradition as found in many of Thyagaraja compositions, many performers emphasise the need to have the emotional and spiritual connect in order to outline the melody as well as the compositions. In fact it is difficult to imagine the music of M S Subbulakshmi without her emotional connect with the words in the compositions. It is this connection that enhanced the richness of her music and many like her.

[23] See for example-"The Gitagovinda of Jayadeva: Love song of the dark lord", by Barbara Stoller Miller.

[24] See for example *When God is Customer* Edited and translated by A K Ramanujan, V Narayana Rao and David Shulman, University of California Press (1994).

[25] Thyagaraja's compositions, more than 700 known, are in more than 200 ragas of which more than 60 ragas are new in the sense he was the first one to create compositions in these ragas even giving names to most of them. Thyagaraja created many types of musical compositions and forms including dance-dramas. Historically, the only parallel we can think of, in terms of creative bursts, is in the mathematics of Ramanujan many decades later. Ramanujan's work lead to so many new sub-fields of mathematics and even today new results are emerging from his early works.

Many musicians and connoisseurs alike tend to downplay the influence of Bhakti (devotional) aspect of the Carnatic music. The musical origins in the Bhakti tradition is merely a scaffolding for some. Once the scaffolding is removed, the pure music based on the melodic outlines has a life of its own. There is some substance in this secularisation of Carnatic music since it is indeed possible to extract the pure musical content out of its original context. However, many performers indeed experience or even need, the emotional aspect of Bhakti enshrined in the words of the composition. The music for them is more than a mere melody– it is divine!

To summarise we may think of Raga as a kind of “Pizza" (using the analogy of Tony Joseph in the book *Early Indians*). The set of notes forms the base and cheese is the movement. These two elements are essential for any music or pizza for that matter. The toppings depend on regions, kind of veggies, fruits etc, which in the case of a raga may be phraseology/accents which could differ from region to region or even from performer to performer introducing the subjective element that in the end gives the best “pizza" that a performer can offer. We are not even talking about sprinkling salt and pepper which may also find an analogy in this musical pizza.

2.3.3 Structure of Compositions

We have already commented on the structure of compositions called padams beginning with Annamacharya which became central to the musical compositions in all the South Indian languages. The compositions involving many emotional elements are the incubators of many ragas. In this sense the period of trinity marks a separation from earlier period of Purandaradasa (1484-1564) and Ooothukkadu Venkatakavi (1700-1765) who composed in Kannada and Tamil respectively (predominantly but also in other languages).

While Purandaradasa is considered the father of modern Carnatic music, what we have from him are a large number of poems not only in praise of his personal god (Purandara Vittala) but also a commentary on the social conditions of the day. He was a social reformer who spread the message criticising the rigid orthodoxy of the times through his poetry and music. There are poems which are severely critical of the caste system, untouchability and the meaningless rituals while at the same time extolling the virtues of Bhakti. We do not know if he himself had set his poetry to music and if so how. As mentioned before, many were set to music in the last century or so by well known musicians. Many versions of his poetry are available through folk musicians which sound very different from the classical versions.

One thing is certain: Purandaradasa did organise the basic teaching elements of Carnatic music which is followed even today and in all probability would have set his poems to music. In this sense, he is the earliest composer in Carnatic music even though we can not identify all the elements of his music historically.

Venkata Kavi on the other hand immediately preceded the trinity and many structural elements that we see in the compositions of the trinity may have been influenced by his compositions. Venkata Kavi’s work however was relegated to the back ground while the work of the trinity dominated the music scene. There is a revival of Kavi’s compositions in the last decade or so.

The structure of Thyagaraja compositions are rather well defined (with honourable exceptions)- the entire composition is framed within the Pallavi (literally meaning repetitive) or refrain. Within this is Anu-pallavi (minor-refrain) which frames the Caranam whose last lines melodically mimic the anu-pallavi. There may be many caranams. The pallavi lays out the melodic outline and the thematic outline of the composition. It is elaborated in the anu-pallavi and concluded in the caranam/s. Most of the compositions are first person accounts of the Thyagaraja’s relationship to his favourite god- Rama in Telugu. This is one important feature of all the Bhakti music where the musician addressed the god directly in first person. Therefore the music has a huge emotional content, some times bringing to fore day to day experiences, difficulties, frustrations and even ethical dilemmas. The emotional outpouring takes different forms based on the mood (happy, sad, meditative, etc) and hence myriad collection of complex melodic outlines. The fact that the origin of so many ragas are attributed to Thyagaraja has its roots in this aspect of his music. *It is the music of the heart.* That is why his music has such a great appeal. [26]

[26]This is not surprising. The development of Blues in Americas has a similar origin- the Blues music grew out as an expression of the sufferings of the Blacks. Most of the early compositions are first person accounts of the huge amount of pain they were going through. The

In Dikshitar compositions we see some elements of the above but also some important differences- while the pallavi and anupallavi may follow the above pattern of Thyagaraja compositions, the caranam may follow a very different pattern with its conclusion involving a set of rapid enunciation in second speed. Again, the caranam may have many stanzas even though each is sung the same way. Nevertheless, the pallavi frames the whole composition (which is true in general for all composers). In contrast to Thyagaraja's compositions, Dikshitar's compositions have a more intricate structure. While being spiritual and meditative, the content also involves a clever play of words with many twists and glides. Dikshitar was influenced by the Dhrupad tradition later in his life when he spent some time in traditional centres of music in the North of India. This influence is evident in many of his compositions.

A careful observation of compositions of Thyagaraja and Dikshitar brings out other differences between them- while Thyagaraja was an vocalist, Dikshitar's main musical outlet was through veena even though the instrument was a proxy for the voice. In Thyagaraja compositions the consonants are interspersed by long periods of vowelisation akin to raga elaboration. These vowel intervals vary within the composition. On the other hand in Dikshitar's compositions consonants tend to appear at regular intervals and periods of vowelisation are shorter. To see this it is necessary to sing the compositions at the same tempo so that the intervals may be compared.

While this may appear to be a minor difference, actually it is not since Dikshitar composed songs that may be played with comfort on a veena. The presence of consonants at regular intervals helps in introducing plucked sounds. The sustenance of the sound on a veena is rather short when compared to a voice[27]. Hence the need to introduce consonants at regular intervals. This allows the performer to present such composition in the same way an vocalist would sing especially when they are presented in a slow tempo. In a oral tradition, the instrumental music does not develop or stand on its own. It will be judged by how well the music resembles the vocal music.

The compositions in the music of the north also vary in style and substance. While Dhrupad compositions are spiritual and solemn, Hindustani music has many different varieties like Khayal, Tarana, Thumri, Bhajan, Ghazal etc. Khayal, meaning imagination, is a style of elaboration that is popular in vocal music. The style also features alaap and Taan as part of the elaboration. The composition itself is short when compared to the compositions now in use in Carnatic music. The Thumri, Bhajan and Ghazal are considered semi-classical forms. Some of these forms are also in use in both North and South Indian music concerts.

Unlike in the Carnatic music where instrumental music essentially follows the vocal music, in the Hindustani system instrumental music often differs from the vocal music. The vocal music is dominated by the Khayal style of singing which goes to elaborate compositions after a short alaap in different tempos. On the other hand, in instrumental music, outlines of the raga are first traced through an elaborate alaap followed by Jod and Jhala which is akin to elaboration of a raga but with a minimal rhythmic base-laya. In this, the presentation follows the example of Dhrupad. This is then followed by a brief composition which is a simple outline of the raga and taans in two to three speeds. An approximate equivalent of this is the Raga- Tana- pallavi part of a concert in ?? music. However the instrumental elaboration can be quite expansive. Often instrumentalists, especially violinists, tend to follow both systems. The semi-classical forms may be common for both vocal and instrumental music performances.

Even within these several forms of rendition there are variations which are some times due to their origin. An interesting concept in the music of the north is the concept of "gharana" (lineage) which links a performer to his or her lineage. Every gharana has a distinct style of presenting its own home-grown compositions which are often guarded zealously. While the concept of lineage is also there in Carnatic music, it is not adhered to as strictly as in the north. The names of the gharanas are geographical indicators which tell you where they originated- thus we have the Khayal gharanas with names like Kirana, Agra, Gwalior, Jaipur, etc; Dhrupad

melodic outlines of these expressions became the standard of Blues music. The music of King or Lightening Hopkins or any other musician during the early part of last century is rooted in the deep emotional experiences just like our Bhakti poets or folk musicians in India.

[27] This is not a problem in recent times since the amplification has changed the nature of music production as also the use of violin and other instruments in Carnatic music.

gharanas carry names like Betia, Darbhanga, Bishnupur, Dagar etc. There are also gharanas in instrumental music as well as in semi-classical music. The exploration gharanas is in itself a huge topic. It suffices to say that a gharana indicates a system of teaching, performance as well as appreciation.

Acknowledgement:

I am grateful to Padma Arunachalam who started me on this project with many questions, Hema Murthy for detailed discussions and critical comments. Thanks are due to Preeti Rao, Venkat Viraraghavan, Raj Narain and G Rajasekaran who went through the draft and provided critical comments and inputs.

Appendix-A: String vibrations

A brief introduction to vibration of strings/columns is given below for those who may be interested. This will help to understand why natural scale is “Natural". Consider the vibrations of a string fixed at both ends. The sound waves are created by transverse standing waves in the strings. In wind instruments the sound waves are created by longitudinal standing waves in air filled tubes. Voice also comes under this category. The different sounds produced by different instruments (including human voice) are distinguished by the “timbre" or quality of sound.

The vibrations of the string are described by the wave equation in one dimension

$$\frac{\partial^2\phi}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2\phi}{\partial t^2},$$

where $v = \sqrt{mT/L}$ is the wave velocity, L is the length of the string fixed at both ends $x = 0, x = L$, T is the tension and m/L is mass per unit length. The solution for a string that is plucked at some point and released at time $t = 0$ is given by

$$\phi(x,t) = \sum_{n=1}^{\infty} a_n \sin\frac{n\pi x}{L}\cos\frac{n\pi vt}{L} = \sum_{n=1}^{\infty} A_n \cos\frac{n\pi vt}{L},$$

where a_n is the amplitude of the $n-$th harmonic which is determined from the initial shape of the string that is plucked. We have assumed the string to be ideal hence no attenuation. For our purposes here this is enough. The fundamental frequency is defined as

$$f = \frac{v}{2L}$$

and the resonant frequencies are $f_n = nf$. Changing L will change the fundamental frequency (moving along the frets in a veena or a guitar will do this job). The solution for a given fundamental frequency f is then

$$\phi(x,t) = \sum_{n=1}^{\infty} A_n \cos(2\pi nft).$$

The standing wave corresponding to the fundamental and its first few excitation are shown in the figure 2.7. Basically the positions of nodes in the standing waves correspond to simple integer ratios- in the second harmonic the node occurs at 1/2, in the third harmonic the nodes occur at 1/3, 2/3 creating a pattern of simple integer ratios. When a string is plucked at the nodal points, the corresponding harmonics are absent (Young-Helmoltz theorem).

Precisely these ratios occur in the natural scale. For example consider the note $M_1 = 4/3$ with respect to S which is taken as the fundamental. Holding one end fixed at M_1 instead of S creates a new fundamental frequency which is $4/3f$. When the string is plucked at holding M_1 fixed will give a solution

$$\phi(x,t) = \sum_{n=1}^{\infty} A_n \cos(2\pi n4ft/3).$$

The third harmonic $n = 3$ of this solution vibrates with the pattern $\cos(2\pi 4ft)$ which corresponds to the solution corresponding to the 4th harmonic of the fundamental corresponding to S. They are said to be resonant or in

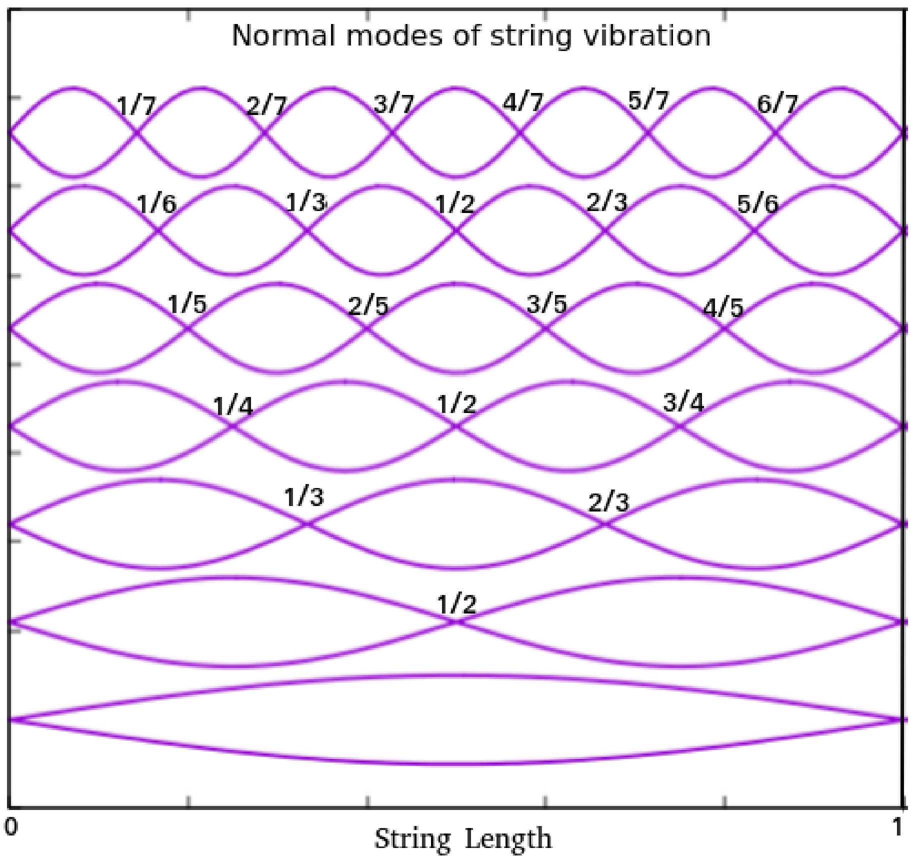


Figure 2.7: Standing wave pattern of the normal modes in string vibration. The first seven harmonics are shown here. Note that the position of nodes, relative to the fixed end at left, forms a Farey Sequence.

general consonant. In general if the new frequency is set at pf/q, where p,q are integers as in natural scale, we have the vibrations given by

$$\phi(x,t) = \sum_{n=1}^{\infty} A_n \cos(2\pi npft/q),$$

and the $q-$th harmonic of this is consonant with $p-$th harmonic of the original fundamental frequency f.

Since all the ratios of the natural scale are of the form p/q (rationals) they are consonant with the fundamental frequency. The resonance can be heard by the ear easily if both p,q are small integers.

We close this appendix by commenting on an interesting connection to Farey sequence of rationals. A Farey sequence can be generated by dividing the interval between 0 and 1 sequentially in the ratio 1/q and 1-1/q where $q \geq 2$ as shown in figure where the sequence is shown up to q=8. The simple ratios of integers are distributed symmetrically around 1/2 in the interval 0-1. We can see that many of the ratios used in the natural scale, after shifting the ratios between $1-2$ to $0-1$, are found in the Farey sequence though not all. However the Farey sequence corresponds precisely to the node positions of standing waves of string harmonics as can be seen from the figure showing string vibrations.

0														*1*
0							*1/2*							*1*
0						*1/3*	*1/2*	*2/3*						*1*
0					*1/4*	*1/3*	*1/2*	*2/3*	*3/4*					*1*
0				*1/5*	*1/4*	*1/3*	*1/2*	*2/3*	*3/4*	*4/5*				*1*
0			*1/6*	*1/5*	*1/4*	*1/3*	*1/2*	*2/3*	*3/4*	*4/5*	*5/6*			*1*
0		*1/7*	*1/6*	*1/5*	*1/4*	*1/3*	*1/2*	*2/3*	*3/4*	*4/5*	*5/6*	*6/7*		*1*
0	*1/8*	*1/7*	*1/6*	*1/5*	*1/4*	*1/3*	*1/2*	*2/3*	*3/4*	*4/5*	*5/6*	*6/7*	*7/8*	*1*

Further iteration may be done using the ratios $2/q, 1-2/q, q \geq 4$. For example with this we get

$$2/12, 2/11, 2/10, 2/9, 2/8, 2/7, 2/6, 2/5, 2/4, 3/5, 4/6, 5/7, 6/8, 7/9, 8/10, 9/11, 10/12$$

which is the same as

$$1/6, 2/11, 1/5, 2/9, 1/4, 2/7, 1/3, 2/5, 1/2, 3/5, 2/3, 5/7, 3/4, 7/9, 4/5, 9/11, 5/6.$$

Notice that now there are new ratios between the natural scale ratios. For example between 1/3 and 1/2 now we have 2/5=0.4. Adding 1 to this we get 1.4 as a ratio between M_1 and P. This is indeed very close to $M_2 = 17/12 = 1.416$. The main difference between the natural scale and Farey sequence is the symmetry of the later around 1/2. It always generates equal number of ratios on either side of 1/2. This is the same as having same number of notes between S–P and P–S which is not the case in the 12 note octave system. The Farey sequence can not be used for producing melodies in conventional instruments like Veena or Sitar due to the design of frets, but electronic musical instruments can easily generate the Farey sequence melodies. It would be interesting to see if it can produce a new Genre of “Farey" Music!

Chapter 3

Salient features of melody, rhythm and lyrics in Carnatic Music

Venkata Subramanian Viraraghavan[a,b] and Sridharan Sankaran[a,c]

[a]Research and Innovation, Tata Consultancy Services Limited, Bangalore, India and
[b]Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai, India
[c]Department of Computer Science, Chennai Mathematical Institute, Chennai, India

3

Abstract

Carnatic music (CM) is a form of music prevalent in South India. In this article we consider the three main areas of CM: Melody (*rāga*, *gamaka*), Rhythm (*tāḷa*) and Lyrics (*sāhitya*). We introduce these three areas and explain the intricacies that pose challenges to the computational analysis and synthesis of CM.

The description of melody in CM is monophonic, and is bound by a *rāga*-system. Melody is constituted by *svara*s, which are approximately equivalent to musical notes. However, *svara*s are enriched and characterised by continuous pitch movements called *gamaka*s. These *gamaka*s are learnt by example in a largely oral tradition. We trace the evolution of the terms *svara*, *gamaka*, and *rāga* in CM musicology and build a landscape that reflects their hierarchy. This landscape serves to define problems in the computational analysis and synthesis of CM melody.

Rhythm in CM is a way of keeping time by counting in beats and units per beat. A *tāḷa* corresponds to a specified pattern of beats and units. When a melody is bound by rhythm, the syllables in the melody have a simple relationship to the units of rhythm. Rhythmic performances typically accompany a main melody. They can range from simply counting units, to sophisticated patterns, whose lengths are co-prime to the underlying *tāḷa*. We explain these concepts and define a few problems in the computational analysis of CM rhythm.

Syllables with extended vowels or nasals carry melody in vocal CM. These syllables can be meaningless, but standardised, or be part of meaningful lyrics. Instrumentalists too imitate these syllables when playing the same tunes. For lyrics with meaning, constraints from poetry (i.e. rhyme) are also adhered to. We show, with examples, how the vowel extensions of syllables carry CM, and describe briefly the rules that govern rhyme in CM lyrics.

3.1 Introduction

Carnatic music (CM) is a form of music prevalent in South India. This article starts with a brief introduction to CM and an overview of the three broad areas of melody, rhythm and lyrics in CM. Detailed discussions of the distinguishing aspects of each area constitute the focus of this article. We also touch upon how these aspects help define research problems in the computer analysis and synthesis of CM.

A basic knowledge of musical concepts, such as note, beat, and lyrics, is assumed, but familiarity with CM is not important. The key CM-terms are defined before use (e.g. *rāga*, *tāḷa*, *kriti*, *sāhitya*, *gamaka*). Other CM-specific terms are translated to English, and the original terms are spelt in brackets using necessary diacritical marks. This enables readers familiar with Carnatic terms to relate to the concept, while other readers can ignore these original terms.

CM is an oral tradition. A teacher (*guru*) sings or plays the music and the students try to sing it or play it on an instrument. Repeated practice is needed for the student to grasp and reproduce the music. The teacher listens and corrects these attempts using qualitative descriptions and may relate to pitch (such as 'raise the pitch slightly', 'the last syllable was too high in pitch') or rhythm (such as a specific 'syllable needs to be shorter'). As the student matures and understands the positions of notes and *svara*s, the corrections may become more granular (e.g. the *gamaka* for a *svara* needs one more pitch movement). However, at no point does notation take precedence over learning by example. Even in the advanced cases, a granular description is used only if the student does not instinctively grasp the pitch and rhythm information[1].

Basic training in CM starts with exercises called *saraḷivarusai, jaṇṭavarusai* and *alankāra*s, which have only sequences of musical notes set to the basic rhythmic patterns. Short songs called *gītam*s, which also include *sāhitya*, conclude the basic training. Most of the basic training material were systematised in the 16$^{\text{th}}$ century by Purandara Dasa, also referred to as the '(grand)father of [Indian] music' (*sangīta pitāmaha*). Renditions taught during basic training are not part of professional concerts. A *varṇam* is an advanced exercise-rendition, and is sophisticated enough to merit a presence in professional concerts. It is often the opening rendition. Usually, it takes several years to a decade of training to perform in professional Carnatic concerts.

3.1.1 Typical Carnatic concert and renditions

In CM concerts (*kuccēri*), a lead musician, who may be a vocalist or an instrumentalist, has one or more accompanists. The accompanying instruments (*pakka vādyam*) are, usually, a drone (*tamburā*) and one instrument each for melody (e.g., violin) and percussion (e.g., *mṛdanga*). It is also common to have an additional percussion accompaniment (e.g. *ghaṭam*, *khanjirā*, *mōrcing*, *tavil*). Electronic and digital versions of the drone (*tamburā*) are also common today. All vocal/instrumental sources barring the drone (*tamburā*) have one microphone each. In professional settings, all sources of audio are fed to a mixer. Thus, multi-track and advanced spatial-sound configurations are possible. In practice, most concert audio recordings are mono. The photographs of a typical Carnatic vocal concert and of an instrumental concert are shown in Figure 3.1. The lead musician chooses the renditions to be performed, which are around 10 in number for a two hour concert. A rendition consists of composed music and often, extempore elaboration. In the composed segments, the musicians perform in unison; in the extempore segments, they perform by taking turns.

The most common form of composed music is called a *kriti*, which has several sections (*pallavi*, *anupallavi*, *caraṇam*). The sections have melody and lyrics bound by a rhythmic pattern. Each rendition is specified by giving the name of the *kriti*, i.e. the opening line of its *sāhitya* (lyrics), its *rāga* (melodic concept), and its *tāḷa* (rhythm pattern). Three composers, Thyagaraja, Mutthuswami Dikshitar, and Shyama Shastri, are credited with creating definitive examples of *kriti*s. These composers are called the Trinity of Carnatic composers. Although the *kriti* is the most popular type of composition, there are other types too. A *tillānā* has a predominance of rhythmic syllables in its first half, and *sāhitya* in the second half. A *tillānā* is usually the penultimate or third-last rendition of a Carnatic concert. *Padam*s and *Jāvaḷi*s are songs that depict separated lovers. In recent decades, *abhang*s, which are devotional songs in the Marathi language, are also included in Carnatic concerts.

Extempore segments usually occur in *kriti*s and can last from a few minutes to nearly half an hour. There are several types of extempore segments, which differ in their focus-areas. An *ālāpana* is an extempore elaboration of the *rāga*, using standard, but meaningless syllables (Section 3.4.1). It is not bound by rhythm and the emphasis is on *rāga* alone. In *kalpanāsvara*, *svara*-names are pronounced, and the emphasis is on *rāga* and *tāḷa*. A *niraval* involves repetition of a chosen line of the lyric and introduces variations in the melody. A *niraval* emphasises *rāga* and *tāḷa*, but is constrained by the *sāhitya*. In a *ślōka* (lyrics in Sanskrit), *viruttam* (lyrics in Tamil), or *ugābhōga* (lyrics in Kannada), the vocalist chooses one or two verses from poetry and sings them in one or more *rāga*s without being bound by a *tāḷa*. Thus, the focus is on the *rāga* and *sāhitya*. The most advanced extempore rendition is the *rāga-tānam-pallavi*. The lyrics are usually short and the focus is

[1]Consequently, the notation for CM is not yet descriptive enough to reproduce pitch curves and *rāga*s faithfully. We discuss this point at length in Section 6.3.

Figure 3.1: Typical Carnatic concert settings.
(Top) A vocal concert. The singer and each accompanist have one microphone each. There is no microphone for the drone (*tamburā*). A *mṛdanga* (sometimes spelt *mridangam*) usually has an additional microphone for its second surface (not visible).
Credits: Kamala L, https://www.flickr.com/photos/50863590@N00/127406087. Annotations added.
(Bottom) Instrumental (*vīṇā*) concert with only a drone and percussion accompaniment. A contact microphone is attached to the *vīṇā*, and another pair of microphones is used for the *mṛdanga*.
Credits: Jean-Pierre Dalbéra, https://www.flickr.com/photos/72746018@N00/7428171194

Table 3.1: Example Concert listing (Music Academy Concert, 27-Dec-2020[a])

Rendition	**Type**	*Rāga*	*Tāḷa*	**Composer**	**From-to (h:mm:ss)**	**Extempore section(s)**
Vanajaksha	*ālāpana* *kriti* *kalpanāsvara*	Gambhira Nattai	 *Ādi* *Ādi*	Mysore Sadashiva Rao	0:00:47-0:10:26	0:00:47-0:01:20 0:05:56-0:10:04
Renukadevi	*kriti* *niraval*	Kannadabangala	*cāpu* *cāpu*	Muttuswami Dikshitar	0:10:45-0:10:26	 0:13:11-0:14:48
Murugane Ka Guha	*ālāpana* *viruttam* *kriti*	Kosalam	 *Ādi*	Koteeshwara Iyer	0:16:25-0:28:30	0:16:25-0:23:44 022:49-0:23:45
Emani ne	*ālāpana* *kriti* *niraval* *kalpanāsvara* *tani*	Mukhari	 *Ādi* *Ādi* *Ādi* *Ādi*	Subbaraya Shastri	0:29:02-0:59:19 *Ādi*	0:29:02-0:40:00 0:46:30-0:54:39 0:46:30-0:54:39 0:54:39-0:58:49
Saramaina Matalu	*kriti*	Behag	*Ādi*	Swati Tirunal	0:59:23-1:02:15	None
Total (approx.)					62 mins[b]	26 mins

[a] Artises: Vivek Sadasivam (Vocal), V.Deepika (Violin), Ambur S.Padmanabhan (Mridangam).

Link: https://www.youtube.com/watch?v=rZ1oP4qg_z8.

[b] By convention, the duration of a rendition counts includes composed and extempore segments.

on *rāga* and sophisticated *tāḷa* patterns (Section 3.3). One extempore segment called the *tani āvartanam* or simply, *tani*, is devoted to percussion accompaniments. The focus of the *tani* is solely on *tāḷa*.

About fifty years ago, Carnatic concerts lasted three to four hours. Current Carnatic concerts are considerably shorter. A list of renditions of a recent concert is given in Table 3.1. This concert consists of only one type of composition (*kriti*), probably because only one hour was allotted to the concert. Yet, we observe that extempore segments occupy more than a third of the duration. Further, all types of extempore segments barring the *rāga-tānam-pallavi* occur in this concert. This indicates the importance of extempore segments in CM.

In CM, melodic instruments (e.g., violin, *vīṇā*, flute) imitate vocal music as far as possible. Even syllable enunciation is marked either in amplitude (attack for a new note) or in pitch (change in pitch in a continuous contour). This is true even for sections of compositions that have lyrics. Clearly, lyrics themselves cannot be discerned. If the listener is familiar with the composition, they may follow the lyrics in their mind. In this case, the same syllables as in vocal music should be perceived in its instrumental rendition. For unfamiliar compositions, trained listeners perceive the *svara*s that are played. While instrument-specific acoustical properties can be studied, a study of melody need not differentiate between instrumental and vocal music.

For percussion instruments, the emphasis is on aligning to rhythmic patterns. A fixed set of syllables (*collu*) are vocalised (*konnakōl*) according to the rhythmic pattern. A similar set of syllables is used to identify specific strokes of a percussion instrument. The same vocal syllable may be associated with different instrument strokes and vice-versa, but both need to match the rhythm pattern exactly. Usually, a study of percussion instruments is preferred to the study of vocalised syllables. Extempore rhythmic patterns are part of the *tani* in a typical CM concert, which is led by a vocalist or melodic instrumentalist. Percussion-only concerts, with melodic accompaniment playing only a refrain, explore rhythmic patterns more comprehensively, but these are, by far, rarer than typical CM concerts.

Irrespective of the rendition type (*kriti*, *tillānā* etc.) and mode (vocal or instrumental), a *rāga*'s identity[2], the *tāḷa*-definition, and rules of *sāhitya* are fixed. Thus, a study of the three areas of melody, rhythm, and lyrics forms a basis to study CM. Sections 6.3, 3.3, and 3.4 describe these three areas respectively.

3.1.2 Definitions, Symbols, and Acronyms

The definitions of technical terms and their symbols as used in this article are given in 3.2. 3.3 gives the mapping between Carnatic *svara*-names and notes in Western music (WM), and is explained further in 3.2.1.

[2] Strictly speaking, a *rāga*'s identity is not static. A *rāga*'s definition and compositions in that *rāga* influence each other over decades in an iterative fashion.

Table 3.2: Definitions and symbols of basic musical terms.

Term	Definition
Pitch	The perceptual property of sounds that allows their ordering on a frequency-related scale[a]. It is approximated as the measured fundamental frequency.
Tonic	The reference pitch of a melody/rendition, which is usually provided by the drone (*tamburā*). Other notes are derived with respect to the tonic.
Stable note	Musical note with a fixed pitch throughout its duration.
Semitone	Notes in tuning systems with uniform[b] or non-uniform[c] subdivisions of an octave.
Integer semitone	Exactly $1/12^{\text{th}}$ of an octave[b].
Cent	Exactly $1/1200^{\text{th}}$ of an octave.

[a]Definition proposed by [KD07].
[b]Equal tempered tuning.
[b]Examples of such tuning systems are just intonation and Pythagorean tuning.

Table 3.3: *Svara* symbols for notes in an octave with the tonic assumed to be C.

Note number[a]	0	1	2	3	4	5	6	7	8	9	10	11
Carnatic	S	R_1	R_2	G_2	G_3	M_1	M_2	P	D_1	D_2	N_2	N_3
Western	C	C#	D	D#	E	F	F#	G	G#	A	A#	B

[a] For equal tempered tuning, the note's pitch in semitones is identical to the note number. For other tuning systems, e.g. just intonation in CM, the pitch is at most 20 cents away.

3.2 Melody in Carnatic music

Melody in CM is defined for monophonic music, which is consistent with the predominance of vocal music. Unlike polyphonic music where different notes are rendered simultaneously, the same melody is rendered by all CM musicians performing together. If the melodies are different, then only one is rendered at a time. Nevertheless, the melody rendered by each of the musicians individually conforms to a common set of definitions and rules. In practice, concert recordings are not monophonic due to the presence of multiple instruments. In composed segments, the concert musicians perform together; in extempore segments, they take turns.

In CM, melody is set to a *rāga*, which is characterised by *svara*-phrases and *gamaka*s. These terms are explained along with other necessary definitions in Section 3.1.2. We trace the evolution of these terms over four centuries of Carnatic musicological literature in Section 3.2.2 to bring out the differences compared to commonly accepted, but inadequate translations to Western music terms. Further, this evolution offers a landscape that can capture musicological literature leading up to current CM practice. It also serves as a background to understand computational problems related to CM melody.

3.2.1 *Rāga*, *Svara*, Phrase, and *Gamaka* in Carnatic Music

A central characteristic of CM is the *rāga* system. Several musicological texts give different definitions of '*rāga*'. In Volume III of a widely accepted text book of South Indian Music, [Sam58] states that a *rāga* is a "distinct musical entity ... and possesses well-defined characteristics." The same volume also summarizes the definitions from *Saṅgīta Pārijāta* of Ahōbala (dated ∼1610 C.E.) as:

> A combination of svaras capable of pleasing the ear constitutes a raga.

and from Mataṅga's *Bṛhaddeśī* (dated 600 C.E. to 800 C.E.) and Śārṅgadēva's *Sangīta Ratnākara* (dated ∼1200 C.E.) as:

> ... a raga is that which is beautified or decorated by the tonal excellence of svaras and varnas and which decoration gives pleasure to the mind of the listener.

In this article, we exclude the subjective aspects of the above definitions, as well as the term 'varna', whose modern meaning is different. Consistent with the above summaries, CM renditions continue to be described in terms of '*svara*s'. As a first approximation, a *svara* is a musical note in the octave. When *svara*s are rendered as stable notes, the musical symbols for *svara*s map to notes in WM as in Table 3.3. *Svara*s are defined relative to an arbitrary tonic (Table 3.3 fixes the tonic at C). For example, if a performer fixes the tonic Sa at C4, the note R_1 would be C#4. If the tonic is fixed at D4 instead, the note R_1 would be D#4. In contrast to the specification of the 'key' in WM compositions, the performer chooses the tonic according to their vocal/instrument's range of pitches. The only constraint in a concert is that all performing voices and instruments, including percussion instruments such as the *mṛdanga* and *ghaṭam*, must use the same tonic note. These tonic notes can be separated by octaves and are standardised. They are typically chosen from {A2, A#2, B2, C2, C#2, D2, D#2, E2} for male singers, and from {D3,D#3,E3,F3,F#3,G3, G#3,A3,A#3} for female singers. For example, if the lead singer in a concert chooses the tonic as A3, the violin will be tuned to A4, and the *mṛdanga*, to A3. In this example, if the main performer is replaced by a flautist, the flute would be tuned to the tonic A5.

The most succinct characterisation of a *rāga* is in terms of an ascending sequence of *svara*s (*ārohaṇam*) and a descending sequence of *svara*s (*avarohaṇam*). These two sequences may be symmetric as in the example of *śankarābharaṇam*.

ārohaṇam: $S\ R_2\ G_3\ M_1\ P\ D_2\ N_3\ \dot{S}$
avarohaṇam: $\dot{S}\ N_3\ D_2\ P\ M_1\ G_3\ R_2\ S$
where the dot above S indicates that the *svara* is one octave higher. An example of a *rāga* with an asymmetric set of sequences is *sāvērī*.

ārohaṇam: $S\ R_1\ M_1\ P,\ D_1\ \dot{S}$
avarohaṇam: $\dot{S}\ N_3\ D_1\ P\ M_1\ G_3\ R_1\ S$

An example of a *rāga* with an asymmetric, zig-zag (*vakra*) set of sequences is *sahānā*. That is, there are descending sub-sequences in the *ārohaṇam* and/or ascending sub-sequences in the *avarohaṇam*. These are marked by bold entries below.

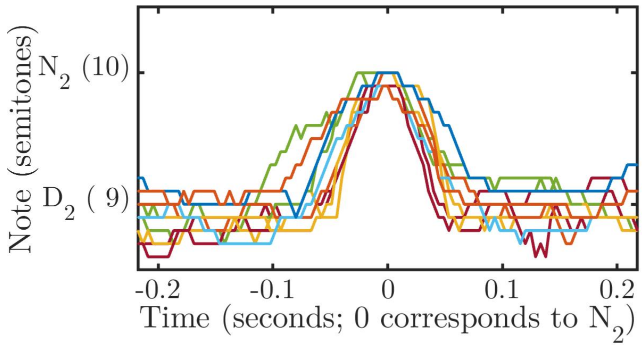
ārohaṇam: $S\ R_2\ G_3\ M_1\ P\ \mathbf{M_1}\ D_2\ N_2\ \dot{S}$
avarohaṇam: $\dot{S}\ N_2\ D_2\ P\ M_1\ G_3\ \mathbf{M_1}\ \mathbf{R_2}\ \mathbf{G_3}\ R_2\ S$

The phrases of the *rāga* may be derived by connecting ascending and descending sub-sequences of the *ārohaṇam* and *avarohaṇam*. Some special phrases that are not derivable from this set are also specified (e.g., $N_3\ D_1\ M_1\ G_3\ R_1$ in *sāvērī*, and $R_2\ G_3\ M_1\ P\ D_2\ N_2$ in *sahānā*).

Since sequences of subsets of the *svara*s in an octave occur in a *rāga*, *rāga*s are likened to scales in Western music (WM). However, in CM, rendering the phrases with stable notes will not have a semblance to the *rāga*. This is especially true for *sāvērī* and *sahānā*. For *śankarābharaṇam*, the major scale may be heard rather than the *rāga*. Hence the need for the qualification 'decorated by the tonal excellence' in the *rāga*'s definition quoted earlier. In the modern context of CM, we interpret this qualification as the use of *gamaka*s, i.e., the use of 'continuous pitch variation.'

Some examples of *gamaka*s are shown in Figure 3.2. In this figure, the y-axis corresponds to pitch. Integer values of semitones mark piano keys. Adjacent piano keys, irrespective of their colour, are separated by one semitone. For more granular measurement, one semitone is divided into 100 cents[3]. If *svara*s are played on a correctly tuned piano, only integer semitones would be possible and the graph of pitch vs. time would have

[3] See Table 3.2 for definitions of 'semitone' and 'cent'. It may be helpful to visualise cents in terms of 100 sub-keys between two adjacent piano keys. Each sub-key has a pitch 1 cent higher than its previous sub-key.



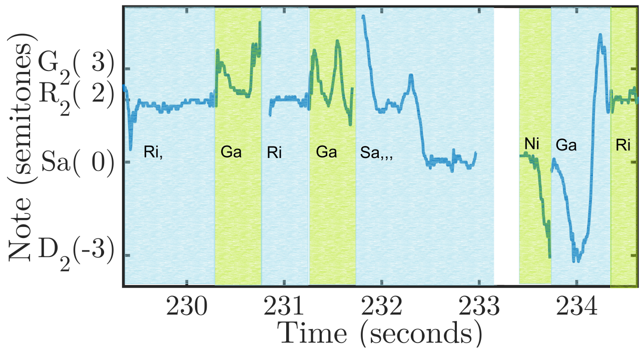


Figure 3.2: (Top) Several examples of pitch curves of a *svara* rendered with *gamaka*, where the *svara*-name depends on the *rāga*; and (bottom) an example of a phrase connected by *gamaka*s, where all pairs of the *svara*s Ri and Ga are connected by *gamaka*s. All examples are from the CM subset [Gul+16d] of the Dunya corpus [PSS13b] released by the CompMusic project [Sri+14c].

horizontal lines at one of the integers. On a violin, however, it is possible to produce pitch values which are in between those of piano keys.

The use of these 'in-between' pitch values differs from one musical system to another. In CM, a single *svara* may be rendered with *gamaka*, as shown by the example at the top of Figure 3.2, or adjacent *svara*s may be connected by a *gamaka*, as shown by the example at the bottom of Figure 3.2. Such continuous pitch variation is essential when a *rāga* is rendered. This is in contrast to WM, where ornamentation is largely optional. For example, the notes in a tune may be sung in an opera with copious vibrato, while the same notes can be played on a piano with no vibrato at all. When it is used, ornamentation in WM is often viewed as a means to approach a note, but the emphasis is still on the note. On the other hand, ornamentation in CM, i.e. *gamaka*s, carry information: if a transient such as the one in Figure 3.2 is removed from a pitch curve, its *rāga* can change. For example, in the phrase Ri,Ga,Ma in the *rāga bhairavī*, Ga may appear only as a transient. If that transient is removed, it may sound like the allied *rāga mukhāri*.

In musicology, 10 or 15 types of *gamaka*s are identified depending on the classification scheme [Sam58]. A few *gamaka*-types (*mudrita* and *nāmita*) are executed by changing a *svara*'s volume or tone (and not its pitch). While these *gamaka*s do contribute to the aesthetic listening experience, they are not mandatory to complete a *rāga*'s definition. Continuous pitch variation, however, *is* mandatory when rendering *rāga*s. In the rest of this article, we use the term '*gamaka*' to signify continuous pitch variation.

As mentioned earlier, melodic instruments need to mimic vocal CM. Thus, CM instruments are created (or adapted) to enable *gamaka*s. Syllables of the lyric are mimicked by introducing a note-onset (or attack) per syllable in the form of a twang, blow or bow, respectively, in the *vīṇā*, flute, and violin. Once the attack is made, the twang is sustained or the blowing/bowing continues. Continuous pitch variation is then produced during this sustained excitation. In a *vīṇā*, the string is pulled at a particular fret[4] perpendicular to its length

[4]See Chapter 1.6 for a more detailed description of the *vīṇā*.

Figure 3.3: Left: South Indian flute. Right: Violin

(away from the player in Figure 3.1). This increases the pitch in a continuous fashion. To decrease the pitch, the string is returned to the open position while still on the same fret. In a south Indian flute (left of Figure 3.3), the number of fully open holes determines the notes $S, R_2, G_3, M_1, P, D_2, N_2$; some examples are given in the left of Figure 3.4. The pitch values between notes are played by partially opening holes, as in the right of Figure 3.4. A *gamaka* is played by controlling the amount of partial opening of one or more holes together. The lone hole, separated from the other eight, is meant for blowing. Another traditional instrument called the *nādaswaram* (or *nāgaswaram*) is similar to the flute, but much longer. In a *nādaswaram*, the variation in the strength of a blow (blow hard vs. blow softly) is also needed to produce *gamaka*s. The violin was inducted into CM concerts in the 19$^{\text{th}}$ century. A few of the piano-note positions for the violin are shown in the right of Figure 3.3. The open strings produce the notes *S* and *P* as marked. The remaining two strings, when open, produce the same notes but one octave lower. *Gamaka*s are played by sliding the finger between notes along the fingerboard, even as the bowing continues. Consequently, it becomes inconvenient to support the fingerboard with a moving palm. Hence the position shown in Figure 3.1 is adopted. Instead of holding the violin nearly horizontally with their hand (as in WM), the violinist rests the instrument on their foot.

While *gamaka*s and notes are crucial to *rāga*s, a phrase-based model of CM *rāga*s also exists. Listeners match entire phrases (called *prayōga*s) to identify *rāga*s. Indeed, the typical initiation into identifying *rāga*s is on the lines of “Song S_1 sounds like song S_2, which is in *rāga* R. Therefore, S_1 is likely to be in *rāga* R.” This view has a musicological backing too: characteristic phrases of a *rāga* are an important part of teaching/learning a *rāga*. The phrase-based model is consistent with the *svara*-based model because phrases are also sequences

Figure 3.4: South Indian flute usage. Left: Two holes closed fully. Right: The second hole is closed partially.

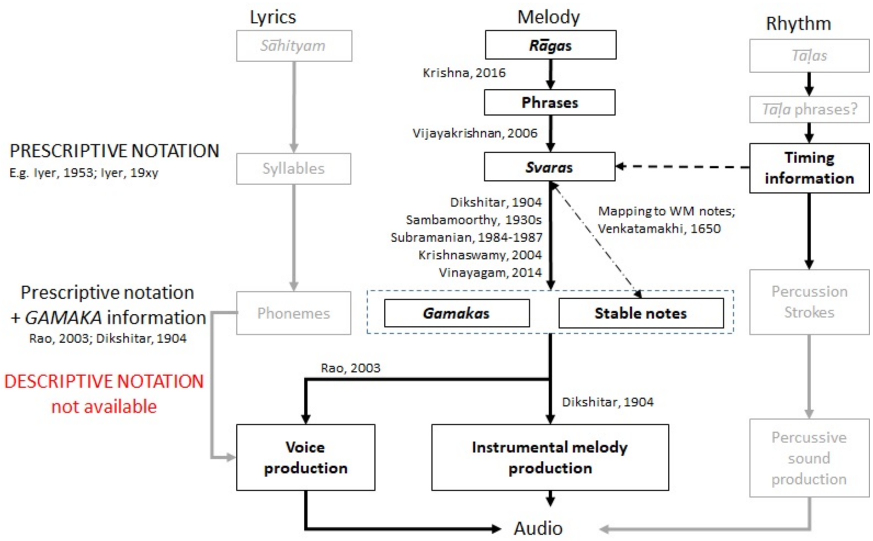


Figure 3.5: Musicological literature arranged according to key concepts in Carnatic music.

of *svara*s (with *gamaka*). A trained listener can recognise individual *svara*s, and the phrases they constitute, to build a picture of the *rāga*, even without having listened to the *rāga* earlier.

The ability to characterise unheard *rāga*s is evidence of a *rāga* being a distinct musical entity. Once the characteristics of the *rāga* have been understood, stating its name is a dictionary-look-up operation. In the two main lineages of CM, beginning respectively with *Thyagaraja* and *Dikshitar*, different names are assigned to the same *rāga*. Such differences are documented [Ram04], and resolving them is not in the scope of computational problems. Therefore, when building databases for CM, a *rāga* is treated as an abstract entity and not classified only by name. For example, a *kriti* by Dikshitar in the *rāga Dēvakriyā* and a *kriti* in the *rāga śuddhasāvērī* by Thyagaraja belong to the same group because the underlying *rāga* is the same.

In summary, phrases constitute a *rāga*, and *svara*s are grouped into phrases. When rendered as stable notes, *svara*s correspond to musical notes as in Table 3.3. However, continuous pitch variation called *gamaka* is needed to render *rāga*s. In Section 3.2.2, we trace the evolution of notation in CM and observe how the *svara*s are not only the musical notes of Table 3.3, but also names for the pitch curves of *gamaka*s.

3.2.2 Evolution of notation in Carnatic music

Figure 3.5 presents a landscape that reflects the hierarchy of the relevant CM terms explained in Section 3.2.1: *Rāga*s consist of phrases, which consist of *svara*s, which in turn can be rendered (i) as stable notes without *gamaka*, or (ii) as *gamaka*s, or (iii) as a combination of the two. This hierarchy proceeds from the cognitively

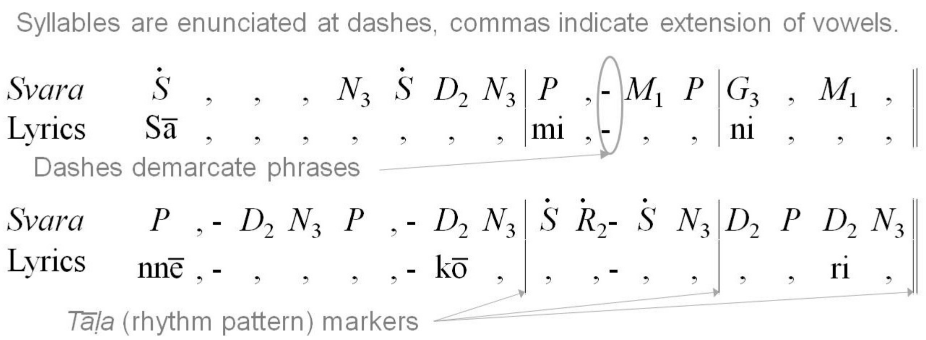


Figure 3.6: Example of a sparse notation for the popular *śankarābharaṇam varṇam*. Notation in black font is borrowed from http://www.shivkumar.org/music/varnams/samininne.pdf, while the description about the notation is in grey font.

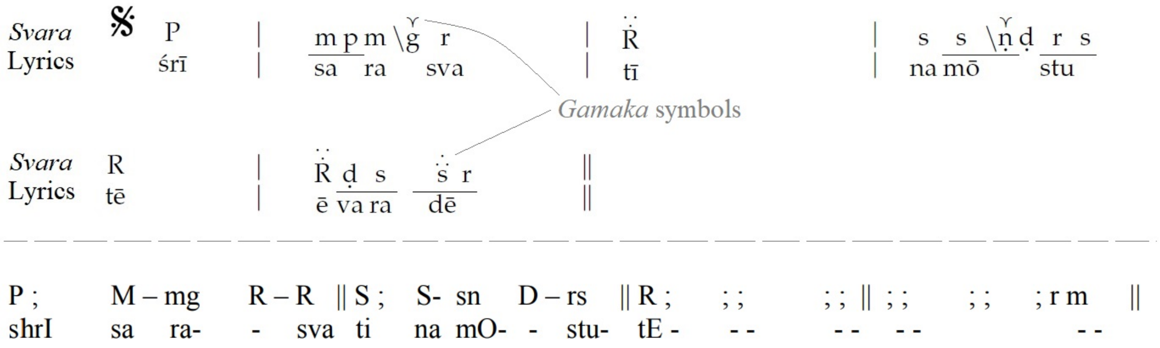


Figure 3.7: Example sparse notations for the *kriti*, the most sophisticated composition-form of CM. The notation above the dashed line, from [Dik04], has *gamaka* symbols, but not the notation below the dashed line (from http://www.shivkumar.org/music/srisaraswati-arabhi.pdf). In the *rāga* (*Ārabhi*) of this *kriti*, the *svara*-symbols S,R,G,M,P,D,N map to $S, R_2, G_3, M_1, P, D_2, N_3$. *Svara*s in upper-case have a duration twice that of *svara*s in lower-case.

most complex (*rāga*) to the least complex (stable notes). For completeness, rhythm and lyrics are shown in grey.

Figure 3.5 also shows CM notation with varying amounts of detail placed at their respective levels. Extant, sparse notation[5] is specified as a sequence of *svara*s with additional timing information for each *svara*. This notation is called sparse because it cannot be used to synthesise music. It does not have information about whether individual *svara*s are to be rendered with or without *gamaka*. It also does not specify how the *svara*s should be connected. However, for *varṇam*s, which are the most advanced practice exercises, the *svara*s to be sung in one phrase are marked by dashes as shown in Figure 3.6.

[Dik04], which is widely read even today, was among the first published works to specify *gamaka*s of individual *svaras*[6]. Notation with and without *gamaka*-information is given for an example-*kriti* in Figure 3.7. The slight differences in the tunes indicate the high complexity of collecting ground truth for CM renditions, even for sparse notation. A further challenge in using sparse notation as ground truth is shown in the bottom of Figure 3.2. The singer chose to render this phrase slightly differently from the widely accepted notation [Dik04; Rao06b]. Such differences are not only acceptable but also appreciated in concerts.

[5] The terms 'prescriptive' and 'descriptive' notation were used first by [See58]. In the words of [Kan07], "Prescriptive notation specifies means of execution". Descriptive transcription, on the other hand, is meant to capture the entire audio that was actually performed. There does not appear to be an equivalent of prescriptive notation in Carnatic music. Although [SWM11a] called Carnatic notation sparse and prescriptive, extant notation does not 'prescribe' as much as it provides a guideline, sometimes only as an aid to memory. We thus use the terms 'sparse notation' (Used in https://www.rajeswarisatish.net/computational-carnatic-music/) and 'descriptive transcription' in this chapter.

[6] [Ram04] observes that the first descriptions of *gamaka*s appeared about 300 years earlier in the *Rāgavibōdha* of Somanatha.

Sparse notation augmented with *gamaka*-information is still not detailed enough for synthesis. This is shown in
https://www.iitm.ac.in/donlab/preview/music/CM_manual_synth.html
This website's example also shows that the typical output representation of WM algorithms is clearly insufficient for CM synthesis. In fact, CM notation synthesised as WM notes has no semblance at all to the *rāga*[7].

Detail that is sufficient for synthesis is expected in a 'descriptive notation,' which is placed lowest in Figure 3.5. The slanting dashed line in this figure highlights the restriction of the mapping between *svara*s and musical notes without ornamentation. The same mapping is ambiguous when *svara*s are rendered with *gamaka*. However, the mapping to *gamaka*s is not arbitrary. To see this, we trace the evolution of the meanings of *svara*-names. A similar evolution in the types of *gamaka*s is relevant. Together, they are used to explain the current research problems related to CM melody.

A seminal work in CM, by the celebrated author, Venkatamakhi, is the *caturdaṇḍiprakāśikā* (CDP), translated from Sanskrit by [Sat01]. Venkatamakhi lived in the same century as composers such as Venkatakavi and Bhadrachala Ramadasa. These composers bridge the gap between Purandara Dasa and the Trinity of Carnatic composers. While the composers created music, Venkatamakhi laid out the theory of CM. He drew upon Ramamatya's earlier work (*Svaramēḷakalānidhi*, dated around 1600 C.E) on a relatively simple *rāga*-classification scheme, and proposed the rigorous 72-*mēḷa-rāga* scheme. (Cite Murthy Sir's chapter) describes this scheme in more detail, but ample explanations would be returned by an Internet search too. A modified version of Venkatamakhi's proposition continues to be the basis of CM-*rāga*-classification. At the time of the CDP, however, there were fewer than 20 *mēḷa-rāga*s. [Sat01] explains other schemes from around the same era. These were more academic in nature and did not find use in practice. Venkatamakhi's scheme found favour with later musicologists and composers, who went on to refine it. Two prolific composers of CM, Thyagaraja (1767-1847) and M. Dikshitar (1775-1835), pioneered systematic composition in most of the 72 *rāga*s and their derivatives. This led to a steady increase in the number of *rāga*s. Although only about 200 *rāga*s may be practised at any era, more than 5000 *rāga*s have been invented in the last three centuries [DP12]. We now describe observations pertinent to the computational treatment of CM.

3.2.2.1 Many-to-many mapping of notes to *svara*-names

The names of *svara*s prior to the CDP are given in the second row of Table 3.4. The notation used in this article is given in the third row. The author of the CDP modified the convention of naming *svara*s, as given in the fourth row of Table 3.4. With this modification, the one-to-many mapping from note-names to positions in the octave (e.g. Ri to semitones 1 and 2, Ga to semitones 3 and 4) became a many-to-many mapping (e.g. Ri and Ga to semitones 2 and 3)[8].

The core contribution of the CDP, namely the 72 *mēḷa-rāga* scheme, occupies six pages in [Rat15]. Yet, no less than three of those pages describe the modified nomenclature for *svara*s, which suggests that the modification was deliberate and unconventional. From a computational analysis point of view, there is now a choice between a musicology-centric alphabet of 16 note-names involving a many-to-many mapping and and an unambiguous computation-centric notation. The former has to be qualified: e.g. [GS19a] had to introduce the complexity of 'forbidden sets' that exclude redundant combinations. By contrast, in a computation-centric notation, G_1, R_3, N_1, D_3 are alternative names for R_2, G_2, D_2, N_2. The additional step of mapping the names according to the *rāga* (e.g. R_2 to G_1 in *varāḷī*), only resolves nomenclature. A similar choice of alphabets affects the treatment of *gamaka*s, which we describe next.

[7]A broader question is whether or not WM transcription algorithms work for CM. Low-level algorithms such as pitch tracking, which were developed in the WM context, may be reused for CM. However, as was discovered by the CompMusic project [Com12], high-level algorithms for WM, such as for transcription, do *not* work for CM. Currently, the evidence is overwhelmingly in favour of pursuing CM-specific algorithms for problems such as transcription. In general, the applicability of an algorithm for one system of music (e.g., WM or CM) on another system (e.g., CM or WM) cannot be assumed. Thus, a comparison with algorithms tested only on WM is unnecessary when proposing CM-specific algorithms and vice-versa.

[8]Venkatamakhi also clarifies that only the *svara*-names in the last two rows of Table 3.4 need to be enunciated when singing *svara*s, and not the mnemonics he used to derive the *mēḷa-rāga* combinations. This practice continues even today.

Table 3.4: *Svara* names for notes in an octave. CDP stands for *caturdaṇḍiprakāśikā*. Rows 5 and 6 have a one-to-one mapping with rows 3 and 4 respectively.

Semitones	0	1	2	3	4	5	6	7	8	9	10	11
Name of note prior to CDP	Sa	Ri		Ga		Ma		Pa	Dha		Ni	
Notation in this article	S	R_1	R_2	G_2	G_3	M_1	M_2	P	D_1	D_2	N_2	N_3
Alternative names proposed in CDP			G_1	R_3						N_1	D_3	
Name of note after CDP	Sa		Ri			Ma		Pa		Dha		
				Ga							Ni	

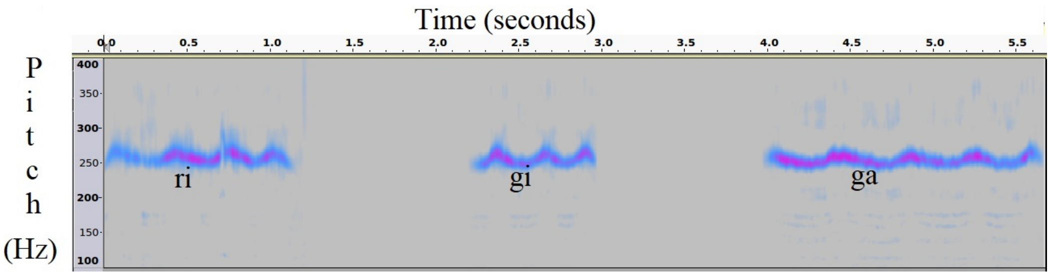


Figure 3.8: Examples of nearly identical pitch curves being mapped to R_2, G_2, and G_1. From a personal collection of live-concert recordings. The *tōḍī* and *varāḷī* examples are available, respectively, at https://youtu.be/zJwTXvFNAZQ?t=52 and https://youtu.be/SLywWVkR-N4?t=27.

In current CM, *svara*-names denote not only note positions in the octave, but also pitch curves. This evolution happened as CM was taught and learned orally from one generation of musicians to the next. There does not appear to be an identifiable point in time or a single treatise that introduced this change, but singing *svara*-names for pitch curves is certainly an inextricable part of CM today. In fact, *svara*s can span more than one octave as observed by [KSS12b]. A consequence of this evolution is the absence of an unambiguous representation of *gamaka*s that can be adopted in computational techniques.

Several attempts at capturing the complexity of *gamaka*s have been made in the past few decades. [Kri04b] partly formalises the mapping of *svara*-names to pitch curves in proposing the use of melodic atoms, i.e. pitch curves for every possible rendition of each *svara* of the third and fourth rows of Table 3.4. This proposal mirrors another recommendation, from [Ram04], to create a repository of audio recordings of the possible renditions of *svara*s and *gamaka*s. While musicians can easily relate to the two proposals, neither enables automatic extraction of *svara*-names from CM pitch curves or audio. There are two impeding challenges. First, demarcating *svara*s in CM is an open problem. Second, multiple names for the same pitch curve are possible. Three such examples are shown in Figure 3.8. The three pitch curves, tracked by the Audacity software, are similar, but map to three different note names: R_2 in the *rāga pūrṇacandrikā*, G_2 in the *rāga tōḍī*, and G_1 in the *rāga varāḷī*. However, Figure 3.8 does not imply that the mapping from pitch curves to *svara*-names is arbitrary. Only a finite subset of possibilities is used in practice, which enables trained students and musicians to start with sparse notation, fill in details according to the context (i.e. choose one of the possible melodic atoms), and render the *rāga*.

The complexity of *gamaka*s in CM can be gauged from the graph on the left of Figure 3.2. The graph shows that the *gamaka* may have continuous pitch variation lasting for around 150 ms. [VAM17a] measured the duration of parts of *gamaka*s (they called them ‘transients’) as ranging from 50 ms to 400 ms. When many such *svara*s with *gamaka*s are concatenated, the amount of information in a descriptive transcription is significantly higher than in sparse notation; about 50% higher according to one estimate [Vir+20a]. Therefore, it is tedious to notate the detail manually. More importantly, it is not clear how challenging it is for a human student to consciously follow such information that changes about 10 times a second, and even higher for faster renditions. The only example of an existing system which needs such detailed information for manual use is the Gamaka-box [Vin12]. However, the extent of its adoption by Carnatic musicians and teachers is unclear.

Another musicological aspect deals with the types of *gamaka*s. The 15 types of *gamaka*s that are recognised today can be rendered by musicians, even on demand. The reverse operation, i.e., identifying *gamaka*-types from pitch curves is harder. It has not been established whether two expert musicians would agree on the segmentation of a melody (or its pitch curve) into constituent *gamaka*s. A hint in volume IV of [Sam58] suggests that the segmentation of CM pitch curves into *gamaka*-types is not unique: The author reduces all *gamaka*s to two basic types: *rava* (vibrato) and *jāru* (portamento). A *jāru* is a glide from one note-position to another. It is instructive that even the *kampita gamaka*, which is quoted most in CM lecture-demonstrations, is classified as the *jāru* type. Going a step further, pitch curves can be segmented uniquely into *jāru*s alone, if a stable note is treated as a *jāru* between two points at the same pitch. Even a vibrato can be explained as a series of concatenated *jāru*s. In this sense, [Ran+19c] segment pitch curves into *jāru*s, while [Vir+20c] segment pitch curves into constant pitch notes (i.e. stable segments) and transients that can be constructed with *jāru*s[9].

3.2.3 Representations of *gamaka*s for computational techniques

In spite of *gamaka*s and pitch-curves taking precedence over *svara*s of fixed pitch, note positions continue to be relevant to describe CM. For example, [Dik04] describes a *gamaka* as:

> In the position of dhaivata with a single pluck of the string, pulling it swiftly so as to sound niśāda, and then returning to dhaivata and then execute the pluck for the next svara.

This description is typical of how musicians (especially instrumentalists) discuss *gamaka*s. The example in the left of Figure 3.2 matches the above description with 'dhaivata' replaced by D_2 and 'niśāda' replaced by N_2. The examples in Figure 3.8, with 'dhaivata' replaced by R_2 and 'niśāda' replaced by G_2, are more telling. They bring together the terms discussed hitherto.

- A singer may enunciate the pitch curves with the *svara*-name 'Ri' (last row of Table 3.4). The specific variant of Ri is R_2 in the third-last row of the same table (*rāga pūrṇacandrikā*).
- The singer may equally well enunciate the same pitch curves as 'Ga', signifying the variant G_2 (*rāga tōḍī*) or the variant G_1 (*rāga varāḷī*).
- The pitch curve may be described in terms of segments using note-names (third row of Table 3.4). In this example, it consists of a *jāru* from R_2 to R_2 (i.e. stable pitch at R_2), followed by a *jāru* from R_2 to G_2, then a *jāru* from G_2 back to R_2, followed again by a *jāru* from R_2 to R_2 (i.e. stable pitch at R_2).

The last point above is an unambiguous description of the pitch curve, and is independent of the *rāga*. Just as the 12-note alphabet avoids the complexity of the many-to-many mapping in a 16-note alphabet, representations using uniquely segment-able *gamaka*s avoid the complexity of the many-to-many mapping between pitch-curves and *svara*-names. Using note-positions instead of semitones or cents will be more accessible to Carnatic practitioners. However, care must be taken to clarify that *svara*-names are not note positions. As mentioned earlier, when using *svara*-names such as 'Ri' or 'Ga', the context of the *rāga*, the choice among possible melodic atoms, and timing information are also needed to describe the pitch curve fully. These representations also provide timing information at a resolution of 0.01 seconds or finer, which enables synthesis of pitch curves.

3.2.4 Summary

*Rāga*s consist of phrases, which consist of *svara*s, which in turn can be rendered as a combination of stable notes and *gamaka*s. Figure 3.5 shows a landscape which captures the relations between these terms.

1. Note-positions in an octave have a many-to-many mapping to *svara*-names.

[9] The need for CPNs in the representation is motivated by the observed difference in precision between CPNs and stationary points [Vir+20b]

2. Pitch curves in CM have a many-to-many mapping to the same *svara*-names.

3. CM *svara*-segmentation is an open problem.

4. For *gamaka*s characterised by continuous pitch variation, musicological descriptions suggest that a representation with more detail than *svara*s is possible.

5. The second row of Table 3.3 provides a sufficient alphabet to describe pitch curves, including stable notes, in terms of a series of *jārus*[10]. If timing information is added, CM pitch curves can be synthesised.

Our notation in Table 3.3, the alternative names, and the the names of the seven notes alone, are all used in the literature on the computational aspects of CM. In fact, the choice of representation considerably alters the problem definition itself. Recent computational methods in CM melody analyse CM in terms of a descriptive transcription such as in the last two rows. This approach treats the *svara*-segmentation problem as a separate step and avoids the ambiguity of the many-to-many mapping. However, *svara*-segmentation remains an important, open problem: Section 3.5 explains that it needs to be solved to understand the interplay between melody and rhythm in CM.

3.3 Rhythm in Carnatic music

3.4 Lyrics in Carnatic music

Historically, lyrics and poetry preceded music and music may have started as an aid to memorising lyrics (Refer Murty Sir's article). However, given that music is now an evolved, sophisticated art by itself, we proceed in the reverse order in this section: from the mandatory role of syllables serving to carry music, then to lyrical constructs that are tightly coupled to music, i.e. their definition depends on the music, and conclude with examples where lyrics are emphasised more than the music.

3.4.1 Lyrics as vehicles for music

Syllables carry vocal CM. The examples of pitch curves shown in Section 6.3 need sustained rendition, which is possible in vocal music through vowels, half-vowels and nasals. These syllables have become standardised by practice. The most common syllables used in an *ālāpana* are ta, da, ri, na. Other syllables used are la and ra. Some syllables may be used more often in one octave than another: e.g. 'i' is used more often in the higher octave Cite Aarathi's work.. In *tānam*, the commonly used syllables are nam and tam. Thus, the most common vowels used for extension are *a,ā* and *i,ī*, and the nasals. We might say that this is the fundamental, culture-independent role of lyrics in music[11].

Indeed, there is an interesting experiment in the Tamil film, *tamizhpaḍam*, which is a spoof of Tamil movies in general. In this film, one of the songs (4 minutes long) has a sophisticated melody, but the lyrics starting with 'O Maha Zeeya' were deliberately written as a meaningless string of syllables ("gibberish" according to https://en.wikipedia.org/wiki/Thamizh_Padam). The song highlighted that the lyrics of modern-day Tamil songs have diverged considerably from those based on 'classical' Tamil in movie-songs a few decades ago. This spoof suggests that most humans expect meaningful lyrics in music, although any extensible set of syllables suffices for melody. Yet, around two centuries ago, CM's most revered composer, Saint Thyagaraja[12], chose to emphasise melody and rhythm over lyrics.

Thyagaraja's compositions are known for vowel extensions. For example, it is possible that the 11-syllable line, *cakkarnirāja mārgamuyunḍaga* is sung for 4 to 5 minutes. In this *kriti*, the distribution of the syllables

[10] This does not preclude the use of non-integer semitones. Appendix 3.6 considers the topic of '22 *śruti*s'.
[11] Discussed in 2017 in a private conversation with Carnatic vocalist, Vignesh Ishwar.
[12] Phonetic: *Tyāgarāja*. A crater on the planet Mercury is named after this composer; see https://en.wikipedia.org/wiki/List_of_craters_on_Mercury.

is given in Table 3.5. If the syllables yuṇḍa are made to match rā, the rhythmic symmetry in the two halves of 32 units each, demarcated by the Subtotals, is evident. The syllable rā lasts up to 18 units and the syllable yuṇ lasts up to 10 units. The composer uses these extensible syllables to introduce nearly 20 variations in the melody alone. Such variations are called *saṅgati*s in CM. Thyagaraja was easily the most accomplished, if not a pioneer, in composing elaborate *saṅgati*s. Typically, the variations proceed from the simplest (slow and spanning 7 semitones in this case) to the most sophisticated (fast and spanning 15 semitones). Further variations have been introduced over the last two centuries, as shown by the differences between current CM schools. But, it was Thyagaraja who enabled such variations by reserving many units for one syllable.

Table 3.5: Commonly sung syllable-arrangement in a *kriti* composed by Thyagaraja.

Syllable	**Vowel**	**No. of units**
ca(k)	a	4
ka	a	2
ni	i	4
rā	ā	12 or 16
ja	a	10 or 6
Subtotal		32
mā(r)	ā	4
ga	a	2
mu	u	4
yuṇ	u or ṇ	10
ḍa	a	2
ga	a	10
Subtotal		32
Total		64

The above discussion applies to compositions that serve as the definitive references for the *rāga*s they are composed in. These compositions are examples where the lyrics play a secondary role. Yet, even in these compositions, the rules of poetry were not compromised. In addition, Thyagaraja himself has composed many songs with around 10 stanzas of lyrics, all set to the same tune. Dikshitar has set Sanskrit lyrics to Western tunes. The lyrics mostly draw upon theological and mythological concepts such as the glory of Gods and Goddesses, philosophical concepts such as rebirth and the ephemeral nature of life, and patriotism during the Indian independence movement.

3.4.2 Lyrical constructs tied to music

As mentioned in Section 6.3, the same *svara*-name can be pronounced for several, different pitch curves. This practice has effectively fixed a set of syllables for carrying music in the form of *svara*-passages. Singing *svara*-passages is often compared to sol-fa singing, but *svara*-passages are far more sophisticated than sol-fa singing might suggest. They are used in both composed and extempore music. In fact, we are not aware of any system of music other than Carnatic and Hindustani, where sol-fa singing is included in a professional concert platform.

The many-to-many mapping of *svara*-names to pitch curves may pose challenges to transcription, but it lends itself to beautiful melodic-lyrical structures such as *svara-sāhitya*, which was popularised by the revered composer Shyama Shastri, who was contemporaneous with Thyagaraja. In this exercise, every *svara* corresponds to exactly one syllable. An example is given in Table 3.6 (Emanine, Mukhari, Adi, Subbaraya Shastri). This feature of CM is possible without compromising *rāga*-characteristics only because the *svara*-names can be mapped to pitch curves. For example, the bold entries in Table 3.6 correspond to the *svara*-names Ga being pronounced for a pitch curve that spans 5 semitones! Another concept is *svarākṣara*, where the consonant in a meaningful lyric matches the *svara*'s consonant. The first three syllables/*svara*s in Table 3.6 are examples.

Table 3.6: Lyrics and corresponding *svara*-names

Lyrics (syllables)	sa	ri	ga	dai	va	mu	gā	na	...	**nā**	tal	li
Svara-names	Sa	Ri	Ga	Sa	Ni	Dha	Pa	Ma	...	**Ga**	Ri	Sa

Mutthuswami Dikshitar, another revered composer contemporaneous with Thyagaraja, chose a direct mapping of syllable lengths to rhythm[13]. In Dikshitar's compositions, the duration of long syllables in the lyrics is twice that of short syllables. We refer to this as the '2:1 rule'. At least in the simpler *tāḷa*s, such as *ādi* and *rūpaka*, this rule is followed invariably. In some songs, exceptions are observed[14]. In other *tāḷa*s, the rule is adapted to suit its structure. For example, for the $3+4$ pattern in the *miśracāpu tāḷa*, some long syllables are 1.5 times as long as short syllables. Irrespective of the *tāḷa*, this 2:1 rule is followed even more rigorously in the fast segments that are twice as fast as the other segments of the song (*madhyamakāla sāhitya*). Dikshitar always included such fast segments in his compositions. An example is in the song *Vātāpigaṇapatim* (Hamsadhwani, Adi), which is shown in Table 3.7. It can be observed that the length of all long syllables is 2 units, and of all short syllables is 1 unit. Interestingly, this rule was followed even by Thyagaraja when he included fast segments in his compositions; an example is given later in Table 3.8. The strict mapping of syllable lengths to rhythm affords little room for syllable extensions. We hypothesise that it is for this reason that Dikshitar's compositions are generally sung at a slower tempo than Thyagaraja's compositions. That is, for the same *tāḷa*, the absolute value of the beat time is larger compared to the beat-time in a Thyagaraja composition. This would allow the longest syllable length to increase.

Table 3.7: Commonly sung syllable-arrangement in the fast segment of a *kriti* composed by Mutthuswami Dikshitar

Syllable	**Vowel**	**No. of units**
vī	ī	2
ta	a	1
rā	ā	2
gi	i	1
ṇam	m	2
vi	i	1
na	a	1
ta	a	1
yō	ō	2
gi	i	1
nam	m	2
Subtotal		16
viś	i	2
va	a	1
kā	ā	2
ra	a	1
ṇam	m	2
vigh	i	2
na	a	1
vā	ā	2
ra	a	1
ṇam	m	2
Subtotal		16
Total		32

[13] Refer Murthy Sir's article explains how Dikshitar's compositions are influenced by the techniques of playing the *vīṇā*.

[14] We estimate the exceptions occur in $< 10\%$ of the cases, e.g., 5 out of 50 in *Vātāpigaṇapatim*, 9 out of 120 in *Rāmacandram Bhāvayāmi*.

The preceding examples suggest that melody and rhythm were often prioritised over lyrics. Yet, the rhyme rules were followed rigorously by the three main composers. The rules are quite elaborate and occupy ten pages in [Sam58] (Book III). There are several examples of rhyme, the most important one being the second-syllable rhyme. The other two are the first-consonant rhyme (called *yati*) and last-syllable rhyme (*antyākṣaraprāsam*). The feature that distinguishes these rules from poetry is that the rules are tied to the *tāḷa*-structure more than lines of poetry. Even if a word breaks across *tāḷa*-cycles, the rhyme rules operate on the syllables starting or ending at *tāḷa*-cycles rather than on the word. These rules are summarised in Appendix 3.7.

The second-syllable rhyme is called *dvitīyākṣaraprāsam* (Sanskrit) or *yadugai* (Tamil). Of the many rhyme rules, this rule seems to be unique to Tamil (c.f. private conversation with a famous Tamil poet and polyglot, Madankarki). This rhyme heavily influenced CM compositions. An example is given in Table 3.8. Each row in Table 3.8 is sung for exactly 8 units. The first syllable has a long vowel (lasting exactly 2 units) and the second syllable is always 'ga'. This example is an extreme case because even one-fourth of a *tāḷa*-cycle shows the second-syllable rhyme, while the rule mandates the second-syllable rhyme only for the first and fifth rows (see Appendix 3.7).

Table 3.8: An example of repeated second-syllable rhyme

nā	ga	rājahṛd
sā	ga	rābja bhava-
sā	ga	rāntaka su-
rā	ga	hara kana-
kā	ga	dhīra sura
nā	ga	gamana śara
ṇā	ga	tābja śrī-
tyā	ga	rājanuta

3.4.3 Emphasis on lyrics

Among the Carnatic trinity of composers, Dikshitar emphasised lyrical beauties. About half of Dikshitar's compositions contain the *rāga*-name. There are several methods that the composer has used to insert the *rāga*-name. In the song mentioned in Table 3.7, the words ***hamsadhvani*** *bhūṣita hērambam* contain the *rāga*-name (*Hamsadhvani*). The meaning of this line is simply, 'adorned by the *rāga* Hamsadhvani' [Rao06a]. This is an example of a straightforward, but contrived, insertion of the *rāga*-name. About half songs that have *rāga*-names in them employ this method. Another method of including the *rāga*-name exploits the meaning of the name. For example, in the phrase ***bhairavī*** *prasaṅgam*, the bold word (*bhairavī*) is the name of the *rāga* and of a deity. Another example is the phrase *raṅganāyakam bhāvayē śrī raṅga**nāyakī**samētam*. The bold word is the name of the *rāga* (*Nāyaki*), but this time it is *part* of the name of the deity. In a complementary example, the phrase *cidbim**bau lī**lāvigrahau*, the bold syllables form the *rāga* name (*Bauli*), but they come from different words in the lyrics. Carrying this further, in the example *bim**bē gaḷa**jitaśaṅkhē*, the *rāga* name (*Bēgaḍa*) is straddled between two words of the lyric, provided the phone ḷ is taken as the acoustically close phone ḍ. Another example of the same type occurs in *samī**pa ṛju**mārgadarśitam*, where the *rāga* name (*Paraju*) is suggested by the phonetically similar *paṛju*.

In the remaining compositions of Dikshitar, the *rāga* name does not appear or is indicated partially. For example, the phrase ***rudra****kōpajāta*, the first two syllables of the *rāga* name (*Rudrapriyā*) appear. Similarly, the phrase *kamala**jānanda**bōdhasukhī*, the first word is part of the *rāga* name (Ānanda*bhairavī*). Just as Dikshitar included the *rāga*'s name in several of his compositions, he also included the references to the temples' locations and the specific stories associated with those temples. Several modern-day composers incorporate *rāga* names in the lyrics.

Dikshitar also employed other lyrical beauties. There is a decreasing pattern (*gōpuccha yati*) of the lyrics in Table 3.9. Although the music follows the rule of *rāga*, *tāḷa*, and rhyme, the emphasis is on the lyrical beauty.

In this example, apart from the pattern itself, the meaning of the first three occurrences of the word *padē* ('one whose feet are ...') is different from that in the last two (together, they mean 'In every step [in a walk]'). The convention in using the same word multiple times is that its meaning in each occurrence should be different. For example, [Sam58] explains that each of the four occurrences of *kamala* in one of Dikshitar's compositions (*Kamalāmbām bhajare*) has a different connotation.

Table 3.9: An example of a decreasing pattern in lyrics

śrīsārasapadē
rasapadē
sapadē
padē
padē

In Section 3.1, it was explained that in some *kriti*s, the same tune is repeated with only lyrics being different. For these cases, the tune is simply referred to as 'Same as anupallavi' in notation books such as [Rao06a; Rao06b]. Ironically, while Dikshitar emphasised lyrics most, he was the first composer who departed from the tradition of having multiple *caraṇam*s with the same tune[15]. This practice has come to stay, and unless a vocalist chooses to emphasise the lyrics and their meaning, multiple *caraṇam*s in major songs are rarely sung today. Dikshitar also deviated from repeating the tune of the *anupallavi* in the second half of the *caraṇam*. However, most other composers including modern day composers, do choose to repeat the tune of the *anupallavi*.

3.5 Conclusion

An introduction to Carnatic music was followed by detailed descriptions of the salient features of CM with respect to the three areas of melody, rhythm and lyrics. Since these three areas are abstract-able and independent of the type of composition or mode of rendition. current computational studies typically consider each area individually Cite Ranjani's article. In studies of melody, low-level descriptions that do not depend on *svara*-names circumvents the problem of many-to-many mapping of pitch curves to *svara*-names. Thus, analysis and synthesis of melody in CM seems a far more tractable problem today than it did 30 years ago. A low-level notation for percussion would help similarly. The study of melody and rhythm are relatively more advanced than the study of lyrics in music, but speech-related techniques may be adaptable to the study of lyrics.

Of course, CM has evolved as an art and will continue to evolve. While modern-day compositions follow the styles prescribed by the Trinity of Carnatic composers, new *rāga*s have been invented recently and compositions have been tuned in them. An example is the *rāga Lavangi*, which departs from the norm that a raga should have at least five *svara*s in an octave. However, as mentioned repeatedly in this article, *gamaka*s are an integral part of CM. It is not clear whether there is a limit to the number of *rāga*s that have information-bearing *gamaka*s, and if so, whether this limit has been reached. Other factors too are shaping CM. The influence of Hindustani music is felt in the choice of *rāga*s not only for composition, but also for extempore treatment. Similarly, fast *tāns* are borrowed from Hindustani music in Carnatic concerts. Very recently, other forms and languages are being used in Carnatic music, but often this 'innovation' is restricted to lyrics from different cultural contexts. There are very few recent examples, if any, of major compositions that emphasise the music as described in Section 3.4.1. In this light, the descriptions of melody, rhythm and lyrics in this article are expected to hold even as CM evolves.

It now remains to bring the three aspects of melody, rhythm and lyrics back together. The syllables of the lyrics in CM need to occur at specific points in the *tāḷa*. These points are fixed at the time of composition. In a *niraval*[16], the syllables of the lyrics must occur at the same points in *tāḷa* as in the original composition, no

[15] When his first composition was criticised for not having multiple *caraṇam*s, Dikshitar 'encoded' this conscious choice into the line *matvātvadēkacaraṇam* of his second composition. In some of the minor *rāga*s, Dikshitar's *kriti*s have only a *pallavi* and and *anupallavi*, and eschews the *caraṇam* altogether. Sometimes, this *anupallavi* is called the *samaṣṭi caraṇam*.

[16] Extempore elaboration of lyrics set to *rāga* and *tāḷa*. See Figure 3.9 for an example.

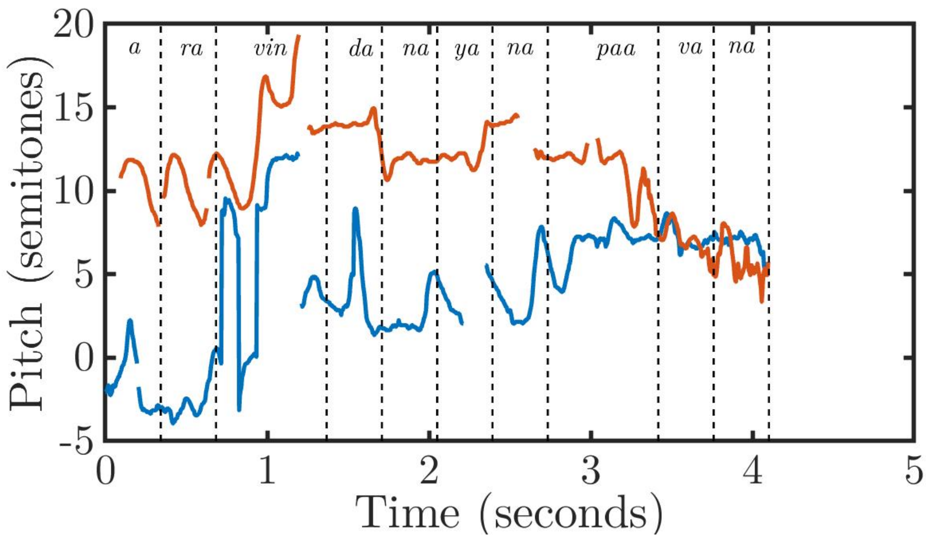


Figure 3.9: Syllables (top) demarcated by vertical dashed lines in two renditions (brown and blue curves) of the same lyrics in a *niraval*. This example is from a professional Carnatic concert.

matter how the melody is varied between them. Thus, although *svara*s constitute only a sparse transcription, they do specify where the syllables should occur in time. This is why, for completeness, the problem of *svara*-segmentation/identification must be attempted. Further, [Vir+19] found that the scaling of pitch curves across tempi is not uniform. Specifically, silence-segments and the stable-pitch segments of pitch curves scale far more than the transient segments. This non-uniform scaling can be addressed partly in a rule-based/statistical manner [Vir+17; Vir+20c]. However, a musicological approach suggests that the locations of syllables of *svara*s or lyrics should scale uniformly across tempi, while the pitch curve between these points scales non-uniformly. Thus, while computational techniques in the areas of melody, rhythm, and lyrics will continue to improve, a paradigm change in terms of the joint study of syllables, rhythm and melody is needed to better understand the sophistication and nuances of CM.

References

[Com12] CompMusic. *CompMusic Project*. http://compmusic.upf.edu/. Accessed: 2017-04-16. 2012.

[Dik04] Dikshita, S. *Sangita Sampradaya Pradarshini*. 2008 web version, 1904.

[DP12] Dhandapani, M. and Pattamal, S. D. *RAGA PRAVAGAM*. Kavitha Publications, 2012.

[GS19a] Garani, S. S. and Seshadri, H. "An algorithmic approach to South Indian classical music". In: *Journal of Mathematics and Music* vol. 13, no. 2 (2019), pp. 107–134.

[Gul+14a] Gulati, S. et al. "Automatic tonic identification in Indian art music: approaches and evaluation". In: *Journal of New Music Research* vol. 43, no. 1 (2014), pp. 53–71.

[Gul+16d] Gulati, S. et al. "Time-Delayed Melody Surfaces for Rāga Recognition." In: *ISMIR*. 2016, pp. 751–757.

[Kan07] Kanno, M. "Prescriptive notation: Limits and challenges". In: *Contemporary Music Review* vol. 26, no. 2 (2007), pp. 231–254. eprint: https://doi.org/10.1080/07494460701250890.

[KD07] Klapuri, A. and Davy, M. *Signal processing methods for music transcription*. Springer Science & Business Media, 2007.

[Kri04b] Krishnaswamy, A. "Melodic Atoms for Transcribing Carnatic Music". In: *International Symposium on Music Information Retrieval*. 228. ISMIR. 2004.

[KSS12b] Koduri, G. K., Serrà Julià, J., and Serra, X. "Characterization of intonation in carnatic music by parametrizing pitch histograms". In: *Proceedings of the 13th International Society for Music Information Retrieval Conference; 2012 Oct 8-12;* 2012.

[PSS13b] Porter, A., Sordo, M., and Serra, X. "Dunya: A system for browsing audio music collections exploiting cultural context". In: *Britto A, Gouyon F, Dixon S. 14th International Society for Music Information Retrieval Conference (ISMIR); 2013 Nov 4-8; Curitiba, Brazil.[place unknown]: ISMIR; 2013. p. 101-6.* International Society for Music Information Retrieval (ISMIR). 2013.

[Raj+17a] Raj, P. V. K. et al. "An Approach to Transcription of Varnams in Carnatic Music Using Hidden Markov Models". In: *Proc. of the National Conference on Communication (NCC)*. 2017.

[Ram04] Ramanathan, H. *Ragalakshanasangraha (Collection of Raga Descriptions)*. 1st. N. Ramanathan, R.A.Puram, Chennai, 2004.

[Ran+19c] Ranjani, H. et al. "A compact pitch and time representation for melodic contours in Indian art music". In: *The Journal of the Acoustical Society of America* vol. 145, no. 1 (2019), pp. 597–603.

[Rao06a] Rao, T. K. G. *Compositions of Mudduswami Dīkshitar*. Ganamandir Publications, 2006.

[Rao06b] Rao, T. K. G. *Compositions of Śyāmā Śāstrī*. Ganamandir Publications, 2006.

[Rat15] Ratanjanakar, S. N. *Chaturdandi Prakāshikā (Hindi Translation)*. Sanskar Prakashan, 2015.

[RPS17b] Ranjani, H., Paramashivan, D., and Sreenivas, T. V. "Quantized Melodic Contours in Indian Art Music Perception: Application to Transcription." In: *ISMIR*. 2017, pp. 174–180.

[Sam58] Sambamoorthy, P. *South Indian Music, (2019 reprint)*. Vol. I to VI. Indian Music Pub. House, 1958.

[Sat01] Sathyanarayana, R. *Chaturdandi Prakasika of Venkatatmakhin (Vols I and II)*. Kalamulasastra Series Vol. 24 and 25, 2001.

[See58] Seeger, C. "Prescriptive and Descriptive Music-Writing". In: *The Musical Quarterly* vol. 44, no. 2 (1958), pp. 184–195.

[Sri+14c] Srinivasamurthy, A. et al. "Corpora for music information research in indian art music". In: *Georgaki A, Kouroupetroglou G, eds. Proceedings of the 2014 International Computer Music Conference, ICMC/SMC; 2014 Sept 14-20; Athens, Greece.[Michigan]: Michigan Publishing; 2014.* Michigan Publishing. 2014.

[SWM11a] Subramanian, S., Wyse, L., and McGee, K. "Modeling speed doubling in carnatic music". In: *ICMC*. 2011.

[VAM17a] Viraraghavan, V. S., Aravind, R., and Murthy, H. "A Statistical Analysis of Gamakas in Carnatic Music". In: *Proceedings of the 18th International Society for Music Information Retrieval Conference*. 2017, pp. 243–249.

[VE75] Von Helmholtz, H. and Ellis, A. J. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. London: Longmans, Green and Company, 1875.

[Vij07b] Vijayakrishnan, K. *The grammar of Carnatic music*. Vol. 8. Walter de Gruyter, 2007.

[Vin12] Vinayakam, R. *Gamaka Box- the first ever Comprehensive Notation System for Indian Music*. http://www.rameshvinayakam.com/research. Accessed: 2016-11-13. 2012.

[Vir+17] Viraraghavan, V. S. et al. "Non-uniform time-scaling of Carnatic music transients". In: *arXiv preprint arXiv:1711.02318* (2017).

[Vir+19] Viraraghavan, V. S. et al. "Visualizing Carnatic music as projectile motion in a uniform gravitational field". In: *Proc. Workshop on Speech, Music and Mind 2019*. 2019, pp. 31–35.

[Vir+20a] Viraraghavan, V. S. et al. “State-Based Transcription of Components of Carnatic Music”. In: *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2020, pp. 811–815.

[Vir+20b] Viraraghavan, V. S. et al. “Data-driven measurement of precision of components of pitch curves in Carnatic music”. In: *The Journal of the Acoustical Society of America* vol. 147, no. 5 (2020), pp. 3657–3666.

[Vir+20c] Viraraghavan, V. S. et al. “State-based transcription of Carnatic music components”. In: *Int. Conference on Audio, Speech and Signal Processing 2020*. IEEE. 2020, accepted.

3.6 The 22 *śruti*s

Carnatic musicology specifies frequency ratios (with respect to the tonic) of 22 note-positions in an octave; these are famously known as ‘the 22 *śruti*s.’ However, in recent literature, [Kri04b] and [Vij07b] suggest that CM may be described fully using 12 notes in an octave, together with pitch curves. [RPS17b] do quantise critical points to the 22 *śruti*-frequencies. However, their formulation is not consistent with the theory of 22-*śruti*s[17]. Further, their results do not conclusively show that using 22 notes is advantageous: the perception ratings differ only marginally. Thus, if a transcription scheme proposes the use of more than 12 notes per octave, the necessity of the higher granularity must be proved first.

Another common, but unverified, argument is that the time-averaged pitch of *svara*s might match one of the values of the 22 *śruti*s. [Vir+19] show across 9 renditions in 6 *rāga*s that the stable notes scale nearly 3 times as much as transients across tempi. The anecdotal example for the *svara* R_2 in the *rāga* Kedaragaula is also telling. Clearly, the time-averaged pitch of this *svara* , and others, is tempo-dependent, but the theory of 22 *śruti*s has no provision for tempo dependency.

Next, some of the ratios given in the 22-*śruti* table of [Sam58] are high-integer fractions. Consider the fraction $\frac{729}{512}$ as a variant of M_2. For a tonic of 200 Hz (quite common for female musicians), this ratio implies that the first overlapping partial is the 729$^{\text{th}}$ harmonic at 145800 Hz, which is beyond the human auditory range. The second overlapping partial for the alternative ratio for M_2, $\frac{45}{32}$, is at 25600 Hz, again beyond the human auditory range (most adults over the age of 40 cannot hear the first partial itself). Clearly, using the cycle of fifths to generate these ratios has discarded the basic premise of low-integer ratios suggested by [VE75]. By contrast, the ratios given in [Raj+17a] have integers smaller than 20, and may be a better hypothesis to test. In fact, [Sam58], who first provided frequency values of the 22 *śruti*s, has acknowledged that the frequency values have not been proven. Yet, his statement below has largely been ignored:

> The frequencies given for the 22 srutis in Table V [of volume IV of the text book] are those which have been arrived at after mature deliberations in the conferences held during recent decades [presumably 1930s to 1950s]. When a suitable apparatus which will indicate the frequencies of notes sung or played is devised, we can experimentally prove the values of these srutis.

Thus, future research depending on experimentally unverified ratios is best avoided. Also, the ‘suitable apparatus’ to measure frequencies is now available. We have observed that the measured note values in histograms do differ by about 20 cents in a *rāga*’s renditions by different musicians. This variation is comparable to the differences between frequencies of *śruti*s for the same note. Thus, the restriction to 12 note-positions does not preclude the use of tuning systems with non-integer semitones. However, the experiments need to consider the precision results of [Vir+20b]. The standard deviation of errors in the pitch of long stable notes (i.e. CPNs with duration more than than 300 ms) is about 10 to 15 cents and the standard deviation of errors in the pitch of stationary points is of the order of 50 cents. In this light, the differences of the order of 20 cents between

[17] A note that quantises to D_2 among 12 notes in the octave may quantise to a variant of N_2 among the 22 notes, which is not consistent with the theory. To test the 22-*śruti* hypothesis, it is necessary to quantise notes other than Sa and Pa to one of the 10 notes first and then substitute each with only one of its two variants (and not any other among the 20 remaining positions).

Table 3.10: Kanjadalayatakshi, Kamalamanohari raga, Adi tala, M. Dikshitar. Pairs of rhyme-lines within a rhyme-segment are boxed together.

Section	**Rule**	**Syllables**				**Rule**
		1st	**2nd**	**Others**	**Last**	
Rhyme-segment 1 (*pallavi* and *anupallavi*; $m = n = 1$)						
Pallavi	3.7.1.1	*k*añ	**ja**	dalāyatākṣi kāmāk	ṣi	
	3.7.1.2:1	*k*a	ma	lāmanōhari tripurasunda	ri	–
Anupallavi (2nd speed)	3.7.1.1	kuñ	**ja**	ragamanē maṇimaṇḍi	ta	–
	3.7.1.2:3	maṅ	**ju**	ḷacaraṇē māmavashi	va	
	3.7.1.1	paṅ	**ja**	raśuki pankajamukhi gurugu	ha	
	3.7.1.2:3	raṅ	**ja**	ni duritabhaṅjani niraṅja	ni	–
Rhyme-segment 2 (*caraṇam*; $m = n = 1$)						
Caraṇam	3.7.1.1	*r*ā	**kā**	śaśivadanē suradha	**nē**	
	3.7.1.2:1	*r*ak	ṣi	tamadanē ratnasada	**nē**	3.7.1.3
	3.7.1.1	*ś*rī	**kaṅ**	cinivasanē surasa	**nē**	
	3.7.1.2:1	*ś*ṛṅ	gā	rāśritamandahasa	**nē**	3.7.1.3
2nd speed	3.7.1.1	*ē*	**kā**	nēkākṣari bhuvaneśva	**ri**	
	3.7.1.2:1	*ē*	kā	nandāmṛtajari bhāsva	**ri**	3.7.1.3
	3.7.1.1	*ē*	**kāg**	ramanōlayakari śrīka	**ri**	
	3.7.1.2:1	*ē*	kām	rēśvarigṛhēśvari śanka	**ri**	3.7.1.3

different tuning systems seem relevant only to the CPNs. Further, the role of 22-*śruti* ratios in distinguishing between *rāga*s is not clear; we believe this role is likely to be negligible.

We also mention that our observations are consistent with most parts of Volume IV of [Sam58]. The *gamaka*-examples in Table V of the same work are valid from a musicological point of view, and serve as a reference for melodic atoms. It is only the assignment of precise ratios to these melodic atoms that is questionable. Further, [Vij07b] argues that 23 (and not 22) melodic atoms are needed to describe *gamaka*s in CM, but this is yet to be proved.

3.7 Rhyme rules

The rhyme rules occupy ten pages in [Sam58]. The rules are best read along with examples. The first composition we consider is Kanjadalayatakshi in Kamalamanohari *rāga*, by Muttuswami Dikshitar and is presented in Table 3.10. It exemplifies nearly all the rules described later in Section 3.7.1. It is thus the typical example in the most commonly used *ādi tāḷa*.

Delving into the example, we first identify the rhyme-segments. The *pallavi* and *anupallavi* together constitute Rhyme-Segment 1. The *caraṇam* is Rhyme-segment 2. In both rhyme-segments, there is one *tāḷa*-cycle per rhyme-line ($m = n = 1$, in the notation of Section 3.7.1). For the sections in 2nd speed (*madhyamakāla* sections), each rhyme-line occupies one half of the *tāḷa*-cycle. The two rhyme-segments are arranged by pairs of rhyme-lines in Table 3.10. Note that only identifying the rhyme-lines depends on the speed of the sections. In this example, the first two rhyme-lines of Rhyme-segment 1 are sung at the original speed, and the last four rhyme-lines of the same rhyme-segment are sung at twice the speed. In terms of duration, the first two rhyme-lines occupy the same number of *tāḷa*-cycles (which is 2) as the last four rhyme-lines. After the rhyme-lines are identified, rhyme-rules apply to all the rhyme-lines irrespective of the speed.

Next, we consider the second-syllable-consonant rule. This is shown by the second and third columns of Table 3.10. The rule (described in Section 3.7.1.1) is referred as 3.7.1.1 in the second column for every alternate rhyme-line starting from the first. Each rhyme-segment has a single rhyme-consonant (*prāsākṣara*). In Rhyme-segment 1, the rhyme-consonant is ‘j’. In the Rhyme-segment 2, it is ‘k’. This is also indicated by the boldface letter of the rhyme-consonant in both segments. We observe that this consonant is identical in all alternate rows of each rhyme-segment. We also observe that the rule only specifies the consonant and the vowel attached to it is immaterial. Further, the first syllable of all alternate rhyme-lines in both rhyme-segments are all of the same duration: a long syllable in this case.

The first-syllable-phone rule is also observed to be followed. (The exact clause that applies is indicated in the second row of each pair of rhyme-lines, and the phones that match are shown in italics.) In the *caraṇam*, the first-syllable-phone rule is followed in every pair of rhyme-lines. In the first pair in Rhyme-segment 1, the first-syllable phones are identical (‘k’). In this regard, the last two pairs in the *caraṇam* also follow the second-syllable-consonant rule. However, this is not mandatory and we thus mark only the first-syllable-phone rule. The *pallavi* of Rhyme-segment 1 also adheres to this rule. The composer has the choice of skipping the first-syllable rule by falling back to the second-syllable rule. In the *anupallavi*, all pairs of rhyme-lines fall back to the second-syllable rule instead of the first-syllable rule. We reiterate that, in the case of falling back, the rhyme-consonant (‘j’) and the length of the first syllable must align to the rest of the rhyme-segment. Additional adherence to the second-consonant-rule (in the lyric *duritabhanjani* or in Table 3.8) is for alliterative effect that is enjoyed by listeners, but not mandatory.

Finally, we consider the last-syllable rhyme. It is desirable to have the same last syllable between rhyme-lines in a pair, if not across line-pairs. However, this rule has the least priority (hence the term ’should’ in the clause on Section 3.7.1.3). We observe that the last-syllable rhyme is followed in some, but not all, pairs of rhyme-lines.

3.7.1 Definitions

In this section, we state the rhyme rules in a rigorous manner. The definitions for the rhyme rules are given in Table 3.11. The three main rhyme rules are described next. They are easier to follow if the reader is familiar with the *kriti*s mentioned in the examples.

3.7.1.1 Second-syllable rhyme

This rule is called *dvitīyākṣaraprāsa* (lit. second-syllable rhyme in Sanskrit) or *yadugai* (Tamil). It is thought that this rule first appeared in Tamil verses and was borrowed in Telugu poetry, which was later used in Carnatic compositions. This theory is consistent with the rule’s absence in much earlier Sanskrit compositions such as the *śl̄oka*, and Jayadeva’s *aṣṭapadi* compositions. Starting from the first rhyme-line in a rhyme-segment, the second-syllable rhyme applies to every alternate rhyme-line in the same rhyme-segment. The clauses to be satisfied simultaneously are:

1. The length of the first syllable of each of the alternate rhyme-lines must be identical (either all short or all long).
2. The second-syllable consonant of each of the alternate rhyme-lines must belong to the same phone-class.

The second-syllable consonant in Rule 2 is called the rhyme-consonant (*prāsākśara*). Each rhyme-segment has one rhyme-consonant. The rhyme-consonants in different rhyme-segments are independent, i.e., they may be different or the same.

3.7.1.2 First-syllable rhyme

This rule is called *yati* (Sanskrit)[18] or *mōnai* (Tamil). The first-syllable rhyme applies to a pair of rhyme-lines in the rhyme-segment. The first rhyme-line in the pair is an alternate rhyme-line, which should also conform to the second-syllable rule. The next rhyme-line is the second rhyme-line in the pair. The rhyme-lines in the pair must satisfy one of the below clauses:

1. The first-syllable phone of each of the alternate rhyme-lines must belong to the same phone-class.

[18] To be distinguished from increasing and decreasing patterns, also called *yati*.

Table 3.11: Definitions of terms used in describing rhyme rules

Term	**Definition and Example(s)**
Rhyme-segment	The *pallavi* and *anupallavi* of a rendition (except *tillānā*) taken together, or Each *caraṇam* of the rendition[a]. Every rhyme-segment is independent of the others.
Rhyme-line	One line of lyrics aligned to the *tāḷa* as follows: In sections that are sung at the initial tempo of the composition, an integer number (say *n*) of cycles of the *tāḷa*. In sections that are sung at double the initial tempo (*madhyamakāla*), an integer number (say *m*) of *half*-cycles of the *tāḷa*. Notes: – There should be no gaps between consecutive rhyme-lines. – All rhyme-lines should together cover the entire rhyme-segment. – Rhyme-lines are aligned to the start of the rhyme-segment or the *sama* of the *tāḷa*, whichever is later. – Often, $m = n$, but this is not mandatory. – Double-*kaḷai ādi tāḷa* compositions need to be expanded into single-*kaḷai* before applying rhyme-rules.
Phone-class	A class of phones with similar sound-production techniques. The following phone-classes are recognised. Subset of vowels: *[a, ā], [i, ī], [u, ū], [e, ē], ai, [o, ō], au* Guttural consonants: *ka, kha, ga, gha* (*ka-varga*) Palatal consonants: *ca, cha, ja, jha* (*ca-varga*) Cerebral consonants: *ṭa, ṭha, ḍa, ḍha* (*ṭa-varga)*) Dental consonants: *ta, tha, da, dha* (*ta-varga)* Labial consonants: *pa, pha, ba, bha* (*pa-varga*) Half-vowels and fricatives: *ya, ra, la, va, śa, ṣa, sa, ha* Subset of nasal sounds: *na, ṇa* Notes: – Sub-classes of vowels are grouped in square brackets. – The diphthong 'ai' belongs to two sub-classes [a,ā] and [i,ī] – The diphthong 'au' belongs to two sub-classes [a,ā] and [u,ū] – The consonant 's' also belongs to the palatal consonants.
First-syllable phone[b]	The first consonant or vowel in the first syllable of a rhyme-line.
Second-syllable consonant[b]	The first consonant in the second syllable of a rhyme-line.

[a] A *samaśticaraṇam* should be treated as the *anupallavi* as is done by [Dik04].

[b] These definitions apply also to syllables consisting of 'half'-consonants. E.g., if the two-syllable word is *vaktra*, the first-syllable phone is 'v' and the second-syllable consonant is 'r'. The consonant 'k' between them does not have a bearing on rhyme-rules.

2. When the first-syllable phone is a vowel and Rule 1 is not satisfied, the other rhyme-line in the pair can start with a consonant, provided the vowel associated with that consonant and the first-syllable phone are within the same sub-class. Sub-classes are indicated by square brackets in the subset of vowels in Table 3.11. In addition, the diphthong ai belongs to both the sub-classes [a, ā] and [i, i=], and the diphthong au belongs to both the sub-classes [a, ā] and [u, u=].

3. If none of the above rules is satisfied, the second-syllable rhyme (including length of the first syllable) must be followed between the same pair of lines. In this case, the rhyme-consonant must match that of the rhyme-segment.

It is due to the fall-back option in Rule 3 that the second-syllable rhyme has to be defined ahead of the first-syllable rhyme. Another convention is that rhyme-lines in a pair cannot straddle different sections of the song (i.e. one of a pair cannot be from the *pallavi*, while the other is from the *anupallavi*. Therefore, the application of this rule depends on the composition structure. The two most common structures of section-lengths (counted in number of *tāḷa*-cycles) are:

- *Pallavi*: 1 cycle, *Anupallavi*: 2 cycles, Each *caraṇam*: 4, 6, or 8 cycles
- *Pallavi*: 2 cycles, *Anupallavi*: 2 cycles, Each *caraṇam*: 4, 6, or 8 cycles

For the first example above (1,2,4), an empty second rhyme-line needs to be introduced in the first rhyme-segment.

3.7.1.3 Last-syllable rhyme

The last-syllable rhyme is different from the common last-vowel/nasal rhyme in English poetry. For example, the opening lines of 'The Inchcape Rock' by Robert Southey:

> No stir in the air, no stir in the sea
> The ship was as still as she could be

are considered as rhyming because 'sea' rhymes with 'be'. However, note that only the last vowel (ī) rhymes, while the consonants (s and b) differ. In CM, and especially in Dikshitar's compositions, the last-syllable rhyme is exactly as the name suggests:

1. The last syllables in entirety, i.e. consonant and vowel/nasal, in a pair of rhyme-lines should belong to the same group.

3.7.2 Examples

In this section we give further examples of compositions that follow the rhyme rules.

3.7.2.1 Srimahaganapatiravatumam

A more involved example is presented in Table 3.12. The *kriti* is Srimahaganapatiravatumam by M. Dikshitar in the *rāga* Gaula. The *tāḷa* is Misra Chapu, which is essentially counting 3.5 beats or 14 units at a time (Section 3.3). The adherence to the rules are shown in the respective columns and using boldface and italicised as in Table 3.10.

The *pallavi*, *anupallavi*, and the 2nd speed section of the *anupallavi* constitute Rhyme-segment 1. The *caraṇam* and its 2nd-speed section constitute Rhyme-segment 2. In both rhyme-segments, two *tāḷa*-cycles form a rhyme-line in original speed and two half-cycles form a rhyme-line in the 2nd speed ($m = n = 2$). From Table 3.12, it is clear that all alternate rhyme-lines in both rhyme-segments follow the second-syllable consonant rule (3.7.1.1) without exception. Similarly, all pairs of rhyme-lines follow the first-syllable phone rule without exception. As in the first example, the last-syllable rhyme, which is optional, is adhered to in some cases.

A few peculiarities in this composition stand out. First, it is rare for compositions in *tāḷa*s like Misra Chapu or Khanda Chapu to have only one cycle per rhyme-line. Therefore, the identification of rhyme-lines is easier if such *tāḷa*s are reckoned as longer equivalent cycles: Tishra triputa or Khanda Ekam respectively in these cases[19]. Next, the second line of the *anupallavi* presents clarity on the strict definition of the second-syllable-consonant rule. The first consonant of the second syllable matches the rhyme-consonant of the rhyme-segment, but this is not adhering to the rule because the first syllable of this rhyme-line (ka) is short, while the first syllables of the other alternate rhyme-lines are long (śrī, kā, kō).

This *kriti* also exemplifies the use of phone-classes in the rhyme-rules. In the last pair of Rhyme-segment 1, the first-syllable rule is followed. The first consonant of the first syllable of the two rhyme-lines are 'k' and 'g',

[19] In this method of reckoning, $m = n = 1$.

Table 3.12: Srimahaganapatiravatumam, Gaula raga, Misrachapu tala, M. Dikshitar

Section	Rule	Syllables 1st	2nd	Others	Last	Rule
Rhyme-segment 1 (*pallavi* and *anupallavi*; $m = n = 2$)						
Pallavi	3.7.1.1	*śrī*	**m**a	hāgaṇapatiravatu	mām	–
	3.7.1.2:1	*sid*	dhi	vināyakō mātangamu	kha	–
Anupallavi	3.7.1.1	*kā*	**m**a	janakavidhīndra sannu	ta	–
	3.7.1.2:1	*ka*	ma	lālaya taṭanivā	sō	
2nd speed	3.7.1.1	*kō*	**m**a	latara pallavapadayu	ta	
	3.7.1.2:1	*gu*	ru	guhāgraja śivātma	ja	–
Rhyme-segment 2 (*caraṇam*; $m = n = 2$)						
Caraṇam	3.7.1.1	*su*	**v**ar	ṇākarṣaṇa vighna	rā–	
	3.7.1.2:1	*jō*	pā	dāmbujō gaura	var–	–
	3.7.1.1	*ṇa*	**v**a	sanadharō phālacand	**rō**	
	3.7.1.2:1	*na*	rā	divinuta lambōda	**rō**	3.7.1.3
	3.7.1.1	*ku*	**v**a	layasvaviṣāṇa pā	śān–	
	3.7.1.2:1	*ku*	śa	mōdaka prakāśaka	rō	–
	3.7.1.1	*bha*	**v**a	jaladhi nāvō mū	la–	
	3.7.1.2:1	*pra*	kṛ	tisvabhāva sukhata	rō	–
2nd Speed	3.7.1.1	ra	**v**i	sahasrasannibhadē	**hō**	
	3.7.1.2:3	ka	**v**i	jananuta mūšikavā	**hō**	3.7.1.3
	3.7.1.1	*a*	**v**a	nata dēvatāsamū	**hō**	
	3.7.1.2:1	*a*	vi	nāśakaivalyagē	**hō**	3.7.1.3

which are both guttural consonants. Thus, Clause 3.6:2 is satisfied. Similarly, the first pair of rhyme-lines in Rhyme-segment 2 use different consonants ('s' and 'j'), but being from the same class (palatal; see note in Table 3.11), the first-syllable-phone rule is adhered to. The second pair of rhyme-lines in Rhyme-segment 2 also employ different first-syllable-phones, but they are from the same phone-class, and similarly for the fourth pair.

Finally, there are four words that are broken across rhyme-lines. These words are marked by a trailing dash in the column 'Last'. It can be observed that the words are broken across rhyme-lines in such way so as to adhere to the rhyme rules. The word *varṇa* is of special interest. It breaks across the second and third rhyme-lines in Rhyme-segment 2 as *var–ṇa*. With this split, the third rhyme-line begins with ṇa. Yet, a popular Sanskrit dictionary [**apta**] does not list even one word that starts with this letter. Thus, the lengths to which the composer has gone to adhere to the rhyme rules are evident. Another related point is that when words are split across rhyme-lines, the last-syllable-rhyme rule necessarily suffers[20].

3.7.3 Further Analysis (Optional Reading)

3.7.3.1 Banturiti

The previous two examples are Sanskrit *kriti*s composed by Muttuswami Dikshitar. Table 3.13 has an example of a Telugu *kriti* composed by Thyagaraja. The *kriti*'s name is Banturiti and is in the *rāga hamsanādam*. It is set to *ādi tāḷa*. In this *kriti*, the lyric starts 6 *mātra*s after the start of the *sama*[21]. The rhyme-lines begin at this point. Since there is no 2nd-speed segment, the value of *n* does not apply. Further, the *kriti* follows the 1 cycle/2 cycles/4 cycles pattern respectively for the *pallavi*, *anupallavi*, and *caraṇam*. Since the *pallavi* has only one rhyme-line, an empty rhyme-line marked by asterisks (*) has been introduced in the *pallavi*. We observe that the mandatory rhyme rules are adhered to, while the optional last-syllable rhyme is absent. Further, the

[20]A notable exception is the composition Balagopala (Bhairavi, Adi) by M. Dikshitar. In the 2nd-speed sections of both *anupallavi* and *caraṇam*, words are split across rhyme-lines even while adhering to the last-syllable-rhyme rule.

[21]There are 4 *mātra*s/beat and 8 beats per *ādi tāḷa* cycle. Section 3.3 defines the terms *mātra*s and *sama*.

second rhyme-line of Rhyme-segment 2 adheres to the second-syllable-consonant rule. Yet, this is not a case of falling back on the second-syllable-consonant rule because the first-syllable-phone rule is also adhered to, which takes priority for the second rhyme-line in each pair.

Table 3.13: Banturiti, Hamsanadam raga, Adi tala, Thyagaraja

Section	Rule	Syllables				Rule
		1st	2nd	Others	Last	
Rhyme-segment 1 (*pallavi* and *anupallavi*; $m = 1$)						
Pallavi	3.7.1.1	*b*aṇ	**ṭ**u	rītikoluvīyyavayyarā	ma	–
		*	*	*	*	
Anupallavi	3.7.1.1	*t*uṇ	**ṭ**a	viṇṭivāni modalaina ma	dā	–
	3.7.1.2:1	*d*u	la	kōṭinēla gūlajēyu ni	ja	
Rhyme-segment 2 (*caraṇam*; $m = 1$)						
	3.7.1.1	*r*ō	**m**āṅ	camane ghanakaṅcuka	mu	–
	3.7.1.2:1	*r*ā	ma	bhaktuḍane mudrapiḷḷa	yu	
Caraṇam	3.7.1.1	*r*ā	**m**a	nāmamane varakhaḍgamulu	vi–	–
	3.7.1.2:1	*r*ā	jil	lunayya tyāgarājuni	ke	

3.7.3.2 Nambikettavar evarayya

We conclude with an example of a Tamil *kriti* by the 20th century composer, Papanasam Sivan. Table 3.14. This *kriti* also shows that the values of *m* can differ from one rhyme-segment to another. In the *pallavi* and *anupallavi*, each rhyme-line lasts for one *tāḷa*-cycle. In the *caraṇam*, each rhyme-line lasts two *tāḷa*-cycle. We again observe that all the mandatory rhyme rules are followed, while the optional last-syllable rhyme is absent[22].

Table 3.14: Nambikkettavar evaraaya, Hindolam raga, Adi tala, Papanasam Sivan

Section	Rule	Syllables				Rule
		1st	2nd	Others	Last	
Rhyme-segment 1 (*pallavi* and *anupallavi*; $m = 1$)						
Pallavi	3.7.1.1	*n*am	**b**ik	keṭṭavar evarayyayā u	mai	–
	3.7.1.2:1	*n*ā	ya	ganai tirumayilaiyin iṛaiva	nai	–
Anupallavi	3.7.1.1	*a*m	**b**u	ligangaiaṇindajaṭḏa	ran	–
	3.7.1.2:1	*a*n	bar	manam vaḷar shambukapāli	yai	–
Rhyme-segment 2 (*caraṇam*; $m = 2$)						
	3.7.1.1	*o*n	**ṙ**u	mē payanillai enṛuṇarnda pinbu uṇḍen	bār	–
	3.7.1.2:1	*o*v	vo	ru manidanum orunaḷinnilai ēyduvuduṙudi adai maṙan	dār	–
Caraṇam	3.7.1.1	*a*n	**ṙ**u	ceyalizandu amarum pozudu civan peyar nāvil vārā	dē	–
	3.7.1.2:1	*ā*	da	lināl manamē inṙē civanāmam colli pazagu anbu	ḍan	–

3.7.3.3 Other examples

In the *kriti vidulakumrōkkeda*, the word *Mārkaṇḍēya* would have been ideal to conveying the meaning easily: Markandeya is the name of a famous sage. Instead, the composer chose the equivalent *Sumṛkaṇḍuja* (lit. the son of *Mṛkaṇḍu*), which is understood by very few listeners. With this word, the rhyme-consonant of the *caraṇam* (‘m’) finds its place without violating the second-syllable-consonant rule. This suggests that, Thyagaraja, who prioritised music over lyrics, strove to diligently adhere to the rhyme rules.

[22]Here again, we revisit the comparison with English. The question is whether *vār* rhymes with *dār* in the first pair of rhyme-lines of the *caraṇam*. According to our definition of the last-syllable rhyme, it does not. If the rule is changed so that last-vowel/nasal rhyme suffices, many more *kriti*s will be found to follow this rule too. M. Dikshitar has followed the strict version of the rule.

In the *kriti Angārakamāśryāmyaham* (Surati ragam, Rupaka talam, Muttuswami Dikshitar), the second rhyme-line starts with *tāśritajana...*. This is an example where the first phone of the first syllable of the first rhyme-line is a vowel ('a'). The first syllable of the second rhyme-line is a consonant ('t'), but the vowel following it is 'ā', which is in the same sub-class as 'a' (Table 3.11 defines these sub-classes). A similar example is found in the *caraṇam* of the *kriti Śrīgaṇapatini* (Saurashtram ragam, Adi talam, Thyagaraja). The two rhyme-lines in the relevant pair begin with *anayamu* and *yāmbuja* respectively.

We now consider two examples where the lyric starts before the start of the *tāḷa*. In the *kriti* Bhaja Re Bhaja Manasa (Kannda *rāga*, Misra Chapu *tāḷa*, Thyagaraja), the *caraṇam* starts with *bhūsama śāntam*. The syllable *bhū* occurs before the start and is therefore not counted. The rhyme-line starts as *sama śāntam*, and the rhyme-consonant is thus 'm'. The corresponding rhyme-line is *vāramakhilada*. Again, the syllable *vā* occurs before the start of the *tāḷa*, and is not counted. The rhyme-line starts as *ramakhilada*, which follows the second-syllable-consonant rule. Two more examples are found in the *caraṇam*s of the *kriti* Maye (Tarangini *rāga*, Adi *tāḷa*, Muttuswami Dikshitar). They respectively start with *upāyē* and *samudāyē*, and the syllables *u* and *samu* are not counted. The rhyme-lines therefore start as *pāyē* and *dāyē*, and the adherence to the second-syllable-consonant rule is evident[23].

3.7.4 Closing Remark

The description of rhyme rules in CM compositions were presented in a formal manner, but there is one open issue. The definition of the rhyme-lines is not specified fully. That is, there are choices to be made. Indeed, in Muttuswami Dikshitar's first *kriti Śrīnāthādiguruguhō* (Mayamalavagaula *rāga*, *ādi tāḷa*), the rhyme rules are adhered to in the *caraṇam* only if $m = n = 2$. That is, in the original speed, a rhyme-line consists of two *tāḷa*-cycles, and in the 2nd-speed segments, a rhyme-line consists of 2 half-cycles. Defining rhyme-lines with $m = n = 1$ would suggest that the rhyme rules are not adhered to. A rigorous definition may therefore need to specify that m and n must be chosen so that the composition's overall-adherence to the rules is maximised.

Measuring the adherence to the rhyme rules could help in deciding the provenance of *kriti*s. In CM circles, there is a notion that only the *kriti*s in [Dik04] are authentic and were composed by Muttuswami Dikshitar, while the same composer's *kriti*s outside this book are mis-attributions. A quick check with the rhyme rules suggests this view may be true. For the listings in the first 24 Melakarta *rāga*s and their derivatives, fewer than 10% of non-adherence to the rhyme rules are observed. In fact, only one *kriti* (Ramachandrena, Manji *rāga*) violates the rule about the length of the first syllable in the second-syllable-consonant rule. For *kriti*s in the same *rāga*s and their derivatives, *kriti*s outside this book (such as in [Rao06a]), the non-adherence is about 50%. Of course, a deeper study is needed for conclusive assertions. In such a study, the optimisation of m, n mentioned in the previous paragraph will play an important role. Irrespective of the outcome of this (possible) study, some potentially mis-attributed *kriti*s are also definitive references for their respective *rāga*s (e.g., Hariharaputram, Vasanta). In this sense, provenance is mainly of academic interest.

3.8 Glossary

The definitions of other CM-specific terms are given in Table 3.15, but the main body of the text uses alternative English descriptions. The terms defined in Table 3.2 are used throughout the article. For completeness, these terms are defined by reference at the bottom of Table 3.15.

[23] This *kriti* is full of alliteration, but presents its challenges. Unlike all other Dikshitar-*kriti*s, it has multiple *caraṇam*s and $m = 2$ would leave each *caraṇam* with a single rhyme-line. The *kriti* may thus not strictly conform to the rhyme-rules described here.

Table 3.15: Definitions of CM-specific terms

Ālāpana	Extempore exploration in a *rāga*; it is not bound by rhythm.
Anupallavi	Second segment of a composition
Caraṇam	Third segment of a composition, often twice the size of the *anupallavi*. The fourth and later segments, if present, are also called *caraṇam*s. Usually, all *caraṇam*s have the same melody and *tāḷa*, but different *sāhitya*.
Collu	Vocal syllable used to mark units of a *tāḷa*
Gamaka	Continuous pitch variation. More generally, ornamentation.
Kuccēri	Music Concert, often CM concert
Kriti	Most important Carnatic composition form, whose melody and *sāhitya*, and bound by a *tāḷa*. A *kriti* has a *pallavi*, *anupallavi*, and at least one *caraṇam*.
Konnakōl	A sequence of *collu*s conforming to a rhythmic pattern
Mṛdangam	TBD
Pakka vādyam	Accompanying instrument
Pallavi	First segment of a composition
Piḍi	Characteristic phrase of a *rāga*; often, but not always, unique to the *rāga*
Rāga	Melodic modes. See Section 3.2.1.
Sāhityam	Lyrics of a composition
Svara	The name of a musical note or of a pitch curve
Tamburā	Drone instrument
Tāḷa	Rhythmic cycles. See Section 3.3.
Tillānā	A CM composition with rhythmic patterns. It uses percussion-syllables in the *pallavi* and *anupallavi*, and *sāhitya* in the *caraṇam*.
Vādyam	Instrument
Karuvi	
Varṇam	The most advanced practice exercise, with *pallavi*, *anupallavi*, and *caraṇam*. The *anupallavi* and *caraṇam* have long *svara*-passages. *Rāga* and *tāḷa* are emphasised in a *varṇam*.

Chapter 4

Of Acoustics and Aesthetics: Analysing Vocal Timbre in the Hindustani Khayal

Srijan Deshpande

Manipal Center for Humanities, Manipal Academy of Higher Education, Manipal, India
srijand@gmail.com, srijan.stck.me

Abstract

This paper will attempt to show how contemporary research in the disciplines of vocology and psychoacoustics, together with recently developed tools for acoustic analysis can be put to musicological use in order to develop a nuanced account of the resonance strategies and timbral aesthetics of master vocalists. This work focuses on the Hindustani Khayal genre of music and uses computational representation in order to empirically examine recorded audio to gain musicological insights of an aesthetic/stylistic nature. It is thus a departure in some ways from recent and ongoing work in this domain. After a brief introduction to recent path-breaking research in psychoacoustics and vocology, this paper will, as a case study, analyse short audio excerpts from recorded performances of two of the most important names of the twentieth century Khayal tradition: Ustad Abdul Karim Khan and Ustad Faiyaz Khan. By examining the computed acoustic spectra of these samples on the basis of the principles propounded by the aforementioned disciplines, we will attempt to examine and explicate anecdotal descriptions as well as scholarly assessments of these masters' voices. While there are many different acoustic signal analyses possible that capture perceptually relevant aspects of the music, this work is restricted to the spectral analysis of held vowels, since it is in held vowels that the timbral aesthetic of a vocalist is arguably most apparent. We will also see how such analyses can shed light on possible aesthetic, stylistic and even physiological reasons for their particular acoustic and performative choices. The paper will conclude with a brief account of the numerous research possibilities this kind of timbral analysis throws up and the importance of such research to the musicology of traditional Indian music.

4.1 Introduction: Empirical Musicology and Acoustic Analysis

The use of technology as an aid to musicological analysis, though not a new activity, has become an exceedingly important part of music studies today. This is largely due to the sophisticated technology available to us now as well as to the insights into the production and perception of the singing voice that science has given us. Within the domain of Indian classical music however, research that uses computational approaches tends to depend heavily upon the codified grammar of this music, in order to either generate music computationally[1] or to examine performance recordings to retrieve structural information about the music from them[2]. This retrieved information tends largely to deal with intonation (shrutīs) and melodic structure (raga phrase-structure) and is often implicitly evaluated against grammatical rules found in traditional pedagogical texts in order to develop accounts of the degree to which the theory and practice of music are compatible in contemporary performance.

[1] Some examples include Das et al (2005 [DC05]) and Srinivasamurthy et al (2012 [SC12c])

[2] See Rao et al. (2014 [RR14]) for an excellent overview of recent research of this kind, particularly within the area of music information retrieval (MIR).

It appears from the thrust of such work that the grammar of Hindustani music, as canonised by the likes of early twentieth century musicologists V.N. Bhatkhande, V.D. Paluskar and their successors exerts an overbearing influence upon the direction this research takes. This is unsurprising given that conformance to a standardized grammar has come to be one of the signifiers of the 'classicism' of this music[3]. While this research certainly has opened up exciting new areas such as music information retrieval (MIR) and continues to enable useful applications such as recommendation engines and pedagogical tools, it must be said that there exists scope for computational research in Hindustani music that deals more directly with performance beyond its compliance with canonised textual grammar. This paper deals with one such research area: that of vocal timbre.

In recent years, the areas of psychoacoustics, signal analysis and vocal pedagogy have all made significant advances, enabling us to fully understand and engage with the phenomenon of vocal timbre with a great deal of nuance. Important works that straddle these areas in an interdisciplinary fashion include Howell (2016 [How16]), Bozeman (2013) [Boz13] and Titze (2015 [TA15]). Although these authors, especially Howell and Bozeman, are voice pedagogues and their work is oriented towards the training and habilitation of the singing voice, their work opens up the possibility of using spectrography and timbral analysis to develop our understanding of how master vocalists use vocal timbre. Such timbral analysis can help us better understand the aesthetic and stylistic goals of these master musicians and describe them in a quantifiable way, while also allowing us to develop a more concrete understanding of anecdotal information about voices and voice use that is available in the tradition[4]. This study brings timbral analysis to the musicology of the Khayal genre of Hindustani music by applying it to the vocal timbres of two of the most important vocalists of the tradition: Ustad Abdul Kareem Khan and Ustad Faiyaz Khan. These musicians have been chosen because their vocal timbres, as much as their overall musical styles, are considered to be two polar extremes within the genre[5]. It is hoped that a comparative analysis of their vocal timbres will serve both to impart clarity to the discussion as well as to highlight the diversity the Khayal genre incorporates.

4.2 The Spectral Envelope and Singing Voice Perception

The following is a brief introduction to basic concepts and terminology in vocal acoustics that are relevant to this paper. The works cited above and within this section can be referred to for comprehensive discussions of the same.

A human voice singing a particular pitch produces a musical tone – a sound that has a fundamental frequency, sometimes described as a lowest common denominator frequency[6]. It simultaneously possesses a number of other frequencies, each of which are multiples of the fundamental frequency and follow the harmonic series. These are known as harmonics and are numbered H1, H2, H3 and so on, where H1 refers to the fundamental frequency. These sets of harmonics are produced at the vocal folds (known as the voice source) and their relative intensities are acted upon by the resonator of the human vocal instrument – the vocal tract, which functions as a filter, since it selectively strengthens or weakens (in intensity) the harmonics in the sound signal fed to it. Depending upon its size and shape at the moment of phonation, the vocal tract has a number of resonant frequencies known as formants, and numbered F1, F2 and so on. The harmonics produced at the voice source that are closer to these vocal tract resonances or formants are boosted in intensity, while those that are further away from them are significantly weakened. The resultant set of harmonics comes to be perceived as a sound that has a composite timbre, the 'colour' or quality of which depends upon the relative intensities of its component harmonics. Because the vocal tract can change its size and shape, the human voice is capable

[3] See Powers (1980 [Pow80]), Schofield (2010 [Sch10]) and Manuel (2015 [Man15]) for more on 'classical' as a social as much as an aesthetic or stylistic construct.

[4] The Hindustani Khayal traditional has remained substantially oral, not just in terms of performance pedagogy but also in terms of its supporting theoretical discourse. The oral discursive tradition thus contains large amounts of valuable anecdotal information, analysing and problematising which is an important function of the musicology of this music.

[5] See Deshpande (1987 [Des87], 41–45) for a discussion of how their styles are in stark contrast to one another.

[6] Bozeman (2013)[Boz13], p. 3

of producing a large variety of harmonic-formant interactions and, thereby, a variety of timbral colours. This understanding of the voice is known as the non-linear source-filter model[7] and forms the basis of this paper.

Recent advances in psychoacoustics have given us more insight into the relationship between our perception of the resultant composite sound described above and the relative intensities of its component harmonics caused by specific harmonic-formant interactions. Three concepts are particularly relevant to this paper: Absolute Spectral Tone Colour (ASTC), Resonance Areas and ‘Roughness’. Each of these will be introduced briefly below[8].

4.3 Absolute Spectral Tone Colour (ASTC)

ASTC is a theory proposed by Ian Howell (2016 [How16]), the basic premise of which is the idea that in our perception of sound, specific frequencies or frequency ranges are associated with specific timbral colours. Just like we see certain frequencies of light as certain colours, we perceive certain frequencies of sound as specific timbral qualities. These qualities or colours are very similar to vowel sounds. Herein lies the crux of the concept of ASTC – that human beings perceive specific frequencies as specific vowel-like tone colours, and that these tone colours are the same irrespective of the source of the sound (hence the term ‘absolute’)[9]. Additionally, these vowel-qualities (and therefore the frequencies associated with them) have perceptual connotations of brightness or darkness. When applied to the harmonic spectrum generated by the singing voice, ASTC gives us a means of associating certain harmonics within the spectrum with specific vowels. Spectral analysers (such as VoceVista Video, the tool used for the present analysis) allow us to visualise the harmonic content of a sound signal and, importantly, to easily notch-filter each harmonic and listen to it in isolation in order to clearly appreciate its ASTC vowel-quality[10]. Importantly, this also allows us to understand vowels as spectral arrangements created by particular harmonic-formant interactions such that the harmonics that lie within the frequency range associated with a particular vowel are boosted in intensity while singing that vowel. Or as Howell puts it, ‘A vowel is not a single color. A vowel is more like a painting made up of many individual colors’ (Howell 2017 [How17], 5). The following graphic that depicts Howell’s ideas provides a good summary of these frequency-vowel associations:

Frequency Range (Hz)	upto 450	450-750	750-1000	1000-1300	1300-1500	1500-1900	1900-2300	2300-3500	3500 Onwards
Perceived Vowel Quality (IPA)	u	o	ɔ	ɑ	a	ɛ	e	i	Bright i
Perceived Brightness	Dark								Bright

Figure 4.1: ASTC Frequency-vowel associations. Based on observations in Howell (2017 [How17], 5)

[7] ’Non-linear’ because it also takes into account the fact that acoustic pressure is not only radiated outwards but is also reflected back to the voice source, influencing its vibration. See Bozeman (ibid) for more.

[8] To be sure, there are other aspects of timbre that are certainly relevant to this paper, namely the temporal aspects of the composite sound, such as attack, decay and release. However, the singing voice typically exhibits long stationary sections and these are what are primarily referred to when describing vocal timbre. Thus attack, decay etc are not as relevant to the singing voice as they are to instrumental timbre (e.g. plucked strings). Additionally, this paper orients itself around the analysis of singular moments within the spectro-temporal flux, as Howell terms it; a kind of analysis made possible by the visual representation of the harmonic spectrum that spectral analysers such as VoceVista give us. Howell provides the following justification for this approach: ‘Two sine tones of the same frequency are identically bright. This is independent of other elements of timbre: attack, decay, release, or general spectro-temporal flux.’ (Howell 2017 [How17], 5). The term ‘bright’ here refers to a psychoacoustical concept established by Plomp, which Howell extends to say that two such sine tones also have an identical tone-colour or vowel-like quality – the crux of ASTC, explained in the following section. This fact makes it feasible for us to avoid engaging with the temporal aspects of the signal.

[9] This fact is especially relevant since timbre is conventionally described as the quality of sound that differentiates two sources from each other, their pitch, loudness and duration being identical. ASTC defines timbre in terms of component tonal colours that are identical between even vastly different sound sources. See Howell (2016 [How16], 20–26) for a detailed discussion of these ideas.

[10] A short but comprehensive and revealing demonstration of this by Howell can be found at https://www.youtube.com/watch?v=TUKYSmF7d10

For the purposes of the present discussion, this opens up the possibility of breaking down a singer's vocal timbre into its component groups of harmonics to see which ones tend to receive energy boosts due to the configuration of the singer's vocal apparatus at the moment of phonation; and to thereby arrive at a nuanced understanding of the singer's vocal timbre.

4.4 Resonance Areas

As Ken Bozeman (2017 [Boz17], 179–80) observes, ASTC allows us to divide the typical range of vocal tract resonances into three broad regions or 'resonance areas'. The first resonance of the vocal tract (the first formant or F1) defines the first of these regions and typically boosts harmonics that lie below or around 700Hz[11]. This region is associated with the vowels /u/ and /o/[12] and contributes 'depth, warmth, and roundness to the tonal construct' (Bozeman 2017 [Boz17], 179). It bears mentioning here that it is the converse of this that is in fact the case: ASTC tells us that 'the vowels /u/ and /o/ are actually defined by the tone colour contribution of F1' (ibid).

The second resonance of the vocal tract (the second formant or F2) defines the second of these regions and typically boosts harmonics that lie between approximately 800Hz and 2100Hz. This region is associated with the vowels /ɔ ɑ a æ ɛ e i/ and is considered critical in giving these vowels their particular identity. While this area covers all the vowels except /u/ and /o/, individual frequencies within this band, when heard in isolation, are perceived to have the quality of individual vowels[13]. This area 'contributes some degree of clarity, projection or brightness' to the tonal construct (ibid, 180).

The third, fourth and fifth resonances of the vocal tract (F3, F4 and F5) together make up the third resonance area. These formants boost harmonics that lie between approximately 2100Hz to 4000Hz and are associated with the /i/ vowel. The frequencies on the higher end of this range take on a very bright /i/ quality and in general, this area contributes the qualities of 'ring, projection, forwardness/brightness' to the overall tonal construct (ibid).

4.5 Roughness

Howell describes auditory roughness as 'a buzzing, sometimes pulsing or beating quality introduced by the inner ear because the cochlea is unable to differentiate simple tones that are very close in frequency' (2017 [How17], 4). Howell goes on to formulate a general rule that says that 'any two simple tones a minor third or closer will give rise to such roughness; the closer, the rougher' (ibid). This interval, called the 'critical band' becomes important because, as Howell demonstrates, all harmonics of any singing voice above H5 'fall within the... critical band of a neighbour'(ibid)[14]. It is important to note here, that unlike ASTC which predicts the vowel-like timbral quality of an overtone based on its frequency, the perceived 'roughness' of an overtone depends upon its relative position in the harmonic series and not on its absolute frequency. Thus, harmonics H1-H5 in the singing voice are perceived as 'pure' sounds as compared to H6 and above which are perceived as 'rough' irrespective of the pitch being sung and the resultant frequencies of those harmonics. The implication of this is that if the harmonic-formant interactions in a singer's voice are such that harmonics above H5 are boosted, the singer's timbre will be perceived as increasingly rough; and vice versa. This general rule will prove important to the discussion that is to follow.

[11] Howell / Bozeman define these areas in terms of their proximity to the treble clef. I have translated them to frequency ranges to avoid bringing staff notation into a discussion on Hindustani music. See Bozeman (2017 [Boz17], 179–80) for formant positions in reference to the treble clef.

[12] This is experimentally verifiable: if harmonics in this frequency range are listened to in isolation, they are perceived as having a distinct /u/ or /o/ vowel-like quality. A sample of a singing voice can be notch-filtered in a spectrum analyser to isolate this area, as demonstrated in the video link mentioned in footnote 10 above. It bears repeating here that this applies to any sound source, not just the singing voice – this is the premise of absolute spectral tone colour.

[13] Thus a tone of 1000Hz will appear to have an /a/-like quality while a tone closer to 2000Hz will have an /e/-like quality and so on.

[14] Also see Sundberg (1994 [Sun94], 115) and Lichte (1941 [Lic41]) for prior work on auditory roughness

Harmonic Number	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	Hn
Pitch Resolvability	Resolved								Unresolved			
Roughness	Pure					Rough, progressively rougher						
Summary	Pure & Resolved					Rough & Resolved			Rough & Unresolved			

Table 4.1: Roughness and Resolvability. Based on observations in Howell (2017 [How17], 5)

Another ancillary and important concept is that of the 'resolvability' of the overtones in the singing voice. The concept of the missing fundamental is firmly established in acoustics and states, in essence, that the pitch of a sound remains perceivable even if the fundamental frequency (H1) is removed. Howell shows, however, that this phenomenon has its limits: 'Generally, the lowest eight harmonics resolve neatly into the pitch, while the ninth harmonic and higher (called unresolved harmonics) contribute progressively less to the perception of the pitch. 'This means that the higher a harmonic is in the series than H9, the more it exists as part of a separate percept, hanging above, rather than coloring the pitch' (ibid, 5). The concepts of roughness and resolvability can thus be put together to give us three perceptual divisions of the voice, as depicted by the figure above: Pure and Resolved; Rough and Resolved; and Rough and Unresolved.

Howell also notes, importantly, that 'Neither roughness nor resolvability exists in the sound wave; the ear creates them' (ibid). As the remainder of this paper will show, the concepts of ASTC, Resonance Areas and Roughness/Resolvability can be put to very fruitful use to describe the timbral-aesthetic choices that musicians make as well as the way their voices are perceived and described by listeners. The next section will apply these ideas to samples taken from the voices of Ustad Faiyaz Khan and Ustad Abdul Kareem Khan after first conducting a brief survey of descriptions of their voices as found in the relevant literature.

4.6 Case Study: Comparing the Vocal Timbres of Faiyaz Khan and Abdul Karim Khan

4.6.1 Descriptions in Scholarly Literature

4.6.1.1 Ustad Faiyaz Khan

The following extended excerpt from Bonnie Wade takes into account the various ways in which the timbre of Ustad Faiyaz Khan's voice has been described. Wade draws upon a large number of commentators to construct her account, and I have retained her citations to make clear her sources for each of the descriptions:

> 'Faiyaz Khan had an unusually low pitched voice[15] which contributed to the effect of his singing; and he had the vocal power that was cultivated by his gharana predecessors (Danielou 1951:57). Descriptions of his vocal quality range from 'clear and sweet' (Vilayat Hussain Khan 1959:115) to 'broad, full husky' (V.H. Deshpande 1973:13). Vasant Thakur dubbed it 'broad, masculine (1966:31) and Dipali Nag reiterated 'masculine'. . . It is possible that the desire to emulate Faiyaz's low voice produced the 'gruff', 'grating', 'husky' vocal quality noted by V.H. Deshpande (1973:16) and Amar Nath (Interview:1978) as characteristic of Agra gharana vocal production.' (Wade 2016 [Wad16], 107–8)

Wade goes on to list a number of descriptors that have been used for the voices of Faiyaz Khan's disciples and successors, including 'beautiful. . . like a lion'; '. . . deep, rich. . . '; 'full-throated'; 'broad, rugged' and

[15] Faiyaz Khan's available recordings have him singing with his tonic note at (approximately) C#3 or C3, and often even as low as B2, very unusual for vocalists of his time. He was also not known to have a very large pitch-range. Anecdotal evidence suggests it stretched to a maximum of 2½ octaves.

'powerful'. (ibid) Ashok Ranade describes Faiyaz Khan's 'unabashed projection of a bass, booming voice' and makes the important observation that 'among his contemporaries, employment of a high fundamental pitch by vocalists was in itself regarded a virtue and a custom religiously followed' and that Faiyaz Khan's choice of a lower fundamental pitch was 'courageous' and set him apart. (Ranade 2011 [Ran11], 122–23). Additionally, Ranade makes the important observation that Faiyaz Khan's vocalisation had a 'marked superimposition of the vowel sound approximate to 'a' (as in 'man')[16] instead of the usually advocated 'a' (as in father)' (ibid). This is a particularly important observation and possible reasons for this vowel choice will be proposed in section 4.9 below.

4.6.1.2 Ustad Abdul Karim Khan

To quote Wade again:

> Two statements have been universally used to describe Abdul Karim's voice: 'sweet' and 'pliant'. To these can be added 'high', for his natural pitch lay in a range that Western practise would call tenor[17]. He produced an effect that was the opposite of, for instance, the powerful, forceful Agra gharana voice. . . The overall effect is of quietness and peace. . . (Wade 2016 [Wad16], 197–98)

V.H. Deshpande characterizes Abdul Karim Khan's voice as 'pointed' and 'sweet' and the resultant music sung using it as 'swara-oriented', to the extent that 'swara-intonation was cultivated [by him and his followers] as a prized possession' so that they made 'purity of swara. . . the very essence of their musical art'. Deshpande also makes it a point to contrast this kind of voice with that cultivated by Faiyaz Khan's followers and says that the latter ended up 'ignor[ing] values connected with swara – its sweetness, smoothness and delicate artistry of tonal nuances'. (Deshpande 1987 [Des87], 41–42)

Ashok Ranade's comments are noteworthy again:

> [Abdul Karim Khan's] vocal tone could be described as thin, pointed, resonant in segments, slightly nasal and occasionally a little hollow. . . The vocal tone so uniformly applied by him was further characterised by superposition of the compound vowel-sound 'o' or a variety of it[18] on the entire vocalising process. . . [this] vowel sound proved a very reliable aid in maintaining 'aas'[19] in his renderings. . . [this] reliance on a compound, resonant vowel sound. . . was a deviation from the normal practice of Hindustani practitioners of vocal music who insist on the use of the vowel sound 'AA'. (Ranade 2011 [Ran11], 60–62)

The remainder of this section will attempt to complicate and make nuanced these analyses of these two voices, while also attempting to explicate the particular timbral choices these two vocalists made, based on the theories of vocal acoustics outlined above.

4.7 Spectral Analysis

The following analysis restricts itself to two audio samples[20]. Both are contiguous segments; the first is 5.5sec in duration and is taken from a recording of Ustad Abdul Kareem Khan (AKK), while the second is 5.8sec

[16] Represented by the IPA symbol [æ], although in the sample analyzed below, he uses its neighbour, [ɛ].

[17] Abdul Kareem Khan's available recordings have him singing with his tonic note at (approximately) F3.

[18] Represented by IPA symbols for rounded vowels such as [o] and [ɔ]

[19] 'Aas' is a term frequently encountered in Hindi/Marathi literature on music and can be roughly translated as 'lingering resonance'. To quote Ranade again, '. . . to have 'aas' in music means that instead of a sudden and certain secession, music fades off without our becoming aware of its stopping or ending'. (2011 [Ran11], 62)

[20] These segments have been chosen as representative of these vocalists' timbre based on the current author's experience with curating and studying their music. The current study is meant to be an introductory, demonstrative study rather than a comprehensive or definitive one, and thus does not take a larger sample set into account. Its findings should thus be considered as a starting point for further research.

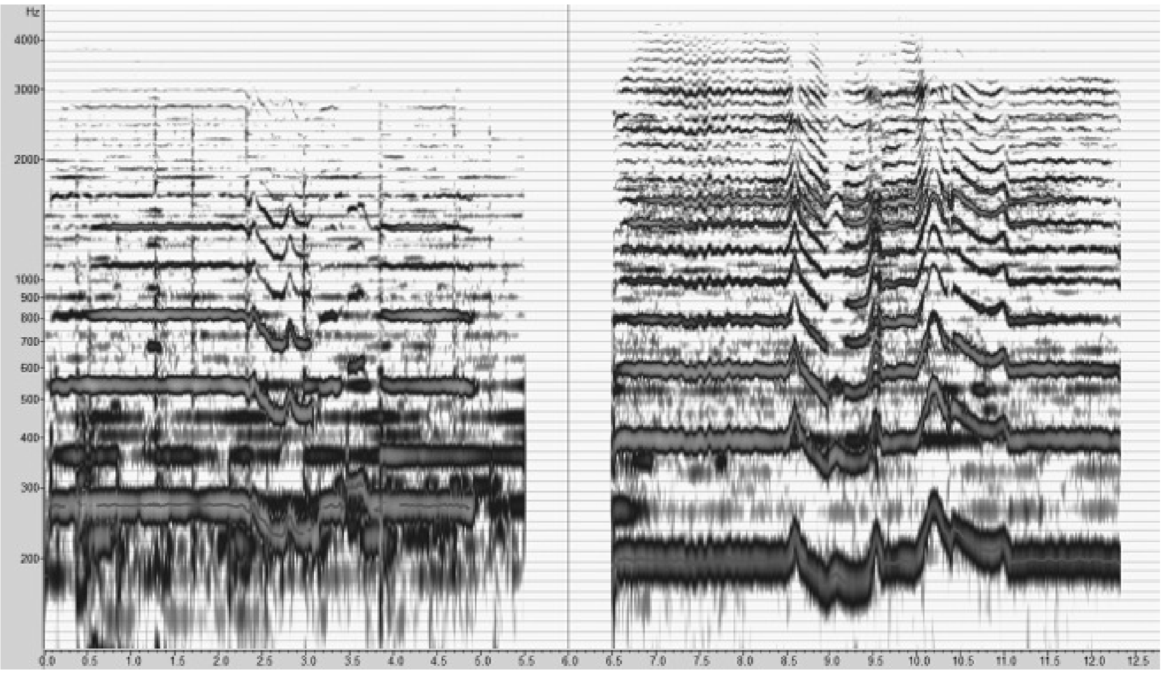


Figure 4.2: Spectra taken from clips of Abdul Kareem Khan (Left) and Faiyaz Khan (right). Source: Author's analysis.

in duration and is taken from a recording of Ustad Faiyaz Khan (FK)[21]. Both recordings were made during the same time period, by the same company, and given the infancy of recording technology at the time, it is reasonable to assume that the equipment used as well as the recording conditions for both artists were similar enough to justify a comparative analysis of this kind.

In figure 4.2, the section on the left is the spectrogram of the sample from the AKK recording while that on the right is from the FK recording. Both excerpts have been chosen to depict the singers singing around the pancham or fifth of their respective keys, and can be said to represent the middle of their tessitura.

After observing both audio clips and taking readings at various points in them, a single moment was identified in each audio clip as being representative of the timbre of each vocalist. Both clips have the artists articulating an /a/ vowel, although their pronunciation of the /a/ is coloured by their individual vowel preference as will be demonstrated below. Both artists sing the pancham (fifth) of their respective tonics in these clips. For AKK, this translates to an F4, while for FK, this is a G3 note. The spectra in figure 4.3 show the intensities of the various component harmonics of the sound produced by each vocalist (AKK on the left and FK on the right) at this selected representative moment:

The table 4.2 contains readings taken from the above spectra:

Note:

- The symbols in the ASTC column represent the vowel qualities ASTC theory ascribes to the frequencies of those harmonics (see Figure 1 above)
- In the Roughness column, Pur = Pure; Rou = Rough; Res = Resolved; Unr = Unresolved. (See Figure 2 above for reference)

[21] See discography section below for details of source recordings. The FK sample considered here starts at 08 seconds into this recording, while the AKK sample considered here starts at 17 seconds into the recording. Both recordings are also available on YouTube, at https://www.youtube.com/watch?v=Wdjlf_mzOiw and https://www.youtube.com/watch?v=UEE4ALUdFAk respectively, as of March 2021.

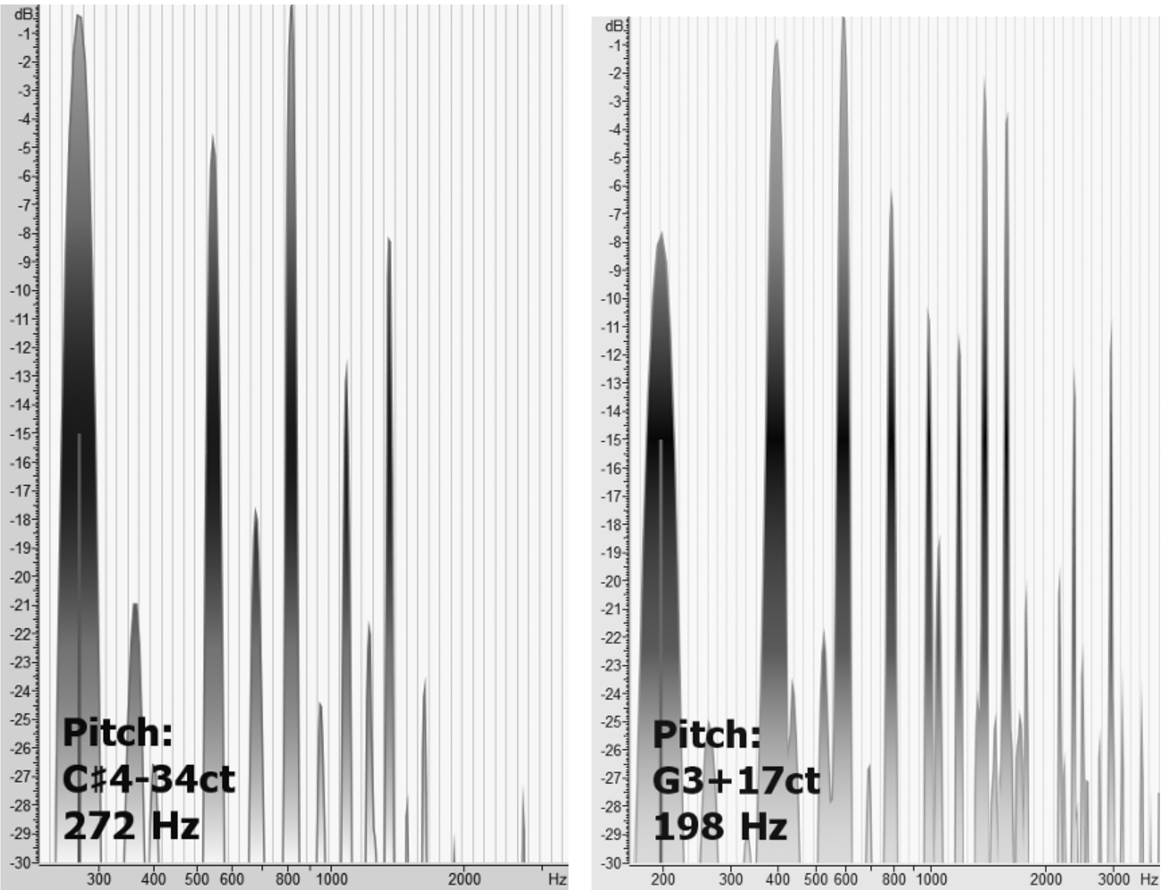


Figure 4.3: Spectrum readings for AKK: 0:19.000s (left) and FK: FK: 0:08.500s (right)

	Abdul Kareem Khan					**Faiyaz Khan**				
Resonance Area 1	Hn	Freq. (Hz)	Intensity (dB)	ASTC	Roughness	Hn	Freq. (Hz)	Intensity (dB)	ASTC	Roughness
<700Hz	H1	269	0	u	Pur, Res	H1	199	-8	u	Pur, Res
(Depth, Warmth,	H2	544	-5	o	Pur, Res	H2	398	-1	u	Pur, Res
Roundness)						H3	592	0	o	Pur, Res
	H3	813	0	ɔ	Pur, Res	H4	791	-6	ɔ	Pur, Res
Resonance Area 2	H4	1082	-13	ɑ	Pur, Res	H5	991	-11	ɑ	Pur, Res
800Hz - 2.1 Khz	H5	1357	-8	a	Pur, Res	H6	1190	-11	a	Rou, Res
(Clarity, Projection,	H6	1631	-24	æ	Rou, Res	H7	1389	-2	a	Rou, Res
Brightness)						H8	1588	-4	æ	Rou, Res
						H9	1782	-20	æ	Rou, Unr
						H10	1992	-31	e	Rou, Unr
	H10	2719	-28	i	Rou, Unr	H11	2180	-20	e	Rou, Unr
	H11	2993	-34	i	Rou, Unr	H12	2379	-13	e	Rou, Unr
						H13	2568	-27	i	Rou, Unr
Resonance Area 3						H14	2778	-25	i	Rou, Unr
2.1Khz - 4 Khz						H15	2972	-11	i	Rou, Unr
(Ring, Projection,						H16	3176	-23	Bright i	Rou, Unr
Forwardness, Brightness)						H17	3375	-31	Bright i	Rou, Unr
						H18	3569	-24	Bright i	Rou, Unr
						H19	3768	-28	Bright i	Rou, Unr
						H20	3967	-28	Bright i	Rou, Unr

Table 4.2: Boosted harmonics in AKK and FK's voices. Source: Author's analysis.

4.8 Observations

The following observations and conclusions can be drawn from the above data, when correlated with the above descriptions of these vocalists' voices:

The first, most obvious observation is that FK's voice has 20 clearly discernible, audible harmonics while AKK's has only 8. This is a significant difference and immediately provides one explanation of the terms 'rich' and 'broad' for FK's vocalisation and 'thin' and 'pointed' for AKK's. One of the reasons for FK's dense harmonic content is that he sings at a lower tonic pitch than AKK does[22]; but this is neither the only nor the most important reason, as we will see below. The following will discuss in detail how and why AKK's vocalisation de-emphasises his upper harmonics while FK's does the opposite, the nature of this selective emphasis and its implications.

4.8.1 Resonance Areas and ASTC

In comparing the resonance areas (RA) of these two voices, we find that AKK seems most focused on RA1, since RA2 and 3 are not nearly as prominent in his voice as RA1. In FK's voice, on the other hand, RA1 certainly plays an important role, but RA2 and 3 are densely populated with a large number of harmonics, many of which are prominent and aurally significant. This fact allows us to draw a number of conclusions. Firstly, RA2 and 3 are both associated with the perception of brightness, and this is the reason FK's voice appears to be 'brighter' than AKK's does[23].

Additionally, within RA1, the relative intensities of the harmonics are significant. For AKK, H1 is the loudest harmonic, while for FK, H2 and H3 are both louder than H1. These intensities are the result of the formant positions of these voices. We have no way of directly measuring the vocal tract shapes and sizes and thereby the accurate formant frequency positions of these vocalists, but estimations of can be made based on the vowels they are singing and on the observable characteristics of their timbre. Formant positions are a result of both the singer's technique as well as the vowel they are singing at the moment of measurement. In our samples, AKK sings a slightly rounded /a/, tending towards an /ɑ/ vowel while FK sings an /a/ vowel that tends towards an /ɛ[24]. The first formant locations of both these vowels are approximately the same (around 600-700Hz), but second formant positions are very different - /ɑ/ has its F2 at around 1000Hz, while /ɛ/-tending vowel has it at around 1500Hz-1700Hz[25]. AKK's /ɑ/ vowel shape tells us that his first formant should be at around 600-700Hz. But AKK's H1 (269Hz) appears to have been boosted by F1 although it is not close enough to F1 to receive such a boost. This tells us that he must be singing with a lowered larynx[26], a common way of elongating the vocal tract in order to lower F1. This is not the case with FK. Since FK's F2 is higher, it boosts harmonics around the 1500Hz region, which can be seen in the boosted intensities of his H7 and H8 harmonics, which harmonics are conspicuously absent in AKK. We know from Bozeman that '. . . the higher the first formant tuning, the more open the pronunciation will sound. . . ' (2013 [Boz13], 14) and vice versa. This explains why FK's sound appears more 'open', while that of AKK appears more 'close'[27].

Clearly then, most of AKK's resonance strategy depends on the importance it gives to RA1. This consistent emphasis on H1 and H2 also presents a possible explanation of Ranade's comment above, that AKK's voice has 'aas'. I contend that the lack of upper harmonics and the emphasis on the resolvable lower harmonics means that the resulting tone is extremely consonant with the droning tanpuras. It is probably the case that this

[22] 'A lower Fo (lower pitch) has both more, and more "closely spaced" harmonics within audibility' (Bozeman 2013 [Boz13], 7)

[23] See Figure 1 above for the dark->bright continuum in the harmonic spectrum

[24] I contend that this 'vowel modification' carried out, perhaps unintentionally, by these vocalists is part of their overall resonance strategy. More on this below.

[25] The above formant positions are approximations based on the authors analysis in VoceVista, which allows formant positions to be approximated based on aurally matching the vowel being sung. (Catford and Catford 1988 [Cat+88]) was also used as a reference.

[26] See section on Physiological Function below.

[27] The term 'closed' is used in informal discourse rather than 'close', but Bozemen et al advocate the use of 'close' instead since 'closed' has undesirable implications in vocal pedagogy. Since this sort of analysis has a direct relevance to vocal pedagogy, I have retained Bozeman's usage.

relatively sinusoidal, 'pure' sound that AKK creates appears so consonant with the tanpuras (tambura) that even when he stops phonating, the tanpuras cause the impression of that sound to linger. Pt. Kumar Gandharva – another vocalist known for his exceptional tonal accuracy – used famously to say that his tanpuras 'carried his sound forward' when he stopped singing. This is a phenomenon that admirers of these musicians' music grow to value greatly and may at least partially be explained acoustically in this fashion[28]. This same lack of upper harmonics, and the resulting sparsely populated signal may also be why Ranade describes AKK's voice as 'hollow' – much like a flute, which is among the most sinusoidal of instruments, is also described.

The harmonics in RA1 have the ASTC quality of /u/ and /o/, which vowels are associated with a 'close', 'dark' sound. These vowels are also associated with 'roundness' and while AKK's voice is conventionally called 'rounded'[29], it bears mentioning that FK's voice also has a roundness component to it. His H1-H3 are not as prominent as AKK's are, but they still contribute significantly to his timbre. FK supplements this roundness, though, with prominent harmonics in the RA2 area, which have the ASTC qualities of /a, æ, e/. This is why his overall timbre appears less rounded than that of AKK's which only has small, less significant boosts in RA2. Another fallout of FK's emphasis on RA2 is that FK's vowels appear to have more clarity than AKK's. As described above, harmonics within this region are responsible for all the vowel colours except /u/ and /o/. FK's emphasis on this area explains this clarity, while also explaining AKK's lack of it[30].

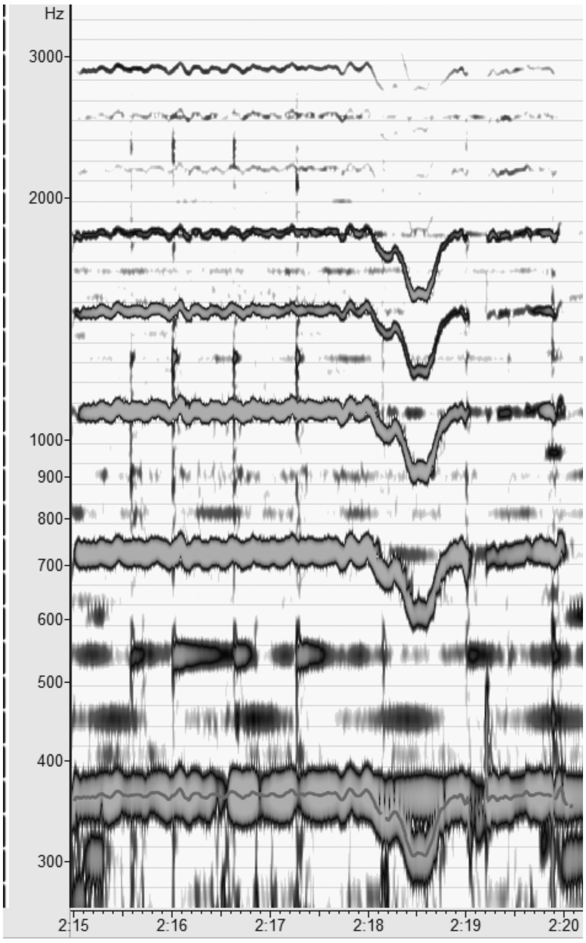


Figure 4.4: Abdul Kareem Khan singing his upper Shadaj (tonic)

In RA3, AKK has only two, barely audible harmonics, while FK has 10, many of which have significant intensity. We know that this area is associated with perceptions of 'ring' and we can ascribe the sharp, ringing quality of FK's voice to the boosts in this region. Psychoacoustics also tells us that the human ear canal has a 'primary resonance in the 3000-4000Hz range' and that 'any harmonics entering the ear canal within those frequencies will receive a strong resonance boost on their way to the eardrum' (Bozeman 2013[Boz13], 17). The fact that AKK's vocal tract configuration boosts virtually no harmonics in this range, while FK's boosts several explains why FK's voice is perceived as 'powerful' and 'loud' in comparison.

Regarding Ranade's comment that AKK's voice is 'resonant in segments', this probably refers to moments in AKK's singing where higher harmonics do get boosted. An example of this can be seen in figure 4.4, taken

[28]It is my opinion that this phenomenon deserves a more thorough investigation, which is beyond the scope of this paper. One hypothesis that comes to mind, especially regarding Kumar Gandharva's vocalization, is that in addition to the emphasized fundamental, there are probably specific upper harmonics being boosted in his voice to match and reinforce those same harmonics in the tanpura's sound.

[29]See Ranade's quote above, which describes this as AKK's 'reliance on the compound vowel 'o''. Other commentators have described this roundness in other ways.

[30]AKK and his followers (of his Kirana Gharana) have gained a certain notoriety for their unclear pronunciation, and for the textual content of their performances being less intelligible. See the following section for a discussion of FK's and AKK's text use and its relationship with their timbres.

from another point in the same AKK recording analysed above. This has AKK singing his high Sa[31]. As seen here, while most of the features of the voice (small number of boosted harmonics, emphasis on lower harmonics etc) are retained, we can see that while AKK is on his high Sa (2:15-2:18), he achieves a boost in the 3Khz region. This boost disappears between 2:18 and 2:19 when he sings lower notes and reappears after 2:19 (AKK sings the /a/ vowel here on the higher as well as the two lower notes). Also noticeable is the fact that H5 is de-emphasised while singing the high sa, but conversely receives a boost when singing the lower notes. We have already seen how frequencies in the 3KHz region are perceived by the human ear as being louder, and it is no surprise that these intermittent boosts are perceived by listeners as 'occasional resonance'.

4.8.2 Roughness and Resolvability

As we have seen, FK has 20 discernible harmonics in his signal, while AKK has only 8. This means that the number of 'rough' harmonics (see Howell on roughness above) present in FK's voice is 15 (H6-H20) while AKK has only 3 (H6, H10, H11). This explains the use of the terms 'pure' and 'smooth' for AKK and 'gruff', 'grating' and 'rugged' for FK. In terms of resolvability, AKK has only 2 unresolved harmonics (H10 and H11) and both of those do not have significant intensity. FK on the other hand, has as many as 12 unresolved harmonics in his signal (H9-H20). While this certainly contributes to the perception of roughness in his voice, I contend here that this might also contribute to the perception of a relative lack of tonal accuracy in FK's singing. Although I have been unable to locate literature that directly correlates the perception of having 'correct intonation' with Howell's 'roughness', the link does not seem a difficult one to make[32]. To quote Howell again, 'If the harmonics forming a spectral peak are low enough in the series, that spectral peak's tone color becomes an aspect of the pitch with a pure quality. If high enough, that tone color elicits roughness and eventually escapes the pitch'(2017 [How17], 5). Clearly, in FK's case, a number of prominent harmonics in his voice 'escape' the pitch and may lend to the overall tonal construct an appearance of having inaccurate intonation even if his H1 is accurately on pitch. It is perhaps the emulation of this quality by FK's followers that led Deshpande to say that they '. . . ignored values connected with swara – its sweetness, smoothness . . . ' (1987 [Des87], 41–42). In contrast, most of AKK's harmonics resolve into his fundamental, resulting in the perception of extreme tonal precision. This seems to be the quality that caused so many commentators to say, like Deshpande, that '[Abdul Kareem Khan] and his disciples carried [tonal accuracy] to such perfection that to them it became the very essence of [their] musical art' (1987 [Des87], 42).

4.8.3 Physiological Function

The spectral envelope of a singer's voice can give us some important clues about the physiology of their vocal apparatus, specifically about the kind of registration they employ, as well as their laryngeal position. A singer's 'registration' refers to how the vocal fold muscles are configured and coordinated in order to regulate pitch and has important implications for the singer's resulting vocal ability and character. Understanding a singer's laryngeal position is critical to developing an understanding of his timbre, since laryngeal position determines the positions of the formants in the vocal tract[33].

FK's spectral envelope tells us that FK was probably making use of what is known in vocal pedagogy as 'pressed phonation', a mode of phonation where the vocal folds are more forcefully adducted and therefore resist airflow from the lungs, thus requiring increased breath pressure. To quote from Bozeman again:

[31] This portion of the recording begins at 2:15min in the same AKK recording analysed above. See discography section below for recording details.

[32] This is probably what Bozeman means when he notes that 'pitch. . . can be [perceptionally] "flatted" or "sharped" from the actual fundamental frequency by resonance factors' (2013 [Boz13], 7 footnote 6)

[33] '[Vocal tract] length determines the general location of the entire formant set [so that] the longer the tube, the lower the frequencies of the formant set, and the more that fall within aural significance' (Bozeman 2013[Boz13], 12) and vice versa. Vocal tract length is determined by the position of the larynx.

> Pressed phonation (a phonation with excess glottal resistance) has a moderate... first harmonic (H1), a gradual roll off in power, and thus relatively more and stronger high harmonics. It is therefore metallic or overly bright in timbre. (Bozeman 2013[Boz13], 5–6)

The above is substantiated by the fact that, as we have seen, FK has a relatively subdued H1, more and stronger higher harmonics and is described as 'bright' in timbre. This leads FK to have what is called a 'chestier' register which has 'a more complex pressure waveform with more, stronger high harmonics' as opposed to AKK's 'headier' register, 'creating a more sinusoidal pressure wave pattern with a strong Fo(H1), but fewer, weaker high harmonics' (ibid, 2013 [Boz13], 6–7)[34].

While Sundberg corroborates the fact that 'pressed [phonation]... acoustically is associated with a decrease of the amplitude of the fundamental', he also attributes this to the singer having a raised larynx (1994 [Sun94], 112). We have speculated above that AKK's laryngeal function was different from this, that he was probably singing with a lower larynx than FK was. Additionally, AKK's voice was characterised as 'pliant' as opposed to FK's 'inelastic' voice[35]. This could lead us to the conclusion that AKK used what is known as 'flow phonation' - a mode of phonation that differentiates itself from pressed phonation by its more balanced harmonic spectrum and lighter adduction between the vocal folds[36]. These insights into these vocalists' physiological function have implications for developing an understanding of their musical choices, as will be discussed below.

4.9 Of Aesthetics And Acoustics

As demonstrated above, data obtainable from timbral analysis serves the important function of allowing us to concretize as well as complicate conventional descriptions of these singers' vocal timbres. In addition, it is also possible to use this analysis to enrich discussions about their larger musical aesthetic and its reception among audiences and commentators. A discussion on these lines is presented below.

FK's music has been called 'speech like' or conversational, both in its usage of dramatic tonal and dynamic contrast as found in animated speech, as well as in its informal, earthy comportment. Commentators describe ?? where he would 'involve' the audience as if in conversation. The comparisons of FK's music with speech also stem from the fact that unlike many other Khayal singers, FK was a native speaker of Braj Bhasha, the language of most Khayal song-texts, and was thus very invested in the literal meaning of the poetry he was singing[37]. Timbral analysis allows us to correlate this 'speech like' quality with the timbral qualities of human speech itself: Howell points out that '... humans frequently speak in a lower and more compact pitch range than they sing. Generally, this fills all ASTC ranges with one or more harmonics... ' (2017 [How17], 7). This timbral insight gives rise to an important question: is there a symbiosis between FK's naturally low tessitura (and the resulting timbre which is also speech like) on the one hand and his affinity for text and the conversational nature of his music on the other?

Abdul Kareem Khan and many of his followers, in stark contrast, were notorious for their unintelligible pronunciation. Their timbral goals again offer an explanation. Intelligible text depends upon clearly articulated consonants. In human speech, consonants exist in the upper ranges of the spectrum and articulating them clearly would therefore mean introducing roughness into the signal[38]. This seems an appropriate explanation for AKK's strategy of deemphasising consonants, given his aesthetic goal of always producing a pure, nearly

[34] See Bozeman for a detailed discussion on these registers, commonly known in vocal pedagogy. The 'chestier' register is also known as thyro-arytenoid dominant or 'Mode 1' while the 'headier' register is known as the crico-thyroid dominant or 'Mode 2' register.(2013[Boz13], 6)

[35] See Wade (2016 [Wad16], 108)

[36] This is speculation based on a cursory study. A more thorough, exhaustive study of these vocalists' recordings would be needed to make a more authoritative statement on their mode of phonation.

[37] See Ranade (2011 [Ran11], 126–27) for a detailed discussion of FK's use of text

[38] or introducing noise, i.e. aperiodic, non-harmonic frequencies

sinusoidal tone. Together with his favouring a round, /o/ dominant timbre, we now have an explanation for AKK's famed unintelligibility[39].

When describing Faiyaz Khan's performative comportment, most commentators call it 'colourful'[40]. This has of course to do with FK's extroverted approach to singing (which included evocative gesticulation), the dramatic dynamic contrasts he employed as well as his characteristic coquettishness, all of which gave his performances the air of festivity and cheer. But this 'colourfulness' is equally a description of his vocal timbre and pointing out the link between this adjective and the fact that his voice literally had a large number of tonal colours (boosted harmonics, each with its own ASTC quality) is irresistible. This is especially so in comparison with AKK's (timbrally) 'monotonous'[41] (literally having a mono, or single, tone) style, which we can now explain by his overbearing dependence on RA1 and an emphasised H1. Importantly, we know now that FK's numerous tonal colours are created, in part, by the overbearing presence of the /æ/ and /ɛ/ sounds in his vocalisation, hence explaining his vowel choice. The vowel preferences of these two vocalists can thus be understood in the context of their resonance strategies, as can be their deviation from the conventional /a/ sound noted by Ranade above. From a musicological perspective, my contention here is that since Khayal is a genre of improvised music that uses flexible, melismatic vocalisation to create complex musical structure, singers are often less concerned with 'correct' vowel pronunciation and more with creating a resonant sound and therefore tend to modify their vowels freely to that end. The fact that the genre allows vocalists this freedom implies both that this music is able to freely incorporate a variety of timbral colours, and that it does not have a theory of 'correct' pronunciation in the way Western operatic music does.

Another important and inescapable connotation timbre has is that of gender. Lower male voices typically have prominent upper harmonics and, as we have seen, consequently more roughness, which higher/female voices don't. The conventional labelling of FK's voice as 'masculine' is clearly a result of this. What makes this straightforward case of gender stereotyping (lower[42] and rougher = more masculine) even more interesting though, is Faiyaz Khan's fondness for 'feminine' coquettishness in his singing, an approach he is known to have picked up from the many courtesan performances he witnessed in his early years. The fact that the singer most known for his 'masculine' voice incorporates a kind expression into his music that is recognised in his time as distinctly feminine, and retains the masculine tag nonetheless raises a number of questions: is FK's (acoustically) 'forward', 'bold', 'unabashed' timbre connected with the tawaaif's identity as a (socially) 'forward', 'bold', 'unabashed' woman? On listening to available recordings of singers from the tawaaif tradition that Faiyaz Khan would have been influenced by (such as Gauhar Jan and Zohra Bai Agrewali), one is struck by the 'rugged' and 'powerful' nature of their vocal timbre, which stands in stark contrast to later ideals of the feminine voice in Indian music (as exemplified by Lata Mangeshkar)[43].

If Faiyaz Khan's music is colourful, Abdul Kareem Khan's music tends to be characterised as spiritual and laden with pathos. 'Khansaheb had a tonal beauty and accuracy which had the power of sending his audience into a trance-like state. . . ' (Deodhar 1993 [Deo93], 150). Or as Ranade says, '. . . it is clear that raga shankara has a mood that is markedly different from that of raga jogiya. . . In [AKK's] music, however, there is only one deep hue of anguish all over and in all ragas' (2011 [Ran11], 71). While Faiyaz Khan's 'colourfulness' had to do with his overall musicianship as much as with his tone, it could be argued that Abdul Kareem Khan's pathos is predominantly a result of his timbral aesthetic – or the fact that pure timbre and intonation were in fact among his primary musical goals. This raises the important possibility of a connection of more sinusoidal,

[39]Another possible reason might be the desire to maintain continuity in the sounded melody, since consonants are largely unpitched sound and function as interruptions in the melodic contour.

[40]Rang or Rangeela is the term that is usually employed. This term is used literally as well as in reference to FK's great-grandfather Ramzan Khan 'Rangeele', indicating that the brand of musicianship FK inherited from him was colourful and that FK retained this quality.

[41]As described by Deshpande (1987 [Des87], 41)

[42]As an aside, an important irony: Faiyaz Khan's voice is often described as a 'bass' voice, while he actually deemphasizes the bass frequencies in his voice as we have seen. In contrast, AKK's voice is described as a 'treble' voice, when what he really does is emphasize the lowest, bass, frequencies in his voice!

[43]While this aspect of FK's music is apparent in his ;khayal performances, it is most dramatically noticeable in his thumri performances. As a contrast, Ranade points out that AKK's single-minded focus on pure tones and intonation is such that it makes his renditions of both thumri and khayal sound equally pathos-laden (2011 [Ran11], 68).

'pure' tones sung with correct intonation with 'spiritual' or 'non-worldly' and 'sorrowful'[44] affects on the one hand; and timbrally richer, upper harmonic dominant tones with 'worldly', 'cheerful/celebratory' affects on the other[45].

As we have seen in the section on physiological function above, timbral analysis shows us that Faiyaz Khan probably used pressed phonation while Abdul Kareem Khan probably used flow phonation, and we have also seen through the course of this discussion how the resulting timbre became an inseparable part of their overall musical aesthetic. This opens up the possibility of linking their acoustic and physiological choices with the formal and structural aspects of their music as well. It is perhaps because of the pliancy of AKK's voice (probably a result of his use of flow phonation) that he was able to 'produce the subtlest variations of swara' and to become so invested in this aspect of music-making that 'his style lacked a formal structure and proportionality... was devoid of rhythm play... and general discipline' (Deshpande 1987 [Des87], 41)[46]. On the other hand, Faiyaz Khan and his followers cultivated a style that 'had neatness and a beautifully proportioned form, it used... dramatic contrasts and rhythm play... but they exaggerated the importance of discipline and order and ignored values connected with swara' (ibid, 42). A fruitful study could thus be made, to see whether singers who favour pure timbres and correct intonation have historically been less inclined to formal, structural pursuits, and rhythmic accuracy[47]. Whether it is timbre and physiology that influence structural aesthetics or vice versa is certainly a chicken-and-egg question, and I believe a serious investigation into this question will prove fruitful.

4.10 Conclusion

If there is one thing the above analysis makes clear, it is that the khayal genre, for all its dependence on a canonical repertoire and standardized theory, incorporates a stunning amount of diversity. Its identity as a 'classical' music often causes researchers to look for patterns of conformance in it[48]. As briefly discussed in the introduction to this paper, this approach, when coupled with computational techniques, often turns into a search for compliance between performance and theory. It is hoped that the present study demonstrates how computational tools can instead be used to create descriptive rather than prescriptive understandings of musicians' aesthetic choices. A case in point is the idea of 'resonance': As we have seen above, instead of using computational tools to develop a standardised understanding of a 'resonant' voice, this study does precisely the opposite – both AKK's and FK's voices are resonant; but in very literal, acoustic terms, their resonance strategies are different – AKK's vocal tract resonances favour lower harmonics, while FK's favour upper harmonics. Thus, both voices are *differently resonant* and their respective admirers tend to ascribe *those meanings* to the term 'resonant', and may not call the other voice resonant for lack of that particular quality.

A study such as this is necessarily interdisciplinary since it brings together the disciplines of musicology, ethnomusicology, computational musicology, psychoacoustics and vocal pedagogy. In doing so it opens up a number of possibilities for research such as:

[44] Pt. Kumar Gandharva, whose timbre is often described as similar in many ways to that of AKK was also known to evoke pathos, and called himself 'shok-piya' – the lover of mourning. Abdul Kareem Khan is known to have descended from a lineage of sarangi players (and was himself also an accomplished been player). The timbre of the sarangi is conventionally described as pathos-laden as well. This is an interesting timbral link that is sure to reward investigation, although the sound of the sarangi is certainly characterised by a richer spectrum than that of AKK's voice.

[45] It bears mentioning that these timbral contrasts exist within the palette of a single vocalist as well. Although it is beyond the scope of this paper to describe in detail, a comparison this author carried out between the present Jaunpuri recording of FK and another one that has him singing Raga Todi showed that his timbre in the Todi is even rougher and timbrally richer than the Jaunpuri. Compared with the Todi, the Jaunpuri sounds more tonally accurate, rounder and smoother.

[46] See Ranade 64-67 for a detailed discussion of AKK's approach to rhythm in particular and formalism in general.

[47] Pt. Jitendra Abhisheki says in an interview that 'in the history of this tradition, it has always been the singers with vocal impediments [inelasticity/intonation problems] who made important contributions to musical [structural] thought' (Abhisheki and Deshpande, n.d., translation mine)[AD]

[48] Ostensibly a result of its 'classical' status as much as of the nature of the genre itself – see Powers (1980 [Pow80]), Schofield (2010 [Sch10]) and Manuel (2015 [Man15]) for enlightening discussions on this subject.

- Bringing nuance to traditional terminology by using repeatable, measurable methods (eg: all musicians want / claim to have a ‘resonant’ voice, but each has a valid but differing definition of resonance)
- Hypothesizing about the possible logic behind singers’ aesthetic, stylistic and even physiological decisions based on available timbral information
- Understanding audience reactions to timbral choices
- Understanding relationships between perceived vocal timbre and affect

It is hoped that this paper has served as an example of how timbral analysis can address the above issues and provide meaningful answers while also giving rise to richer questions.

4.11 Discography

(All recordings sourced from the archival collection of the Manipal-Samvaad Centre for Indian Music, Manipal Centre for Humanities, MAHE Manipal)
KHANSAHEB ABDUL KARIM KHAN; Raga Jhinjhoti: Piya Bin Nahin Awat; HMV STC 02B 6231 [cassette, 1989].
USTAD FAIYAZ KHAN SAHIB; Phool Banki Gendan – Jaunpuri; His Master’s Voice – EALP. 1292 [Vinyl, LP, Album, 1965]
USTAD FAIYAZ KHAN SAHIB; Garwa Main Sang Lage – Todi; His Master’s Voice – EALP. 1292 [Vinyl, LP, Album, 1965]

References

[AD] Abhisheki, J. and Deshpande, S. *Pt Jitendra Abhisheki - Samvaad Foundation Interview*. Samvaad Foundation, Mumbai.

[Boz13] Bozeman, K. W. “Practical Vocal Acoustics: Pedagogic Applications for Teachers and Singers”. In: *Vox Musicae Series* (2013). https://books.google.co.in/books?id=HT8qmgEACAAJ.

[Boz17] Bozeman, K. “The Pedagogic Use of Absolute Spectral Tone Color Theory”. In: *Journal of Singing* vol. 74, no. 2 (2017), pp. 179–183.

[Cat+88] Catford, J. C. et al. *A practical introduction to phonetics*. https://books.google.co.in/books?id=bEViAAAAMAAJ. Clarendon Press Oxford, 1988.

[DC05] Das, D. and Choudhury, M. “Finite state models for generation of Hindustani classical music”. In: *Proceedings of International Symposium on Frontiers of Research in Speech and Music*. 2005, pp. 59–64.

[Deo93] Deodhar, B. R. *Pillars of Hindustani Music*. https://books.google.co.in/books?id=cAkUAQAAIAAJ&source=gbs_similarbooks. Popular Prakashan, 1993.

[Des87] Deshpande, V. H. *Indian Musical Traditions: An Aesthetic Study of the Gharanas in Hindustani Music*. https://books.google.co.in/books?id=IDjaAAAAMAAJ&source=gbs_similarbooks. Popular Prakashan, 1987.

[How16] Howell, I. *Parsing the spectral envelope: Toward a general theory of vocal tone color*. New England Conservatory of Music, 2016.

[How17] Howell, I. “Necessary roughness in the voice pedagogy classroom: The special psychoacoustics of the singing voice”. In: *VOICEPrints* vol. 14, no. 5 (2017), pp. 4–7.

[Lic41] Lichte, W. H. “Attributes of complex tones.” In: *Journal of Experimental Psychology* vol. 28, no. 6 (1941), p. 455.

[Man15] Manuel, P. “The Intermediate Sphere in North Indian Music Culture: Between and Beyond “Folk” and “Classical””. In: *Ethnomusicology* vol. 59, no. 1 (2015), pp. 82–115.

[Pow80] Powers, H. S. “Classical music, cultural roots, and colonial rule: an Indic musicologist looks at the Muslim world”. In: *Asian Music* vol. 12, no. 1 (1980). https://doi.org/10.2307/833796, pp. 5–39.

[Ran11] Ranade, A. D. *Some Hindustani Musicians: They Lit the Way!* Promilla & Company, Publishers in association with Bibliophile South Asia., 2011.

[RR14] Rao, S. and Rao, P. “An overview of Hindustani music in the context of computational musicology”. In: *Journal of new music research* vol. 43, no. 1 (2014). https://doi.org/10.1080/09298215.2013.831109, pp. 24–33.

[SC12c] Srinivasamurthy, A. and Chordia, P. “Multiple viewpoint modeling of north Indian classical vocal compositions”. In: *Proceedings of the International Symposium on Computer Music Modeling and Retrieval*. Citeseer. 2012.

[Sch10] Schofield, K. B. “Reviving the golden age again:“Classicization,” Hindustani music, and the Mughals”. In: *Ethnomusicology* vol. 54, no. 3 (2010). https://doi.org/10.5406/ethnomusicology.54.3.0484, pp. 484–517.

[Sun94] Sundberg, J. “Perceptual aspects of singing”. In: *Journal of voice* vol. 8, no. 2 (1994). https://doi.org/10.1016/S0892-1997(05)80303-0, pp. 106–122.

[TA15] Titze, I. R. and Abbott, K. V. *Vocology: The science and practice of voice habilitation*. https://books.google.co.in/books?id=RubjjgEACAAJ. Compton Publishing Limited, 2015.

[Wad16] Wade, B. C. *Khyal: Creativity within North India's classical music tradition*. https://books.google.co.in/books?id=MiE9AAAAIAAJ&dq=aneesh pradhan&source=gbs_similarbooks. Munshiram Manoharlal Pub. Pvt. Ltd., 2016.

Chapter 5

Rhythm and structural segmentation in dhrupad bandish performance

Martin Clayton[a], Preeti Rao[b], Rohit M A[b]

[a] Department of Music, Durham University, UK
[b] Department of Electrical Engineering, Indian Institute of Technology Bombay, India

5.1 Introduction

In this chapter we approach the challenge of computationally segmenting performances of dhrupad bandish. Dhrupad is an important vocal style in the North Indian raga tradition, of great historical importance and considerable popularity in present times, albeit practiced less frequently than the predominant khyal style. Amongst the distinctive features of dhrupad are the extensive alap sections (unmetered raga expositions, generally without drum accompaniment), the use of distinctive talas (especially chautal and dhamar, neither of which is used in khyal), and a highly syllabic style (in which one text syllable is performed to one or two pitches, rarely more, unlike khyal which uses extensive melismas; Clayton 2020: 48-52). Our focus here is on the question of whether rhythmic features can be used computationally to structurally segment audio recordings. In this chapter we discuss the rhythmic and structural properties of dhrupad bandish performances, illustrating these with the help of computationally-extracted features, and also report on novel approaches that may contribute to a future system for automatic structural segmentation. These draw on approaches such as vocal syllable and drum stroke onset detection as well as tempo tracking and pitch tracking.

In a previous chapter, we presented structural analyses of music audio that explored, in particular, rhythmic organisation with examples drawn from dhrupad alap performances. The present chapter does not consider alap sections but only those presenting bandishes (vocal compositions), accompanied by the pakhavaj barrel drum and set in the most common dhrupad tal, chautal (of 12 matras). These bandish performances comprise the basic elements of

1. the composition itself,
2. (optional) 'fixed laykari'[1] or lay bant, in which the bandish is performed at different speeds relative to an unchanging drum accompaniment, and
3. 'free laykari' or bol bant, in which the text is used as the basis of melodic and rhythmic variation.

As we shall see these elements do not occur in a fixed order; in fact, their use and ordering may be distinctive of gharana or individual style. It would be interesting from a musicological point of view, therefore, to detect and label these sections. Within the 'free laykari' portion, the performance can be divided into a string of discrete episodes in which a rhythmic idea is presented, developed and completed: detecting this episodic structure can also help to track musical processes and has applications in musical analysis.

[1] The terms 'fixed' and 'free' laykari were coined by Sanyal and Widdess (2004: 239)[SW04]. On the terms lay bant and bol bant see also Thielemann (1997: 13-4)[Thi97]. Also see Thielemann (1997: 79ff)[Thi97] for a discussion of the differences in rhythmic terminology between Dagar and Darbhanga traditions of dhrupad.

One approach to analysing bandish performances is, therefore, to identify the metric cycles of the tala and then assign cycles to these different categories: bandish, free and fixed laykari. Within those sections we can look at the breakdown of laykari into discrete episodes, identify features such as the typical use of the text (e.g. whether only one line is used or several, whether whole lines of text must be retained or the words can be broken up at will; see Sanyal & Widdess 2004: 259 [SW04]), the nature of the pakhavaj accompaniment and the interaction between singer(s) and pakhavaj player. Given a sufficiently large and representative sample of recordings and an efficient way of segmenting them – in our approach, using computational methods – comparative study of individual and group styles would be facilitated.

Following an introduction to the typical rhythmic features of chautal bandish performances, we introduce a test dataset and the processes of manual annotation and computational analysis used to explore it, illustrating these with examples drawn from these performances. We explore the potential of current computational approaches to assist with rhythmic and formal analysis of dhrupad performances, and finally discuss computational challenges for future research.

5.2 Rhythm and form in dhrupad bandishes

Dhrupad bandishes, which normally follow unmetred alap sections in raga performances, are set in one of a small number of talas. As noted above, the most common by far is the 12-matra chautal, which is strongly associated with the genre, and which we focus on in this chapter. Next most common is the 14-matra dhamar, which is associated in particular with holi-themed songs: it serves as an alternative for slower bandishes. The other talas used are jhaptal (10 matras, the only dhrupad tala also common in both khyal and instrumental gat genres), tivra tal (7) and the fastest, sultal (also 10 matras). Singers may pair slow and fast tempo bandishes (for example in chautal and sultal), but this is much less common than in khyal or gat performance: often a single bandish is performed in each raga. As in other genres of Hindustani music, tala is maintained through a combination of hand gestures and stereotypical drum patterns called theka. The use of tala-related hand gestures is more common than in other genres: Talwandi Brothers perform a rigid sequence of gestures throughout;[2] in other styles it is not so systematic, but it is common for singers to reference the gesture sequences from time to time while singing.

Dha	Dha	Din	Ta	Tete	Dha	Din	Ta	Tita	Kata	Gadi	Gana
X		0		2		0		3		4	
Clap		Wave		Clap		Wave		Clap		Clap	
1	2	3	4	5	6	7	8	9	10	11	12

Table 5.1: Chautal: theka and hand gestures. The top line represents the theka, or stereotypical drum pattern; the second line gives conventiomnal notational symbols for the 6 sections (vibhag) of chautal's cycle, where X = sam, 0 = khali, and numerals indicate other talis (claps). The third line indicates the standard clapping pattern, and the last the matras.

Drummers accompany on the barrel drum pakhawaj. They play a mixture of theka patterns (standard tala-identifying sequences), solo sections (typically while accompanying the statement of the vocal bandish at the beginning, sometimes at other times in the performance), and a more active accompaniment style in which drummers respond to and interact with the rhythm of the vocal line, sometimes playfully (for instance, singer and drummer appear to be trying to put each other off the tala counting) and only intermittently referencing the theka. As Sanyal and Widdess write,

[2] See e.g. https://www.youtube.com/watch?v=9ng4xwhkdt4& from about 24 minutes (Talwandi Brothers perform Raag Chandni Kedar) or https://www.youtube.com/watch?v=4u-kL_YShp4 from about 12 minutes (Talwandi Brothers perform Raag Bageshri). In their case, the 'wave' gesture on matras 3 and 5 is performed as a sharp downward movement of the right hand angled away from the knee.

"It is unusual... for the pakhāvaj player to play the hekā for more than a small proportion of the performance. More typically, rapid, rhythmically intricate patterns are played, which disguise the progress of the tāl-cycle, and so far as possible the pakhāvaj player simultaneously imitates improvisations of the singer [...] He may return to the hekā for a few beats after each episode of improvisation, but will begin to improvise again as soon as the singer does so... During the laykārī improvisation by both performers, the pakhāvaj player may include hints of the hekā from time to time to aid the singer in keeping tāl... But while the singer is singing the composition, which is rhythmically much simpler and more predictable, the pakhāvaj player may exploit the opportunity to embark on much longer improvisations that depart completely from the hekā."

(Sanyal and Widdess 2004: p.9)

5.3 Dataset

In order to explore this practice in more detail, and test computational approaches, this chapter makes particular reference to a convenience sample of 4 dhrupad performances, all set in chautal. A number of talas are used in the dhrupad style, of which chautal has particular prominence: almost all dhrupad recitals include at least one chautal piece, and chautal is almost never used for any other genre. In selecting the pieces for this chapter, we considered the difficulty of obtaining a large enough sample to meaningfully compare structure and rhythmic style between talas, and opted instead to focus only on the most common, chautal.

Code	Style	Singer(s)	Pakhawaj	Raga	Year	Duration
GB_AhrBrhv_Choutal	Dagar	Gundecha Brothers (Ramakant and Umakant Gundecha)	Akhilesh Gundecha	Ahir Bhairav	2011	14:23
UB_Malkauns_Choutal	Dagar	Uday Bhawalkar	Manik Munde	Malkauns	1997	06:47
PNM_Bihag_Chautal	Darbhanga	Prashant and Nishant Mallick	Mrinal Mohan Upadhyay	Bihag	2016	10:30
RCM_Bhrv_Chautal	Darbhanga	Ram Chatur Mallick and Abhay Narayan Mallick	Ramashish Phatak	Bhairav	1988	08:55

Table 5.2: Dataset

The other main consideration was coverage of stylistic schools of dhrupad performance. As in other genres, performance style is partly a matter of individual choices and temperament, and partly something inherited through teaching lines. In current dhrupad performance, Sanyal and Widdess distinguish three main approaches, termed the Dagar, Darbhanga and Talwandi gharanas. The Dagar tradition is the best known, with several branches of the extended Dagar family and their students performing professionally; well-known non-family performers include the Gundecha Brothers[3] and Uday Bhawalkar. The Darbhanga gharana, associated with the court of Darbhanga in Bihar, has produced several noted dhrupadiyas, including those of the Mallick family. The Pakistan-based Talwandi gharana offers some distinctive approaches, especially in the use of fixed laykari; unfortunately, good quality recordings are not easily available, therefore they are not represented in this dataset.

We decided, therefore, to concentrate for this discussion on a set of four chautal performances by prominent representatives of the Dagar and Darbhanga traditions. They include both solo and duet (jugulbandi) performances, a common feature of dhrupad. Since details of presentation can depend not only on the gharana or individual singer but also on the raga and situation (for example, the duration allowed by the concert or recording session), we do not propose that these performances be taken as encapsulating and representing individual styles, or the differences as summarizing those between gharana styles. However, using a range of performers ensures a degree of variety in the dataset and points to areas that may be compared in a large-scale study.

Manual annotation of the four performances was carried out, with each sam (cycle boundary) marked: from these annotations, matra positions are estimated by dividing the time of each cycle equally into 12, to give

[3]Umakant Gundecha and Ramakant Gundecha (d.2019).

a reference. (There is no objectively correct time stamp for each matra: rather, the timings can be estimated by observing the onset times of relevant drum strokes, or by tapping along with the beat to give a subjective estimate. Times derived by dividing cycle durations, as here, should be considered indicative of approximate matra locations.) We also annotated the main features of each cycle, for example which line of the bandish was sung or what kind of improvisation was used. In each section and episode of the bandish performance the singer(s) selects what to perform: bandish, free or fixed laykari, the laykari based on particular speeds (subdivisions of the matra); the drummer complements this, sometimes matching speed and rhythmic patterns, at other times complementing it. Each section may therefore present a distinctive combination of speeds and rhythmic emphases. This ‘speed’ can be represented as a ‘lay ratio’ (i.e. subdivision of the underlying matra pulse),[4] and changes in this lay ratio help the listener to parse the concert. The lay ratio was manually annotated for each cycle in both vocal and drum parts.

5.4 Automatic rhythm analysis

We used a number of different computational techniques to complement and compare with manual annotations. One of the main challenges was to obtain the lay ratios computationally. In order to do so we need to identify onsets of sung syllables (for the singers) and drum strokes (pakhavaj players) and compare their density to the underlying matra. A specific event in the audio stream such as a sung note or a drum stroke can be detected by its onset that is associated with a rapid rise of signal energy. Signal energy variations are more pronounced in certain frequency regions depending on the acoustic timbre changes that accompany the onset. In the vocals, for instance, the consonant-vowel transitions in the singing of the lyrics are most strongly cued by increases in the mid-frequency region of 600 Hz to 2400 Hz while drum onsets are more uniformly spread across the entire range. Differencing of the short-time energy in the specific spectral regions thus provides an onset detection function (also called novelty curve) whose peaks indicate the time location and strength of the onset. Important parameters in the analysis include length of the smoothing window

We are interested in the individual rhythm patterns performed by the different artists. This requires us to separate the two streams of onsets. This is most easily done with a prior step of audio source separation as illustrated in the Figure [block diagram]. Source separation has been a widely researched topic in MIR for applications such as polyphonic music transcription. More recently, deep neural networks that have been extensively trained on multi-instrument ensembles of melodic and percussive instruments (mostly Western popular music genres) have been released in the public domain. We use one such source separation tool, Spleeter, with a convolutional neural network (CNN) U-net architecture that has been trained on large datasets to generate ‘masks’ for the recovery of each of vocals and accompaniment streams from stereo mixes (Hennequin et al. 2020 [Hen+20]). Due to the peculiar nature of Indian drums such as pakhavaj, and in particular, due to the harmonic nature of the strokes that overlaps with vocal characteristics, the separated streams are not perfect. The separated streams nonetheless turn out to be helpful for the rhythm processing task due to the integrity of the onset information corresponding to each of the vocals and drums.

[4] Clayton (2000: 154)[Cla08] coined the term ‘lay ratio’ for the ratio between the fastest sustained pulse level performed to the matra pulse level.

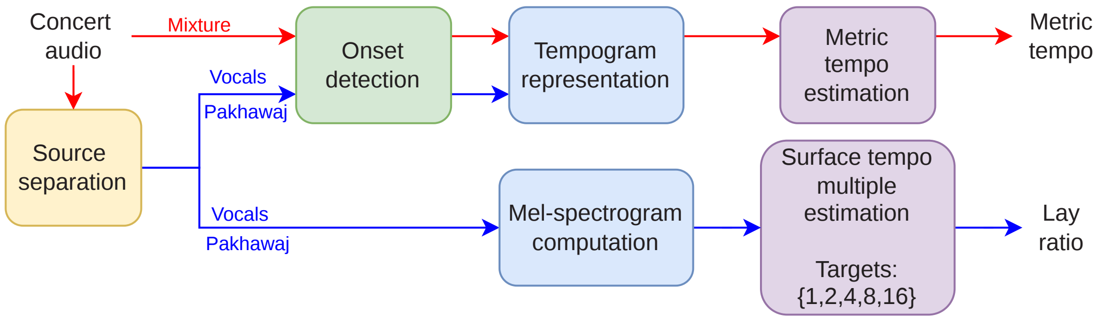


Figure 5.1: A block diagram showing the stages of our computational algorithm for the estimation of the metric tempo and the surface tempo of each source in the mixture.

A novelty curve of onset strength versus time is thus obtained based on the short-term spectral changes in each source-separated stream. The pattern of peaks in terms of their relative strengths and the durations of the inter-peak intervals captures the rhythmic emphases and density. The subsequent stage of rhythm analysis is applied to the computed novelty curve of each of the separated vocals and drums. The local rhythmic density is estimated by computing the periodicity of the novelty curve in a neighbourhood of the time instant of interest. That is, a windowed segment of this signal is compared to shifted versions of itself to obtain the autocorrelation function versus the amount of shift (lag). The autocorrelation of the novelty curve is expected to show peaks at lags equal to the inter-onset interval and its multiples. The autocorrelation function, computed at intervals of 0.5 s over windowed signal segments of duration 5 s for short and 10 s for long audio excerpts, can be converted to a 2-dimensional tempogram by inversion and interpolation of the autocorrelation values as discussed in Chapter MSA.

While the tempogram provides a visual illustration of the syllable or stroke density at any instant of time by way of the vertical separation of the peak contours in the 2d plot as discussed in Chapter X, computing quantitative estimates of the tempo (BPM) itself is more challenging. The automatic detection of tempo is typically plagued by octave errors, i.e. the optimality criterion employed can easily pick a sub-multiple of the true tempo ('true' meaning, in this case, the rate of the matra pulse). Prior knowledge about the tempo range can greatly help resolve such ambiguities. Another important point to note is that we are interested in estimating the surface tempo as well as the metric tempo. The latter is related to the inter-matra interval that provides the scaffolding for the instrument rhythmic patterns.

The metric tempo is estimated from the full mixture audio (i.e. both vocals and percussion contributing) by finding the lag value that maximizes the mean of the peaks of the autocorrelation function at that lag and the corresponding lag multiples (Vinutha et al. 2016)[Vin+16]. Further, restricting the metric tempo search range to 40 – 100 BPM reduces the chances of octave error. The surface tempo may be obtained in the similar manner but by biasing the estimate towards lower lags or higher tempi. Given the vast range of tempi spanned by the surface tempo, and given that it is expected to be a multiple of the metric tempo, it is also possible to invoke machine learning trained on labelled examples to directly infer the rhythmic density relationships (Rohit et al. 2020)[RVR20]. As depicted in Fig. 5.1, the algorithm works on the spectrogram of each separated source stream to classify each time frame into one of a set of expected tempo multiples. This set excludes the less frequently occurring non-binary subdivisions (4:3, 3:1, etc.) for simplicity and for their relatively rare appearances in the data set used to train machine learning algorithms. The challenge of accurately detecting onsets is overcome by this more implicit processing but, of course, at the expense of having at hand the requisite training data. Finally, algorithm outcomes at the frame level are subjected to a step of temporal smoothing to bring in the musical constraint that tempo changes occur at relatively long time scales, with the metric tempo typically evolving only gradually across the concert while the surface density registers abrupt jumps at musical section boundaries.

In the subsequent sections of this chapter, we apply our computational tools to the selected audio concert examples to capture the salient musical features of the bandish and layakari sections that characterise the dhrupad form while also serving as anchors for individual and stylistic variations. In particular, we capture the surface tempo (sub-division of the beat) for each of the vocals and pakhawaj. Our automatically detected onsets are superposed on the matras of the chautal cycle, estimated as described above.

Apart from representations of rhythmic structure, we present a plot of the melody as it evolves in time. The computed pitch of the singing voice represents the sung notes with their precise intonations, and is obtained with periodicity estimated via the autocorrelation of windowed segments of the source separated singing voice. The melodic analysis can provide complementary information and better insights in certain contexts. Finally, we show the spectrogram, a common visualization used to observe the development of musical parameters in time. Both the syllable onsets and melodic pitch information are derived from this basic time-frequency audio representation.

5.5 Bandish structures and text setting

Dhrupad bandishes are characterised by a syllabic style (Clayton 2000: 48ff)[Cla08]. Text setting in bandishes respects the metrical qualities of the verse (chand) to a significant extent. As Sanyal and Widdess point out, short syllables (notated S or U) are normally assigned one matra and one note, while long syllables (L or –) can take up two matras and typically 1-2 pitches (Clayton 2000: 11)[Cla08]. This is open to some variation though, for example with theoretically long vowels being treated as short to allow accommodation within the tala cycle. As Clayton points out, many chautal bandishes follow a pattern in which the text naturally falls into four 3-matra units (for example in sequences LS LS LS LS), creating a cross-rhythm when set against a clapping pattern that marks out 6 groups of 2 matras (Clayton 2000: 116)[Cla08]. This is not the only pattern that bandish melodies may follow, however; some bandishes also include a mukhra: that is, they begin a few matras before the sam.

5.5.1 Example 1: Bandish first line

Uday Bhawalkar's Rag Malkauns bandish, 'Jayati jayati Shree Ganesha', falls rhythmically into the pattern of 4 groups of 3, albeit the word Ganesha crosses this rhythmic boundary.

u	u	u	u	u	u	–		u	–		u
Ja-	ya-	ti	ja-	ya-	ti	Shree		Ga-	ne-		sha
1	2	3	4	5	6	7	8	9	10	11	12
dha	dha	din	ta	kita	dha	din	ta	tita	kata	gadi	gana

Table 5.3: Uday Bhawalkar, Rag Malkauns, first line. The top line indicates the length of each vowel in the text (second line). u = short syllable, – = long syllable: note that long syllables take up two matras each, short syllables one matra. Vertical lines indicate a rhythmic division into four groups of three matras. The bottom two lines show how this pattern is aligned to the tal structure and theka.

Figure 5.2 illustrates the first cycle of the bandish. The black curve in the top frame is the computed onset detection function, or novelty curve, of the singing voice. It is superposed on the audio waveform. We see prominent peaks in the novelty curve coinciding with those matra locations that contain actual sung syllable onsets. Most of the text syllables can be clearly located: an exception is the 'ya' in the second 'jayati' (matra 5) – as a semivowel this is less likely to sound as a clear rhythmic articulation. The spectrogram (lower frame) shows clearly, in the vertical black stripes, the word breaks (which divide the cycle into 3+3+2+4 matras): the singer takes a breath at these moments. The pitch track illustrates how most matras are marked by a rise in pitch, either up to the main svara (e.g. matras 1-3) or past and then back down to the svara (e.g. matras 9 & 10). The second matras of the syllables shree and Ganesha show a clear articulation in the pitch track (middle frame), but (not surprisingly) these do not show up strongly as syllable onsets in the upper frame.

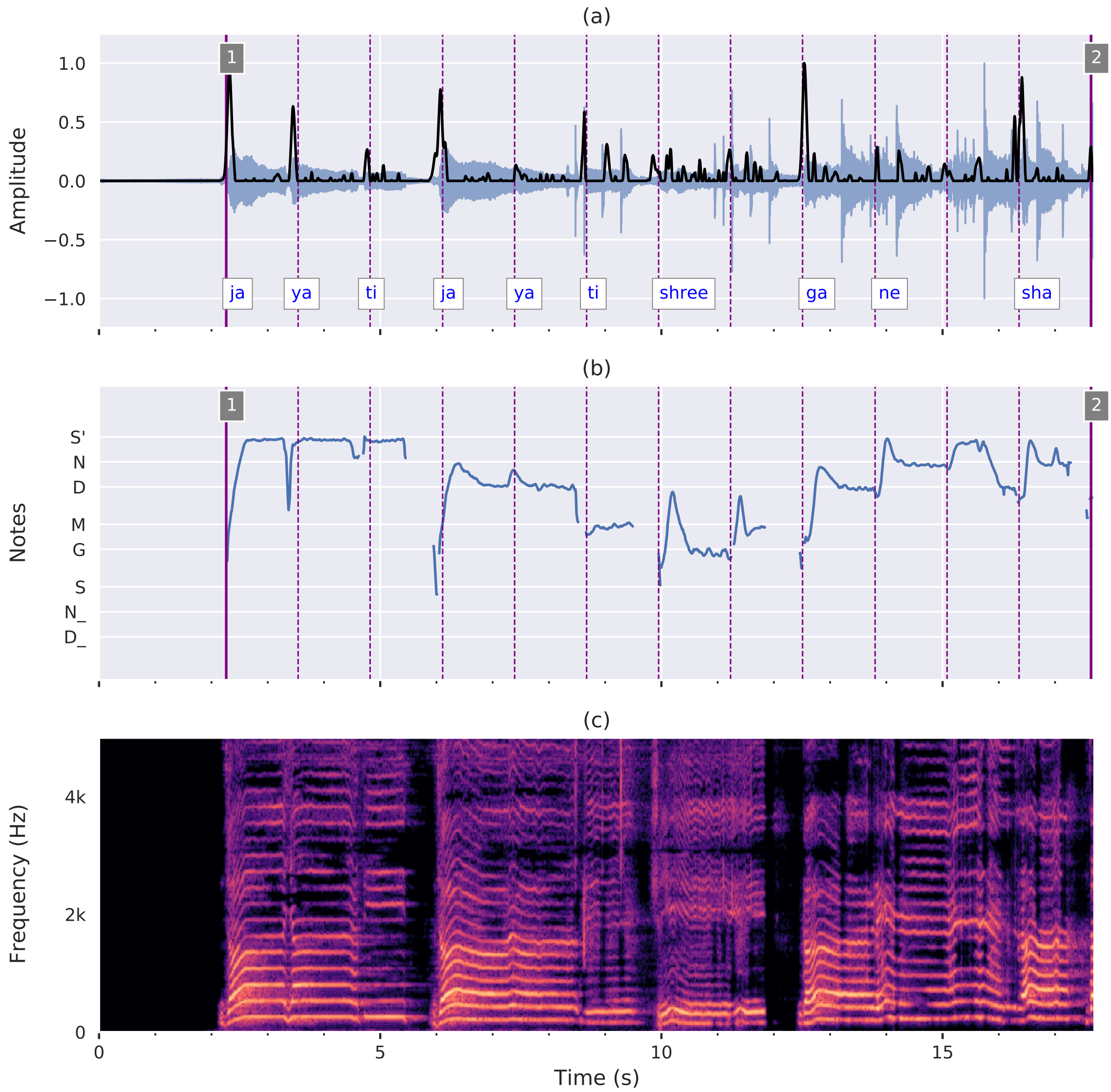


Figure 5.2: Uday Bhawalkar, Rag Malkauns, cycle 1. From the top: (a) waveform with the output of a vocal syllable onset detector in black, (b) pitch track; each has the 12 matras of the chautal cycle overlaid as dashed vertical lines, and (c) spectrogram.

5.5.2 Example 2: Bandish first line with mukhra

In contrast, Ram Chatur Mallick's bandish in Rag Bhairav, 'Lambodhara…', breaks the 12 matras up in an asymmetrical fashion, and starts from matra 8. Starting the bandish a few matras before sam – often, as here, 5 matras before – is typical of khyal and instrumental gat compositions, but as we see here also encountered in dhrupad. Compared with the previous example, we can see differences other than that the text line overlaps the tala cycle. The syllable onsets tend to occur just after rather than on the matras, and in general the matras themselves are not strongly emphasized in the vocal line (see Figure 5.3).

These two examples cannot illustrate the full variety of ways in which text lines can be set in chautal, of course. In practice, even if the first line follows the 4 x 3 pattern (in terms of rhythm, if not word breaks), other lines in the bandish will include some rhythmic variety; the mukhra, where used – as in Rag Bhairav example above – is not necessarily carried over to the later sections of the bandish.

–		–			u	u		u	u	–	u
lam-		bo-			dha-	ra		ga-	da	aa-	na
8	9	10	11	12	1	2	3	4	5	6	7
ta	tita	kata	gadi	gana	dha	dha	din	ta	kita	dha	din

Table 5.4: Ram Chatur Mallick, Rag Bhairav, first line of bandish text. Top line indicates the length of each vowel in the text (second line). U = short syllable, – = long syllable: note that the relationshuip between syllable length and matra count is now looser, although the first two long syllables receive two and three matras respectively. Vertical lines indicate a rhythmic division of the line, given as 5 matras (the mukhra, the portion before sam) + 3 + 2 + 2 matras. Bottom two lines show how this pattern is aligned to the tal structure and theka.

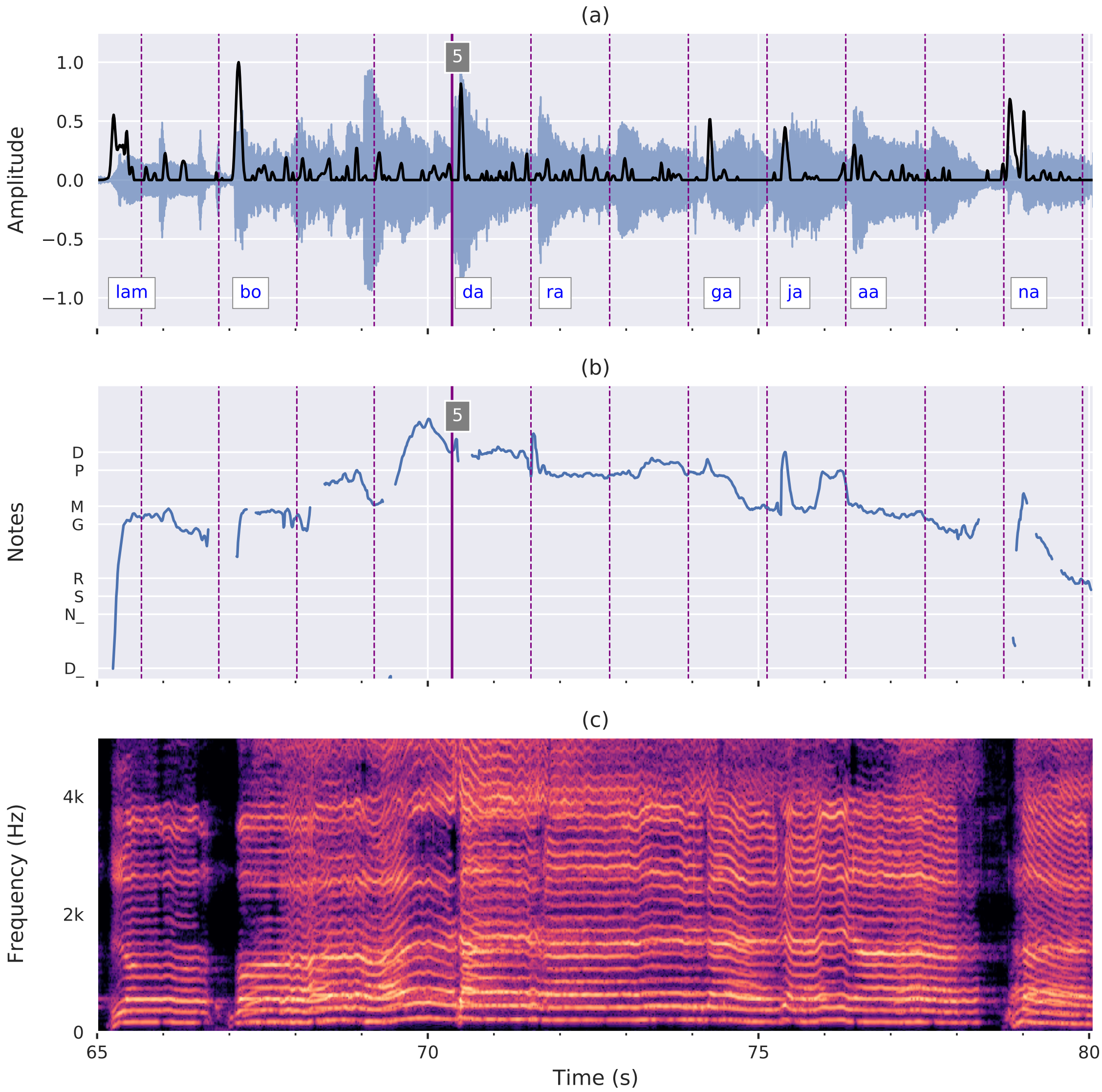


Figure 5.3: Ram Chatur Mallick, Rag Bhairav, first line of bandish text. From the top: (a) waveform with the output of a vocal syllable onset detector in black, (b) pitch track, and (c) spectrogram; each has the 12 matras of the chautal cycle overlaid as dashed vertical lines. In the vocal line, note that the syllable peaks fall just after the matra annotations. Although it is important to bear in mind that the latter are estimates and may not reflect the exact position of drum strokes marking the matras, in this case it does reflect the singer's placement of the syllables just after the beat. In the pakhavaj part, although the pattern is based on the theka there is significant variation in the last few matras before the sam (marked as 5) and on matra 5 of the following cycle.

5.6 Bandish sections

In many accounts, chautal bandishes have four sections, named sthayi, antara, sanchari and abhog, each comprising four cycles and four lines of text; the majority of recorded performances feature only two sections however, sthayi and antara. According to Sanyal and Widdess, in the Darbhanga tradition the final two sections are often merged into one, thus we have three sections, sthayi, antara and sanchari (2004: 223)[SW04]: in RCM_Bhrv_Chautal the sanchari and abhog are sung together without break as if a single section, while the other examples in the dataset include just sthayi and antara. In most cases the whole bandish is sung at the start; all or part of it may be repeated later, most commonly at the end of the performance. The four examples used here demonstrate different approaches:

GB_AhrBrhv_Choutal	Sthayi and antara sung in full at start
UB_Malkauns_Choutal	Sthayi and antara sung in full at start
PNM_Bihag_Chautal	Sthayi and antara sung in full at start, but separated by a section of free laykari
RCM_Bhrv_Chautal	Sthayi and antara sung in full and repeated at start; sanchari then sung once before free laykari begins

Table 5.5: Summary of approaches to bandish presentation in the dataset.

5.7 Laykari

Following the presentation of the bandish itself, singers engage in variations on the text, set in the same raga and tala. This may be termed laykari (rhythmic play). A characteristic feature of dhrupad is that it is sung exclusively to the text of the bandish. Laykari can be divided, in Sanyal and Widdess's terms, into 'fixed' and 'free' kinds. In 'fixed' laykari or lay bant a section of the bandish is sung as normally but at a different speed, typically double, triple and/or quadruple speed (thus a line of text occupies 6, 4 or 3 matras, with the rhythmic patterns retained). The use of such fixed laykarai varies greatly: in Talwandi gharana style it can take up the majority of a performance; Thielemann reports that it is used much more in Darbhanga than in Dagar style, although the Gundecha Brothers provide an example of Dagar gharana artists who employ this technique (1997: 80-81)[Thi97].

GB_AhrBrhv_Choutal	Free laykari, mostly binary subdivisions (2:1, 4:1) but also 3:1; Fixed laykari in 2:1, 3:1 and 4:1 to end.
UB_Malkauns_Choutal	Free laykari only, binary subdivisions
PNM_Bihag_Chautal	Free laykari only, mix of binary subdivisions with ternary (3:1, 6:1) and also 7:1
RCM_Bhrv_Chautal	Free laykari only, binary subdivisions

Table 5.6: Summary of approaches to laykari and lay ratio in the dataset.

Free laykari (bol bant) can be divided into a series of episodes in which the singer develops an idea for one or more tala cycles. Episodes are rounded off with a concluding (cadential) phrase at the end of the cycle. The most recognisable form of cadential phrase is the tihai (a thrice-repeated pattern which ends, in the simplest form, on sam). Tihais are used more commonly in Darbhanga than the Dagar style: more commonly in the latter, simple rhythmic figures are used to signal the end of an episode (Sanyal & Widdess 2004: 259)[SW04]. In our dataset, tihais are used heavily in PNM_Bihag_Chautal and more sparingly in the two Dagar style recordings.

Division of a performance into episodes is not always unambiguous. There are occasions when a singer appears to be concluding an episode, but then continues immediately in the following cycle; in jugulbandi performances, one singer may conclude his episode but in the expectation that their musical idea will be continued by their partner. Parsing performances into episodes often, therefore, involves decisions on the part of the listener, and different interpretations may be equally valid. In our annotations, episodes are identified on the basis of a tihai or other clear cadential figure, and/or the repetition of a portion of the sthayi section from the next sam.

Whereas lay bant typically accelerates through different speed levels – basic, double, triple, quadruple – bol bant accelerates in a more variable and less premeditated way, while the basic tempo also tends to gradually accelerate. What can be observed is that singers may begin to improvise bol bant with a pulse of one note/syllable to the matra; then the matra is subdivided in two and the greater rhythmic possibilities are exploited for a while; then (often finally) the matra is subdivided into four pulses. Other subdivisions are used: 3:1 (not in all performances, often only briefly), even 5:1 or 7:1, although these subdivisions are more rare. The vocal line may lie in another relationship to the matra beat, for example 3:2 (which could be achieved by subdividing the matra in 3 and then articulating every second pulse) or 4:3 (pulse subdivided in 4, every third pulse sung). PNM_Bihag_Chautal is perhaps the most systematic of our examples, progressing through 6:1, 7:1 and 8:1 before dropping back to 3:1. More typically we may observe a tendency to increase speed gradually but without observing a clear model, using mainly binary subdivisions with others (especially ternary) used occasionally for variety.

Pakhavaj accompaniment also features different speed levels, but the pakhavaj does not always match the speed of the singing. Sometimes, for example, a fast vocal passage will be accompanied by a slower theka-based accompaniment; or the opposite, the initial presentation of the bandish by the singer may be accompanied by a fast solo improvisation by the drummer. In PNM_Bihag_Chautal, however, the pakhavaj player tends to match the speed of the singers in the free laykari sections.

5.8 Example 3: Fixed laykari

The Gundecha Brothers use a section of fixed laykari to finish their performance in Rag Ahir Bhairav. In cycles 62-3 (from 783.8 secs) they perform the four lines of the sthayi at double speed, thus over two cycles. In the following two cycles they sing the same four lines at triple speed, repeating the last line three times.[5] In cycle 66 they sing all four lines in one cycle, before repeating the first line at the original speed to conclude.

In this case the pitch contour (see Figure 5.4) is most helpful in observing the ‘fixed laykari’ process. The contour of the complete sthayi is visible over two cycles (62 & 63 – descending-ascending-descending pattern). The same contour is repeated in more compressed form in the next 1 1/3 cycle, with the last section repeated to complete cycle 65. The same contour is compressed into a single cycle in cycle 66. The same pattern can be observed in the tempogram representations. The tempo is around 60 bpm (actually averaging 58 bpm for these cycles) and the most prominent dark line in each case runs at a tempo of c. 60 bpm. In the dugun section there is one faster lag on the vocal tempogram at around 120 bpm; in the tigun section there are two, around 90 and 180 bpm; in the caugun cycle, there are three, at roughly 80, 120 and 240 bpm (the last of these represents the 4:1 ratio; 80 bpm emerges as the rate at which groups of three pulses pass, and is not musically significant). In the pakhavaj tempogram , in contrast, the dugun and caugun sections show the 4 subdivisions (at about 60, 80, 120 and 240 bpm), while the middle, tigun section shows just two divisions. Cycle 63 has more lines due to the higher level of ornamentation in the pakhavaj part in this cycle. This indicates that rather than matching the singers’ ternary subdivision, the pakhavaj player plays more sparsely and concentrates on keeping the basic tempo.

5.9 Example 4: Free laykari (binary subdivision)

In ‘free’ laykari or bol bant (‘division of the text’), the singer may improvise on the text. Practice varies as to whether singers employ only the first line of the sthayi for these sections, the whole sthayi, or also the antara (bol bant rarely if ever employs the other sections). There is also variability in the way the text is deployed: whether the whole text line should be sung in full (but to different melodic and rhythmic patterns); or whether words or phrases may be repeated. Although some singers are stricter than others, generally speaking an important principle is that the meaning of the text should be respected.

[5]Note that this is a standard procedure, matching that set out by Thielemann (1997: 48)[Thi97].

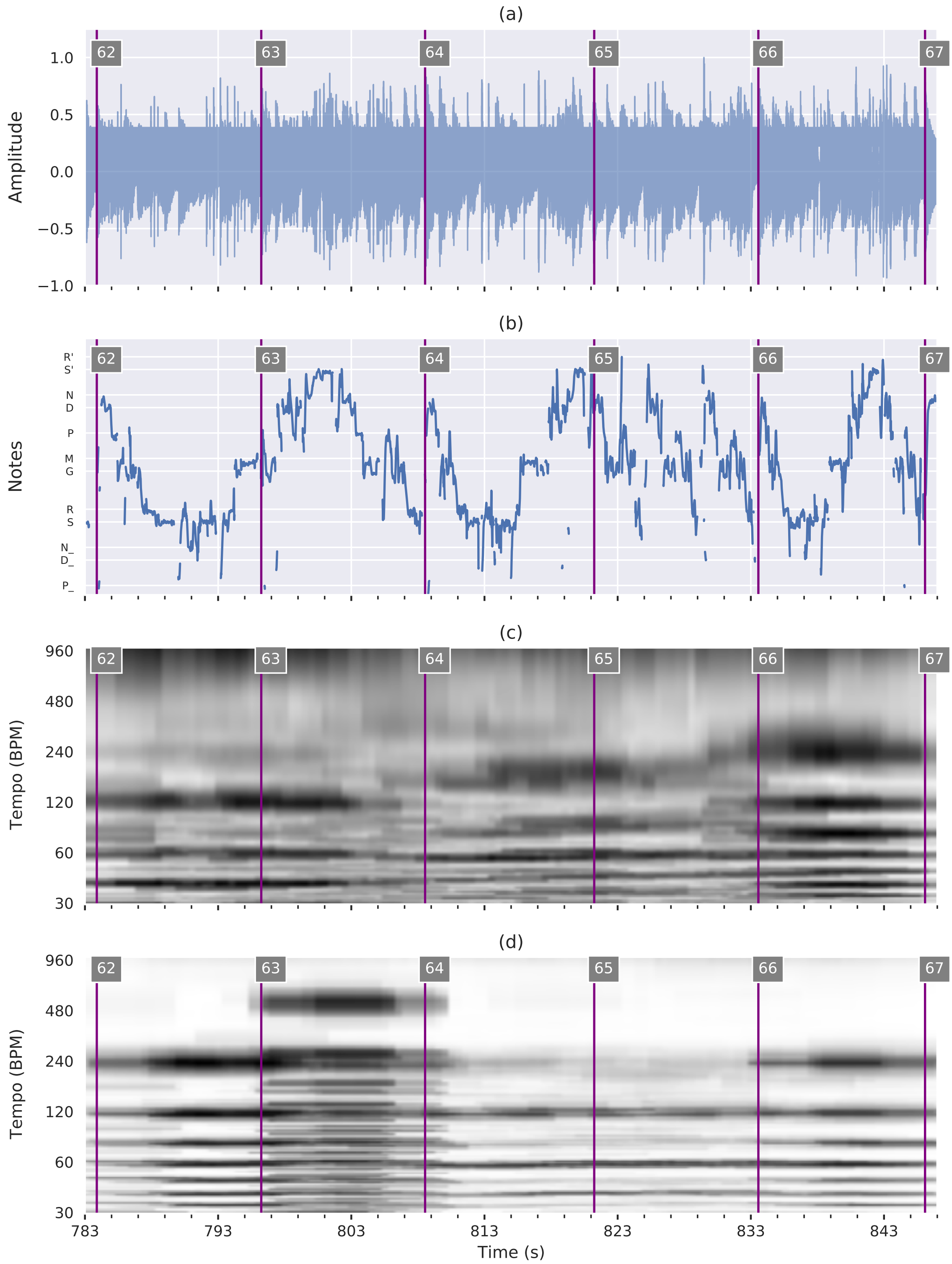


Figure 5.4: Gundecha Brothers, Rag Ahir Bhairav, lay bant (fixed laykari). From the top: (a) waveform with tala cycle boundaries overlaid, (b) pitch contour (from source-separated vocals), (c) vocal tempogram and (d) pakhavaj tempogram. Cycles 62-3 = dugun (2:1), cycles 64-5 = tigun (3:1), cycle 66 = chaugun (4:1).

The most common lay ratios are based on binary subdivisions[6] (2:1, 4:1, and rarely 8:1; 16:1 may occur in pakhavaj parts but is generally much too fast to sing), and the most common pattern for dhrupad performances

[6]We use numerals for the ratios (2:1, 3:1 etc) and the terms 'binary' and 'ternary' subdivision to group them. We also refer to common Hindi terms such as dugun (double), tigun (triple) etc. For a more comprehensive summary of traditional terminology see Clayton 2000[Cla08]:155ff.

sees a gradual acceleration of the vocal line. This example from Uday Bhawalkar's Rag Malkauns comprises one chautal cycle of 'free' laykari in binary subdivision: the episode comprises this and the previous cycle but only one cycle (18) is illustrated in Figure 5.5 for visual clarity. It illustrates a typical relationship of bol bant to the tala: the episode is concluded with a cadential figure (the rhythm clearly suggests a conclusion, and leads to the start of the bandish, 'Jayati. . . '. (The end of the previous cycle also has a 'closing' feel, but in that case Bhawalkar decided to continue the laykari into the next cycle.) Cycle 17 started with a clear 4:1 subdivision in the vocal part, slowing towards the end of the cycle; at the start of cycle 18, illustrated here, he only articulates a fraction of the 4 subdivisions, placing accents off the matra, before speeding up in the second half of the cycle. This kind of fluid shifting between different binary subdivisions, and more on- and off-beat feels, is typical of free laykari. The pakhavaj plays at a basic 8:1 subdivision, but slows towards the end of the cycle as the drummer attempts to match the singer's rhythmic patterns.

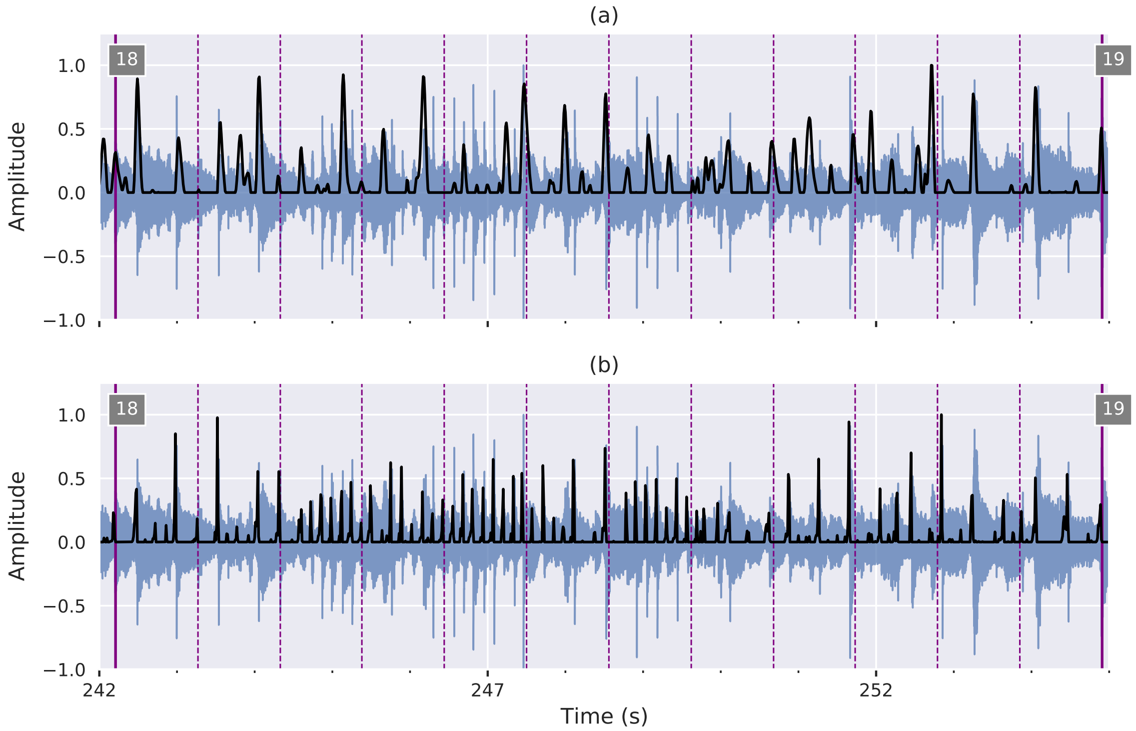


Figure 5.5: Uday Bhawalkar, Rag Malkauns, free laykari in caugun. Waveform (mixture audio) with matra boundaries overlaid and (a) syllable onset, and (b) drum onset detection functions in black. The drum onsets (lower pane) show the basic 8:1 division of the matra, especially between matras 3-7. The vocal onsets (upper pane) shows that a 4:1 division is more prominent, and the accented syllables are not always those falling on the matra (for example, matras 2-5 do not coincide with a syllable onset).

5.10 Example 5: Free laykari (ternary subdivision)

The next most common subdivisions are ternary (3:2, 3:1, 6:1; 12:1 in pakhavaj parts). Prashant & Nishant Mallick's Rag Bihag provides a good example of the use of ternary subdivision here (Figure 13.5). The pakhavaj accompaniment matches the vocal, also using a 3:1 subdivision, and the cycle is concluded with a tihai. The syllable onset function shows that in both the vocal and pakhavaj parts, there is a clear division of each matra in three. Not all subdivisions are articulated in the vocal line; for instance, the first subdivision of matra 8 is silent; the tihai features a fast articulation of three syllables on the second subdivision followed by a gap and the second singer calling out 'ek' on matra 11, 'do' on matra 12 before the climax on the following sam.

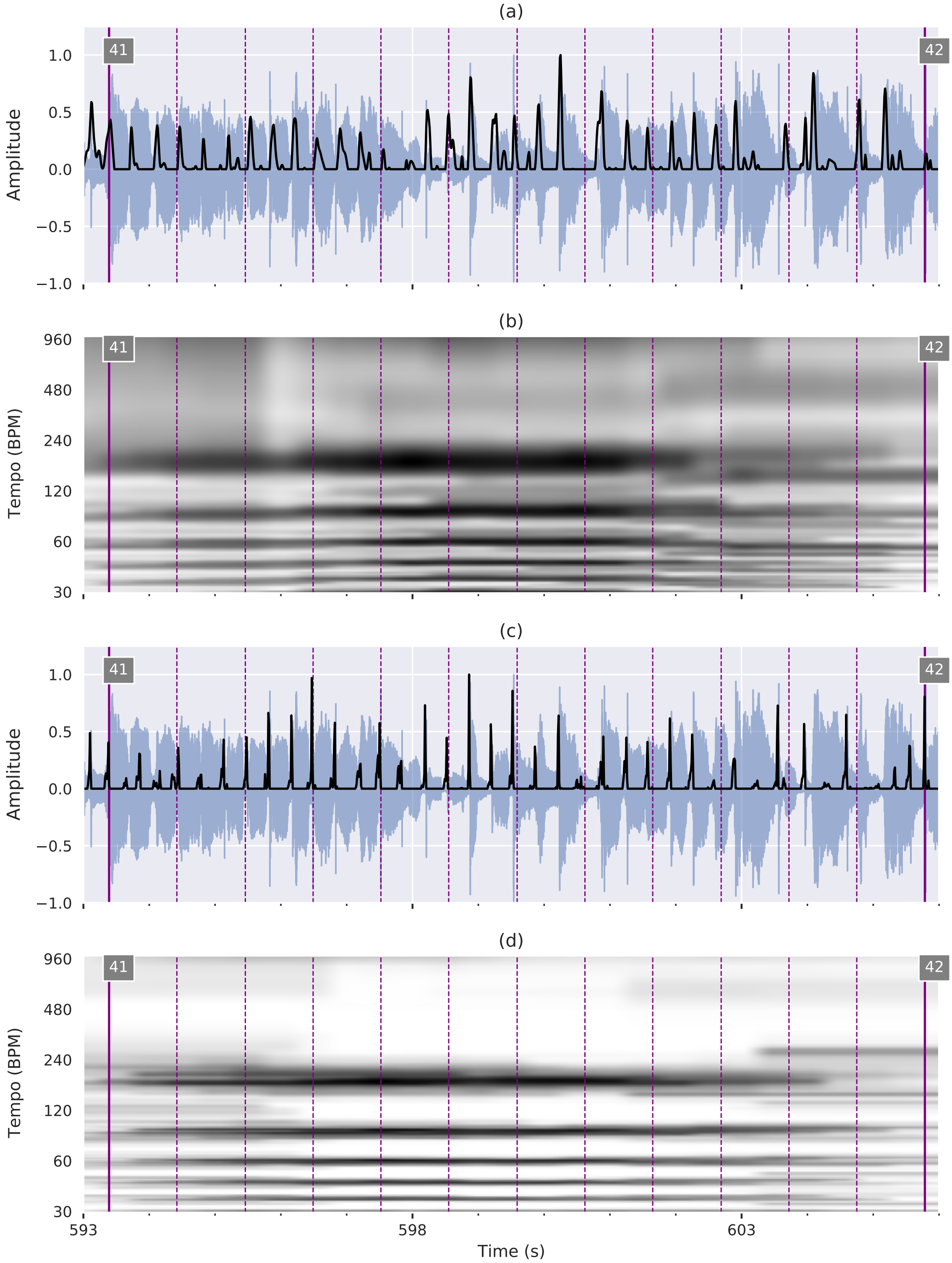


Figure 5.6: Prashant & Nishant Mallick, Rag Bihag, cycle 41. From the top: (a) waveform with the output of a vocal syllable onset detector in black, (b) vocal tempogram, (c) waveform with drum onset detection function overlaid in black, and (d) pakhavaj tempogram.

The pakhavaj plot also shows a clear 3:1 subdivision; in the last few matras when the tihai is sung, the strokes appear to come ahead of the matra subdivisions: in this part he is neither matching the vocal line nor maintaining the steady three, but responding in the moment to the tihai pattern. This pattern is supported by the tempograms, which show periodicities at close to 60 bpm (it is actually about 58 bpm), 90 and 180 bpm. The pattern is more clearly defined in the pakhavaj part, but similar in both.

5.11 Example 6: Free laykari in saatgun (7:1)

Some performances use different varieties of subdivision based on division into 5 (panchgun) or 7 (saatgun). This example of Prashant & Nishant Mallick's Rag Bihag features the latter (Figure 13.6). In this cycle (32) the pattern of dividing the matra into 7 is clearly audible from matra 3. On the onset detection this shows as a high density of peaks, although it is not obviously visually that the division is into 7. On the pakhavaj onset function on matras 9 & 10 in particular we can see a distinctive uneven division into three, where he plays a grouping of 3+2+2, articulating only the first of each group. On the tempograms we have added a reference line at 392 bpm, which is 7 times the metric tempo of 56 bpm at this point: we can see dark bands on both tempograms in this region.

5.12 Example 7: Free laykari in complex subdivision (4:3)

Singers also sometimes use more complex lay ratios, where rather than just subdividing each matra, they appear to subdivide groups of matras. The simplest example of this is dividing two matras into three pulses (3:2), known as derhi or derhgun, as mentioned above. This example from the Gundecha Brothers' Rag Ahir Bhairav illustrates another possibility, a 4:3 ratio. In fact, this passage belongs in the binary category: the 4:3 effect is achieved by dividing the matra pulse into four and then articulating each third pulse.

This can be seen in Figure 13.7. In the vocal track there are 16 main syllable onsets in the 12-matra cycle, as seen by the automatically detected syllable onset peaks. Against a basic tempo of (again) about 58 bpm, the main periodicities are around 40 and 80 bpm, which is consistent with the 4:3 subdivision. In contrast, in the pakhavaj part the main subdivisions are the half-matras and the tempogram shows periodicities at 60 and 120 bpm. Although both vocal and drum parts are consistent with an underlying 4:1 subdivision, the vocal part emphasises every third subdivision and the pakhavaj part (especially in the last four matras, where he plays 'tita kata gadi gana') every second subdivision. We can also see what appears to be a slight imprecision in onsets falling slightly ahead of the matra. We should not read too much into this: in this case the sam of cycle 52 is manually annotated a little after the drum stroke (about 65 ms) and this affects the placing of the reference lines.

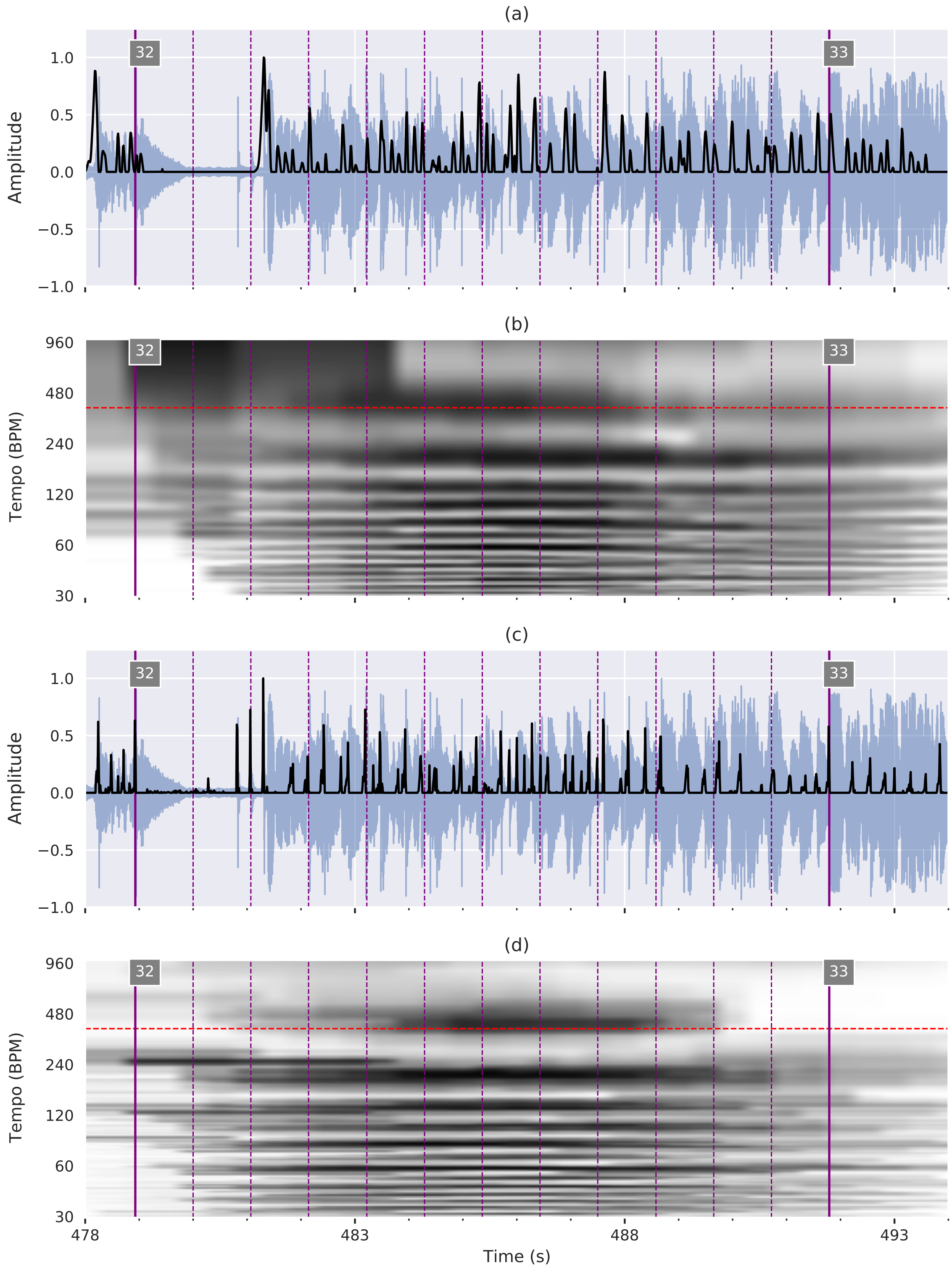


Figure 5.7: Prashant & Nishant Mallick, Rag Bihag, cycle 32. From the top: (a) waveform with the output of a vocal syllable onset detector in black, (b) vocal tempogram with reference line at 392 bpm (= 7:1 ratio), (c) waveform with drum onset detection function overlaid in black, and (d) pakhavaj tempogram.

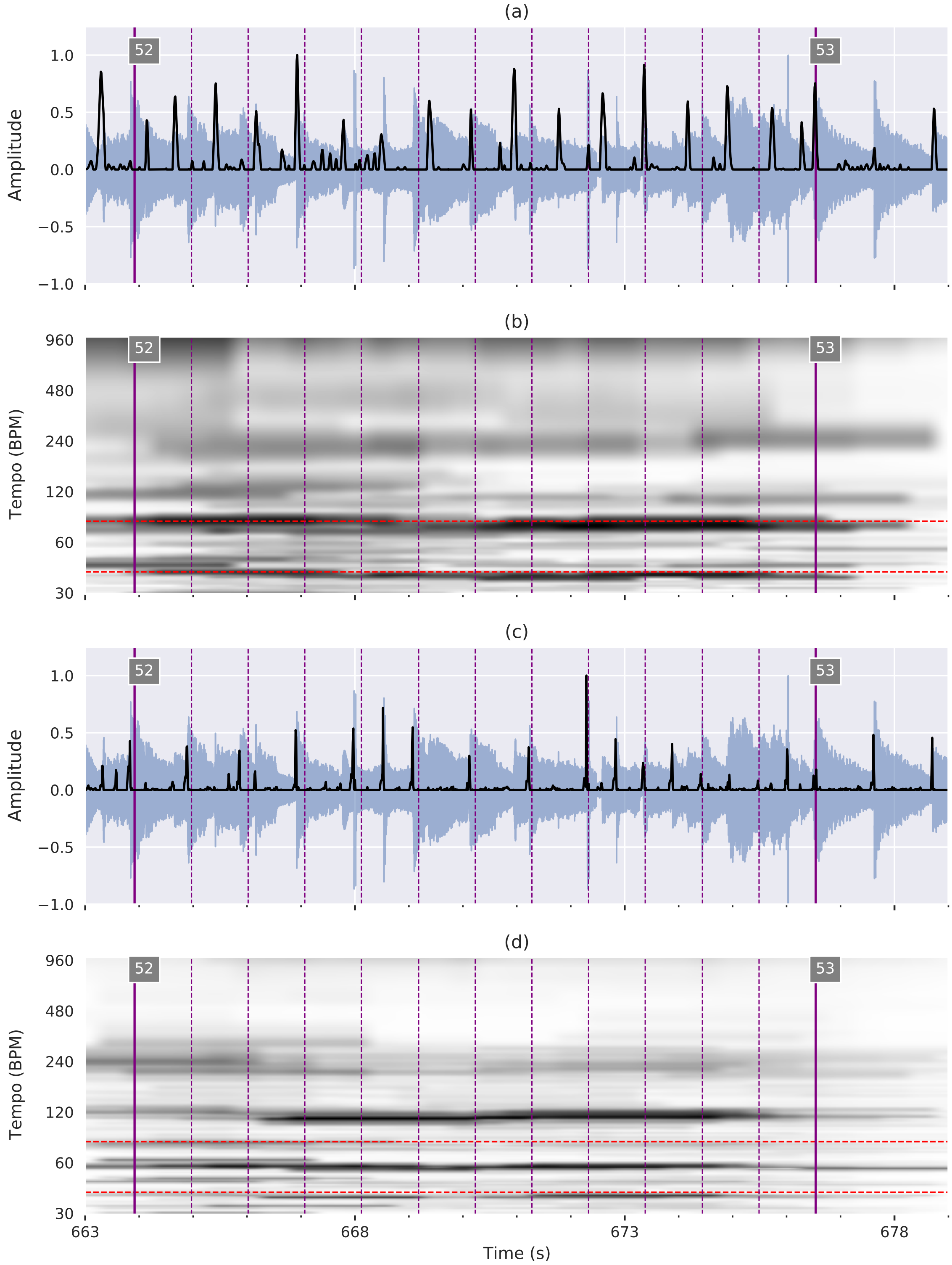


Figure 5.8: Gundecha Brothers, Rag Ahir Bhairav, cycle 52. From the top: (a) waveform with the output of a vocal syllable onset detector in black, (b) vocal tempogram, (c) waveform with drum onset detection function overlaid in black, and (d) pakhavaj tempogram. Dashed horizontal lines in the tempogram plots mark 40 and 80 BPM.

5.13 Example 8: A complete performance

Each of the preceding examples has illustrated a particular stage of a typical dhrupad performance, be it bandish presentation, free or fixed laykari, and illustrated how these musical features can be observed in tempograms, onset detection functions, spectrograms and pitch tracks. The last example pulls out to the level of a complete performance, and demonstrates how these techniques can be used to parse the performance into its constituent sections or phases. This approach is a work in progress, hence we discuss the strengths and weaknesses of our approach to date. Figure 13.8 illustrates the overall structure of a dhrupad bandish performance, the Gundecha Brothers' Rag Ahir Bhairav, illustrating the increase in tempo, division into sections and episodes, and the different lay ratios of vocalist and pakhavaj. Tal cycle boundaries are not included for readability; rather, manual annotations of episode and section boundaries are overlaid on vocal and pakhavaj tempograms.

The 'tempogram' representations highlight periodicities in the two parts, and can be related to the lay ratio estimations in the bottom panel. In the initial section the singers perform the bandish while the pakhavaj solos; a similar pattern is repeated in cycles 39-43 (roughly 500-560 secs): this can be seen in the relatively low density in the vocal tempogram and high density for the pakhavaj tempogram for these sections. The vocal lay ratio estimation (second panel from the bottom, red line) correctly plots the gradual increase in density from cycle 1 through to 38, then a drop back when the bandish is repeated and a further increase up to cycle 57. However, the automatic estimates of 8:1 are not accurate: the first occasion (just before 500 secs) may be due to bleed from the drum part, which is playing at 8:1; in the second '8:1', around 720 secs, they actually sing at 3:1 with a gamak (shaking oscillation in the voice) which means it could be perceived as 6:1; however, as mentioned earlier, the system has not yet been trained on ternary subdivisions, and so is unable to pick up the 3:1 here or in the fixed laykari section (cycles 64-5, roughly 808-833 secs).

In the pakhavaj tempogram we see the gradual intensification right up to 16:1 in the initial bandish presentation, then a more gradual speeding up to match the singers' free laykari. Then again, we see the fast solo passage, and thereafter it seems to jump about as the drummer adjusts his playing to match the vocal part.

5.14 Summary

An expert listener can parse a dhrupad bandish performance in different ways. The most unambiguous is the tala structure: the tala is in principle never ambiguous, and generally the only reason it may be so is if one of the performers makes an error. The sung content can be divided into the bandish itself, lay bant ('fixed laykari', optional) and bol bant ('free laykari'). The bandish will typically use a maximum of one syllable per matra; the steady rhythm of the singing may be contrasted with much faster and more intricate pakhawaj patterns. Lay bant will be characterised by a clear step-wise increase in the speed of the vocal line. Bol bant material is more diverse but will be characterised by a gradual increase in speed and rhythmic interest: basic subdivisions of the matra would be the binary 2:1, and 4:1 but other ratios, especially ternary, are heard more rarely. Bol bant sections tend to progress as a series of episodes whose end is marked by short closing figures and a reiteration of part of the bandish (usually the sthayi). Pakhawaj accompaniment uses either the theka (with a basic speed of one stroke per matra, although a couple of matras may use a 2:1 subdivision and the last couple of matras may be more dense than the rest); solo sections, typically 8:1 or faster; and 'sath sangat' in which the drummer matches and responds to the vocalist's patterns.

It should be clear from the above that some – but by no means all – of the most important factors according to which we might construct an analytical description of a bandish performance are essentially rhythmic. Identifying the tala cycle boundaries is essential – something we have done manually here but could in future be done computationally – also because in this genre, laykari episodes tend to take up a multiple of the tala cycle, and section boundaries are thus a subset of tala cycle boundaries. A key feature that distinguishes sections and allows us to track what the musicians are doing is the subdivision of the matra pulse, which we term lay ratio. Identification of this ratio separately in the two main parts – voice and drum – should be an important

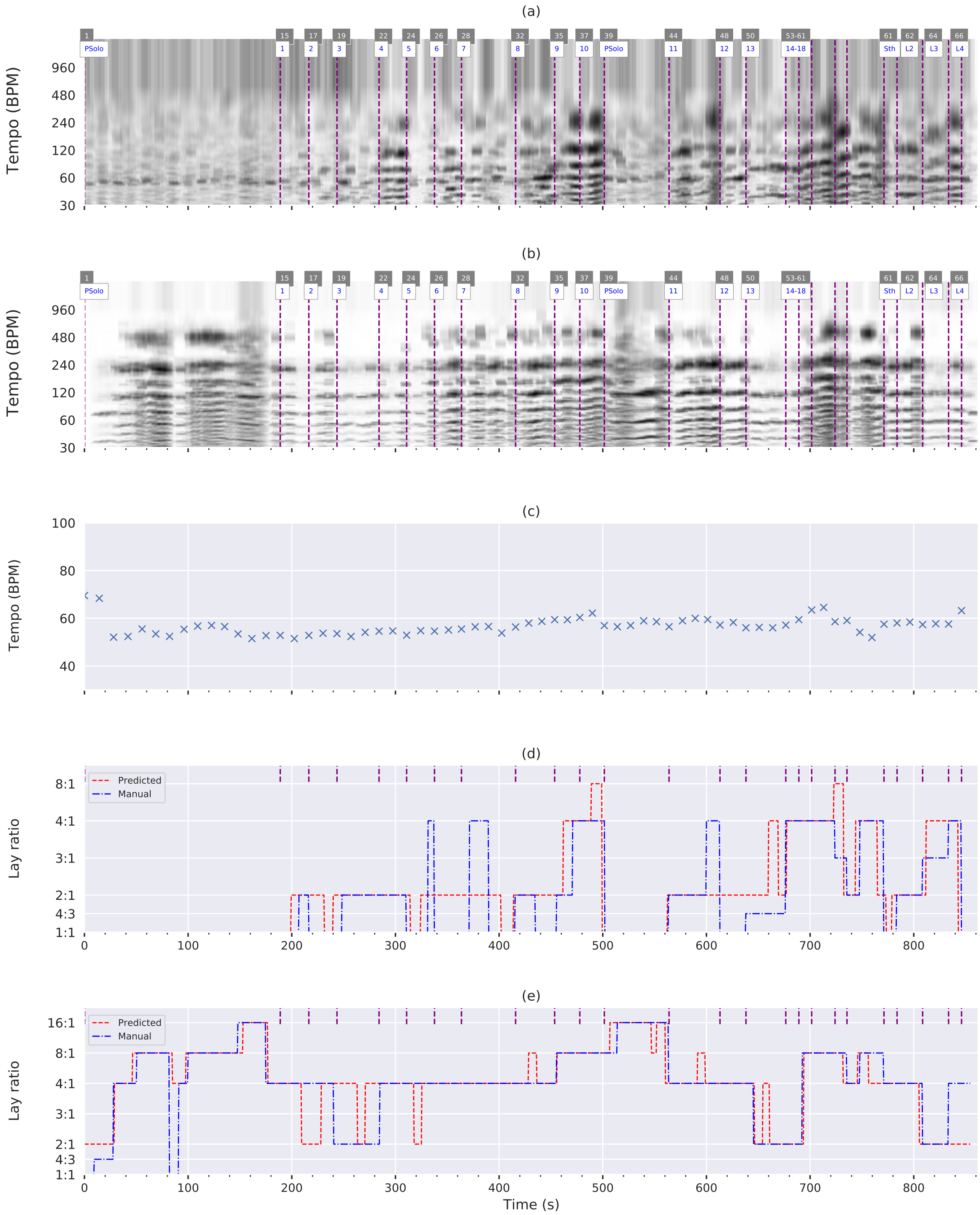


Figure 5.9: Gundecha Brothers, Rag Ahir Bhairav, complete bandish performance. From the top: (a) vocal tempogram with manually annotated structural boundaries overlaid (Numbers in grey boxes are cycle numbers corresponding to episode boundaries; numbers in blue font in white boxes indicate vocal episode numbers; PSolo is Pakhavaj Solo; Sth is sthayi; L2, L3, L4 are episodes of fixed laykari (see example 3) in double, triple and quadruple speeds); (b) pakhavaj tempogram with manually annotated structural boundaries overlaid; (c) computed metric tempo; (d) computed and manual lay ratios for vocals and (e) for pakhawaj. The metric tempo is computed on mixture audio (Vinutha et al 2016[Vin+16]) and lay ratios from CNN (Rohit et al 2020[RVR20]) of vocal and pakhawaj from source separated audios.

component of any computational approach to parsing dhrupad bandish performances. Other features would depend on approaches such as timbre, pitch or text recognition. This chapter does not set out to cover all such possible approaches, although we did show in Example 3 how pitch information may be useful in identifying passages of fixed laykari. What we do aim to do is to implement and test approaches including source separation and lay ratio estimation, and evaluate their success in the structural description of performances.

The musicological benefit of such a system would be that manual annotation and description of performances is time-consuming, and this makes corpus-level analysis prohibitively costly in terms of expert time. In order to study differences between individual or group styles, or the historical development of performance practice, the work of musical analysis could potentially be supported by computational means: a system capable of breaking a performance down into sections and episodes and characterising different parts as bandish, lay bant or bol bant; and of tracking the relative speed of those sections and episodes, could potentially enable such studies and help musicologists to direct their attention where it is needed.

References

[Cla08] Clayton, M. *Time in Indian music: Rhythm, metre, and form in North Indian rāg performance*. Oxford University Press on Demand, 2008.

[Hen+20] Hennequin, R. et al. "Spleeter: a fast and efficient music source separation tool with pre-trained models". In: *Journal of Open Source Software* vol. 5, no. 50 (2020), p. 2154.

[RVR20] Rohit, M. A., Vinutha, T. P., and Rao, P. "Structural segmentation of dhrupad vocal bandish audio based on tempo". In: *Proc. of int. soc. for music information retrieval conf.(ismir).(Montreal, Canada)*. 2020.

[SW04] Sanyal, R. and Widdess, R. *Dhrupad. Tradition and Performance in Indian Music.* Ashgate, 2004.

[Thi97] Thielemann, S. *The Darbhanga Tradition: Dhrupada in the School of Pandit Vidur Mallik*. Indica Books, 1997.

[Vin+16] Vinutha, T. P. et al. "Structural Segmentation and Visualization of Sitar and Sarod Concert Audio." In: *ISMIR*. 2016, pp. 232–238.

Chapter 6

Music Information Retrieval in Carnatic music

Ranjani H. G.[1] **and Hema A Murthy** [2]

[1] Indian Institute of Science (IISc), Bangalore, [2] Indian Institute of Technology Madras (IITM)

6.1 Introduction

Music Information Retrieval (MIR) is an interdisciplinary research area in music technology which aims to analyze, understand, represent music data and finally model abstractions of music from the acoustic signals that are sensed. Music signals, within a cultural framework, invoke perceptual and cognitive abstractions at various levels, from acoustic features (such as pitch, loudness, rhythm) of constituent instruments to lyrics, music composition and rendering styles, genre. In this Chapter we document the literature on computational analysis of Carnatic Music, an Indian art music form that comes from Southern India. Indian music is primarily an oral tradition that has been passed on from the teacher to the taught. Further, Indian art music also lends itself to significant extempore exploration which leads to improvisation. This improvisation enables artistes to present their view of a melody/composition. This is referred to as *manodharma*. Owing to significant differences in presentation across artistes, processing Indian art music is a challenge. In recent times, there has been an impetus to processing Indian art music. The primary objective of analysis is to enable music retrieval, rich transcription and generation. Musicological analyses of Carnatic music can be dated back to (Chaturdandi prakasa) While there have been sporadic efforts to analyse Carnatic music from a computational perspective, the attention towards systematic computational approaches is quite recent **??**.MIR research in the past two decades has resulted in progress in representation, understanding, retrieving, organizing and interlinking large music content. An interest towards developing and promoting cultural specific perspectives in MIR, in the past decade, has led to identifying musically relevant problems in cultural contexts and developing suitable generic or specific approaches towards these.

This chapter provides an overview of recent work on MIR with the focus on Carnatic Music (CM), referred to henceforth as Carnatic Music Information Retrieval (C-MIR).

The Chapter is organised as follows. Section 6.2 gives a brief overview of Carnatic Music concert. Section **??** identifies the various challenges in MIR for Carnatic Music (C-MIR). Section **??** reviews the various efforts in data driven approaches to MIR, and finally Section **??** attempts to suggest new MIR problems that can be addressed. **??** is a good starting point for young researchers in the field of Carnatic Music.

6.2 Carnatic Concert

In India, Carnatic music is more prevalent and performed in the southern regions of India. Carnatic music is performances consist of a concert that lasts for about 3 hours. The items that make up a concert are seldom pre-planned. The music is based on "gayaki" style or vocal style. A conventional Carnatic Music concert has a lead performer and accompanying artists. Lead vocalists are predominant. While concerts can have lead instrumentalists, the instrumentalists also mimic the oral style of performance. Each concerts consists of a

Figure 6.1: Carnatic katcheri with Vid. MSS as lead performer.

variety of different types of items performed in different *rāgas*. There are about 200 *rāga*s that are regularly presented in concerts. Every concert consists of a main item that is performed in a *rāga*s that lends itself to significant improvisation. Such *rāga*s are small in number. The items themselves come from many different forms, namely varnam, swarajati, kriti, padam, javali, paasuram, thevaaram, viruttam, thillana, sloka. A *kriti* is considered an important and serious item in a concert. The lead artist(s) chooses compositions (typically to suit the occasion) which showcase their proficiency (eg, variety of *rāga* and *tāla* with adherence to tempo) while also simultaneously catering to the needs and/or requests of the audience based on their assessment of the audience's taste and expertise. The audience too responds to the artists through tapping of the tala, clapping to acknowledge the creativity and/or performance as well as requesting for a few compositions. The number of expositions can range between 5 - 10 along with sufficient time for improvisations and showcasing of proficiency of accompanying artists. With respect to rhythm cycle, everyone keeps their own count of beats. This gives freedom to a rhythm accompanist to improvise as well. The musical ideas of the performer are presented in both compositional and extempore formats, in various ragas and talas. Generally, the performers avoid repeating the raga in order to institute variety in the concert as well as to be able to demonstrate the proficiency across ragas, forms, compositions as well as improvisation well within the framework of the music.

Figure 6.1. shows a typical Carnatic music katcheri. A kutcheri typically has at least 2 performers - lead performer, percussionist. The percussion accompaniment is provided with mridanga, ghata, kanjira, thavil or morsing. The Sruti (or drone) is provided either by an accompanying performer, sruti box or electronic tambura. The lead performer can be either a vocalist or instrumentalist. Typical instruments used for melody are vina, violin, flute, nadaswara, mandolin, saxophone, guitar, jala-tarang. Off-late whistling is also used in performances. Whatever be the instrument/vocal in which the lead performer performs, presentation basically mimic the oral performance. The prowess of an the instrumentalist is evaluated by how well s/he is able to even mimic lyrics.

A typical CM concert begins with a varna or short composition rendered without much fanfare. Longer and main compositions follow the varna. These mainly consist of sub-sections alapana (extempore exposition of a raga with vocal syllables such as ta, da, na, ri, without explicit rhythm accompaniment), tanam (also extempore exposition which blends melody with medium-paced rhythm to exhibit the intricacies of the raga using syllables such as a, nam, tam, na, tom, ta, nom, such that the audience perceive the words anantham - endless and anandam - happiness in a continuum), followed by the composition and swarakalpana (explicit vocalisation of swaras). The accompanying melodic instrumentalist also showcase their proficiency during the alapana, and the tanam. The percussion performance known as tani-avarthana follows the main composition. With multiple percussionists, sequential alternate performances is the convention, concluding with all of them converging to the same rhythm pattern. The final parts of the concerts consist of audience requests, followed by a thillana (beat synchronous melodic utterances of rhythmic syllables such as dem, ta, tha, na, dir, ki jhum, and the likes within the raga framework) and/or shorter compositions. The conclusion is with a mangala - a thankful prayer.

The current popular compositions have a rich history while also including the recent popular compositions. They range from those composed in 12th century (eg., Basavanna, Akka Mahadevi, etc), through 15th century (eg., Purandaradasa, Arunagirinathar, Kanakadasa,) to those in the late 18th and early 19th century (eg., Thyagaraja, Muttuswami Dikshitar, Syama Sastri, Muttiah Bhagavathar, Patnam Subramanya Iyer, Swati Tirunaal). There are popular compositions from the 19th century (eg., Subramanya Bharatiyar, Jayachamarajendra Wodeyar, Mysore Vasudevachar, Papanasam Sivan) and 20th century as well (eg., M. D. Ramanathan, M. Balamuralikrishna). The adherence to the original compositions is not very clear as the tradition followed guru-shishya parampara and relied on aural and oral accounts. The composer is referred to as vaggeyakara. Vaggeyakara is etymologically a kaara (one who does implying composes), vak (lyrics) and geya (sings) and is therefore the person who composes lyrics and sets it to the raga and tala - implying profound proficiency in both language, prosody, melody and rhythm aspects so as to retain the bhava of the raga and the lyrics. Most Carnatic compositions are in Sanskrit, Kannada, Tamil, Telugu, Malayalam or Marathi languages. Performing artists can choose compositions composed by these vaggeyakaras for the Kutcheris. In any Carnatic kutcheri, the performing artists have the opportunity to showcase their creativity through extempore and improvisational aspects within the framework of the composition while retaining the bhava of the raga and lyrics.

6.2.1 Wishlist for C-MIR

Pedagogy and applications are the drivers for music information retrieval. Indian art music is primarily an oral tradition where the music has been handed down via the *gurukula* system, which leads to significant variations in renditions of compositions and rāgas, and is continuously evolving. While many of the MIR tasks of Western music are relevant for Indian art music too, Indian art music has many challenges owing to its continuous evolution. Unlike Western music which is polyphonic, an ensemble where the conductor plays an important role, Indian music has a main performer, generally a vocalist/instrumentalist that conducts the proceedings. Most often the musicians meet on the stage for the first time, but nevertheless, perform very well together. Even the structure of a concert is hardly decided apriori. In Carnatic music, the concert is like a story that evolves over a three hour period. Carnatic music is very composition oriented. The prowess of a performer is established by the way s/he renders the composition bringing out the beauty of the lyrics, while adhering to the grammar of the *rāga*. Even the rhythm for the same composition performed across concerts by the same artist need not be the same. Despite the freedom that is allowed in this form of music, there are structures that can be quantified. Today there are some websites that archive Carnatic music, but unfortunately, they hardly give details. The first step in C-MIR is to identify the various tasks that can be useful for the listener, musician, learner, and researcher. Broadly the tasks can be classified under the following heads:

1. Melodic analysis: There are about 200 melodies that are actively performed in concerts today, each with its own structure. A number of songs are composed in each of these melodies. Understanding the phraseology, and nuances of melodies referred to as *rāgalakshana* is an important aspect. Inorder

to perform melodic analyses the tonic has to be identified, as any composition can be rendered in any key. Generally an artist fixes the tonic for an entire concert, and all items are performed at the specified tonic. Melodies with identical *ārohana* and *avarohana* can also exhibit significantly different phraseology. Transcription is thus a difficult task, in that even arriving at a universal transcription approach is challenging. The different items that make up a concert can themselves belong to different forms. Each type composition has its own structure, and segmenting these structures is yet another task. From a listener's perspective identifying the name of the *rāga*(s) that make up an item is important. Different schools render the same melody differently, and quantification of these variations can help a learner.

2. Rhythm analysis: There are about 175 different rhythm patterns in Carnatic music. Identifying the rhythm pattern associated with a composition, the downbeat, the segmentation of a composition based on rhythm are tasks that are useful. Similarly, there are different types of rhythm instruments, and melodic patterns are associated with these instruments. Identifying rhythm sequences is another important task in MIR. Here again, the language of rhythm is ill-defined. The different levels of rhythm seen in CM include akshara period, laya, sama, anga, edupuare some concepts that are unique to Carnatic music.

3. Structure analysis: Items in Carnatic music can belong different types, namely, kriti, varnam, viruttametc. Determining the identity of each item is indeed an important task. A CM concert has about 10-12 items. Segmenting the concert into items is useful from an MIR perspective. An item like kritican be further divided into its subparts. Given the improvisational nature of CM each of these subparts can also provide additional meta-data which can be tagged with the item.

On a similar note, consider the classification/verification task. While in the context of MIR in the context of Western music, cover song detection is an important problem, in Carnatic music the compositions are composed in different melodies by a number of composers. While the objective in Western Music is to find melodically similar songs, the objective in C-MIR, reproduction of the same composition with melodic improvisation is key. Unlike Western Music, the key for any song is defined by performer. The task can get more involved with increasing the depth of the problem statement such as discovering phrases and other structures of a raga, differences and commonalities between ragas, identifying the music composers through current renditions or phraseology and different schools of learning.

The chapter is intended to cover many aspects of C-MIR. This introduction section has provided a brief background of a Carnatic concert, followed by a wishlist of C-MIR. The subsequent sections cover melodic aspects (6.3), rhythmic aspects (Section 7.3) and timbral aspects of C-MIR (Section 7.4). A context is provided to every section by providing an introduction to the music concept and the research problem being addressed, the potential MIR applications, and the challenges involved in the MIR task. The approaches are then detailed followed by a summary of the approaches.

6.3 Melodic Analysis

6.3.1 Context

Melody, in the context of C-MIR, can be defined as [Gul17]:

> *an auditory object that emerges from a continuous series of transformations along six dimensions: pitch, tempo, timbre, loudness, spatial location, and reverberant environment, by the dominant individual melodic source in a music ensemble.* Details of different aspects of *rāga* in Indian art music can be found in 9.1 and 2.3.3.

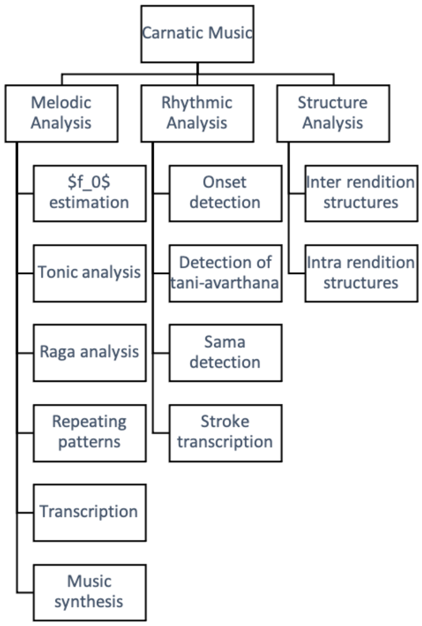


Figure 6.2: Overview of the chapter

Table 6.1: Summary of C-MIR melodic analysis

Author/Report	f0 estimation/ predominant melody estimation	Tonic identification	Analysis of melodic patterns	Raga analysis	Melodic transcription	Music generation and synthesis
M Subramanium, [MSu09; Sub02; Sub12; Sub99]	✓			✓	✓	✓
A Krishnaswamy, [Kri03a; Kri03b; Kri04a; Kri04b; Kri04c]	✓			✓	✓	✓
Srikumar, PhD thesis				✓	✓	✓
S Gulati, Masters thesis	✓	✓				
A Bellur, Masters thesis	✓	✓				
V Ishwar, Masters thesis	✓					
S Dutta, Masters thesis			✓	✓		✓
R Rajan, PhD thesis	✓					
S Gulati, PhD thesis		✓	✓	✓		✓
G K Koduri, PhD thesis			✓	✓	✓	
Padmasundari, Masters thesis			✓			
Ranjani, PhD thesis		✓	✓	✓	✓	✓
Venkat, PhD thesis				✓	✓	✓

6.3.2 Overview of melodic analysis

Melodic analysis in CM ranges from (but not limited to) estimating the pitch as a low-level feature from a recording or performance to mapping them onto higher level aspects such as shadja, micro-tonal structures, gamakas, motifs, raga. Table 1 summarizes the different sub-problems addressed. Most of the works are collaborative efforts. Some of the works are detailed in subsequent subsections. A summary of the datasets used/available for each of the sub-problems is detailed in Table 2.

It is worth highlighting that music recordings are protected by Intellectual Property Rights. The above datasets are created only for research purposes and not for distribution and/or broadcasting.

Table 6.2: Summary of datasets for C-MIR melodic analysis tasks

Dataset	C-MIR Tasks	Availability
MIREX datasets	f0 estimation/ predominant melody identification	Publicly available
Compmusic Tonic dataset	Tonic frequency estimation	Available here https://compmusic.upf.edu/iam-tonic-dataset
Compmusic Carnatic dataset	Tonic frequency estimation, raga\analysis	Available here https://compmusic.upf.edu/node/328
Charsur dataset	Tonic frequency estimation, raga\analysis	Available on registration at https://www.charsur.com/
Sangitapriya dataset	Raga\analysis, Knowledge representation	Available on registration at http://www.sangeethapriya.org
Varnam dataset	Transcription, raga\analysis, music synthesis	Available here http://compmusic.upf.edu/carnatic-varnam-dataset
IISc tonic estimation database	Tonic frequency estimation	Available on request to authors of [RAS11]
Shivkumar online audio and prescriptive transcript dataset	Raga\analysis, transcription, music synthesis	Publicly available
IITM concert database	Tonic frequency estimation, motif analysis, Raga\analysis	Available on request to authors of [**Ashwin_Istanbul2012**]
IISc transcription database	Transcription	Available on request to authors of [RPS17a]
IITM multi-track concert database	Tonic frequency estimation, motif analysis, Raga\analysis	Available on request to authors of [San17]

6.3.3 f_0 contour estimation

6.3.3.1 f_0 contour as melodic representation

The melody of a CM rendition can be characterized in terms of gamakas characteristic phrases, grammar, raga chosen, improvisations within a rendition and teacher-student lineage. It is imperative that any representation of a melodic piece must capture these aspects of the rendition; such a complete representation is essential for a listener to (i) reproduce the melody either as a hum/whistle (ii) capture the essence of music [Pol+07]. Given the heterophonic nature of CM and the MIR tasks to analyse the melodic components of music from audio recording, it is possible to capture almost all of the above mentioned aspects of melody through a sequence of finely sampled pitch estimates. A reliable, robust estimate of predominant pitch or multi-pitch is necessary for the melodic analysis. Although pitch refers to perception of fundamental frequency (f_0), for all practical purposes in C-MIR, f_0 is considered to be an objective estimate of pitch and hence, when we refer to pitch estimation, we indeed refer to estimating f_0.

Consider a sample melodic clip[1] of CM. This audio has been synthesized using only predominant pitch extracted from an actual recording, along with the associated corresponding harmonics and timbre (the number of harmonics and timbre kept constant). It is evident from listening that the clip has the melodic component of CM. A sample of the spectrogram is shown below (as captured using Audacity tool) with the red pointer indicating towards the f_0 of the lead melody. The other contours above it correspond to the harmonics of f_0. The blue pointer indicates the f_0 of the drone in the background. The f_0 values (y-axis) correspond to pitch variations and hence swara-marks, and movement of f_0 over time results in the melodic contour. Although timbre is important for perception, for computational purposes in the works discussed in the Chapter, we use the predominant pitch is considered as a representation of a raga.

[1] https://tinyurl.com/ya5tnv8d

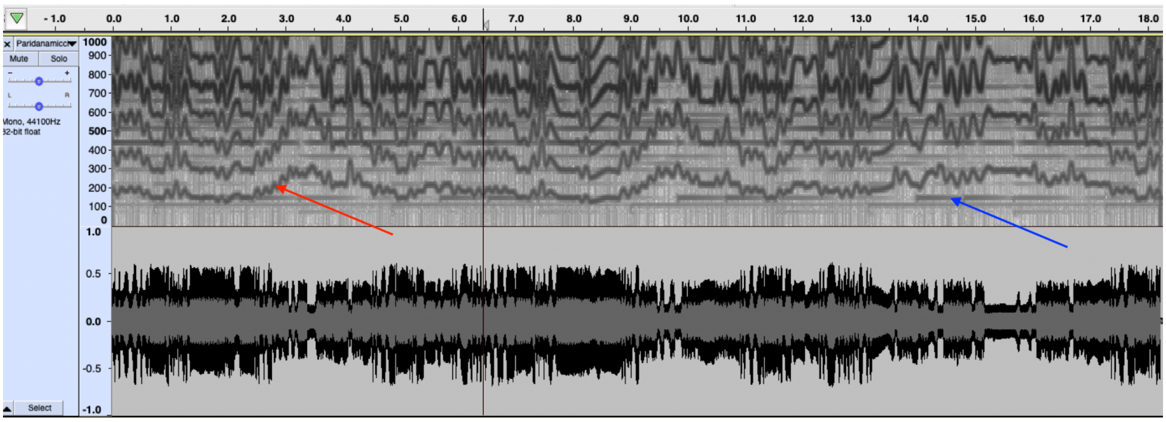


Figure 6.3: Spectrogram of an audio clip - https://tinyurl.com/ya5tnv8d. Red pointer indicating towards the f_0 of the lead melody, while the blue indicates towards the f_0 of the background drone.

6.3.3.2 Challenges

The heterophonic nature of CM leads to varying a single melody through multiple instruments. The lead and the accompanying instruments play the same melody but can be tuned to different octaves depending on the type of accompanying instruments, and the tonic of the lead performer. Thus harmonic frequencies of lead and accompaniments can overlap or be close, requiring well resolved f_0 estimates. Also, the accompanying percussion instruments of CM display harmonic nature and are tuned to the tonic frequency of the lead performer; this further poses challenges for f_0 estimation.

6.3.3.3 Approaches

There exists a considerable number of algorithms for estimating f_0 starting from [Boe93], [De 02] (which are generic approaches for any audio signal) to [AB15; SG12a; ZCY20] (which addresses melody in polyphonic audio signals in particular).

If the music considered is dominated by a single instrument/vocal lead (monophonic), then the fundamental frequency and its harmonics are smooth over time and can lead to efficient estimation using autocorrelation based pitch estimation approaches such as [Boe93], [De 02]. [Boe93] makes use of the autocorrelation of the windowed signal, coupled with cost functions to reduce octave jumps and voiced/unvoiced transitions. In [De 02], an average magnitude difference function (AMDF) approach is considered in order to reduce the estimation errors; further fine-tuning is achieved through error reduction approaches such as integration and thresholding followed by best local estimate. Alternatively, modified group delay based (MGD) method is proposed in [RM13] - the MGD of the cepstrally smoothed signal produces peaks at multiples of the pitch period. Consistency across frames is used to ensure the consistency across frames in the pitch tracking. A performance comparison with YIN and ESPS techniques on MIREX data sets shows the competitiveness of the technique. However, ground truth is not available for Carnatic music, and hence evaluation is through perception of synthesized audio from pitch contours.

Although these do not address heterophony music, the possibility of estimating multiple pitch tracks is preferable while using concert recordings. A method proposed in [SG12a] picks out predominant voices that reflect the desired pitch even in the presence of other periodicities. Sinusoidal components are extracted using STFT and low frequency estimation errors are corrected by using instantaneous frequency (IF) and amplitude of the phase spectrum. The salience is computed as the sum of the weighted energies of harmonics of that frequency. Thus, the sinusoidal components along with their salience are grouped along time using heuristics of auditory streaming cues and constitute the multi-pitch estimates of the audio. [SG12a] has almost become the defacto method for estimation of pitch for heterophonic music.

The works of [Sub02] and [Kri03a] are few of the earliest to advocate use of pitch trackers for melodic analysis of CM. The pitch contours of certain ragas rendered by various artists are extensively analysed in terms of the theoretical note position vs those seen in data. In the work of [Kri03a], audio clips of several recordings were considered and 12 clearly distinguishable intervals were observed from melodic analysis. Further, variations in intonations and presence of inflexions in CM are also observed. The following work of [Kri03b] studies the pitch perception vis-a-vis the f_0 tracks obtained using STFT. In a subsequent work of [MSu09], the *gamaka* analysis of thodi raga is studied using pitch contours extracted from Praat software [Boe93]. Subsequent works on melodic analysis have used either mono-pitch estimates or multi-pitch estimates which pick out the predominant voices.

Extracting pitch corresponding to vocal melody has been proposed in [Ish14] using a machine learning based approach. Here, candidate pitch contours are extracted using predominant melody estimation [SG12a]. Timbral representative features are derived from the extracted harmonics of the candidate pitch contours and are used to classify the pitch contour candidates to vocal and non-vocal classes. The timbral representative features used were spectral centroid, energy, modulation energy ratio of various harmonics, their dynamic representation using mean, deviation and delta features along with pitch contour features such as length, mean pitch, salience.

6.3.3.4 Summary

f_0 estimates are an adequate representation of melodic contours; these can be estimated from audio recordings, which are extensively available for a wide variety of recordings in CM. These contours are observed to be information rich but do not map to the melodic components (such as swaras, gamakasand ragas) of the CM directly [KI12b] or even shadjaalso referred to as tonic. Tonic estimation is essential for any melodic analysis of Indian art Music. This is the focus of the following section.

6.3.4 Tonic Analysis

6.3.4.1 Tonic frequency

As mentioned in the introduction, unlike Western music, the key for a composition is chosen by the lead performer, unlike Western music where the key is defined by the composer. In a typical CM performance, the drone instrument such as tanpura, tamburi or *śruti*-box aids in providing tonic reference to the lead performer and other concert accompanists. The percussion instruments such as *mridanga*, *tabla*, *pakhawāj*, *khōl*, *ghatam* and *morsing* are also tuned to the tonic frequency. Based on the octave in which it is rendered, it is referred to as mandra shadja(tonic frequency of the lower octave), adhara shadja (tonic frequency of the middle octave) and tara shadja (tonic frequency of the upper octave). The chosen tonic frequency can vary even for a particular singer across renditions of the same composition (across concerts, with the exception of cases when composition is in madyama sruti or has sruti bedha). In the context of C-MIR, the term shadja refers to the tonic frequency of the middle octave i.e., adhara shadja unless otherwise mentioned. The tonic frequency is said to be achala (immovable in Sanskrit) and must be maintained steadfast throughout the performance. The positions of the rest of the swaras are derived in relation to shadja. The relations between tonic frequency and other swaras are as per the ratios in [Sam98a].

A listener makes use of a combination of the following cues to identify the tonic frequency:

- Melodic characteristics of the swaras
- Presence of drone in the background
- Knowledge of the raga

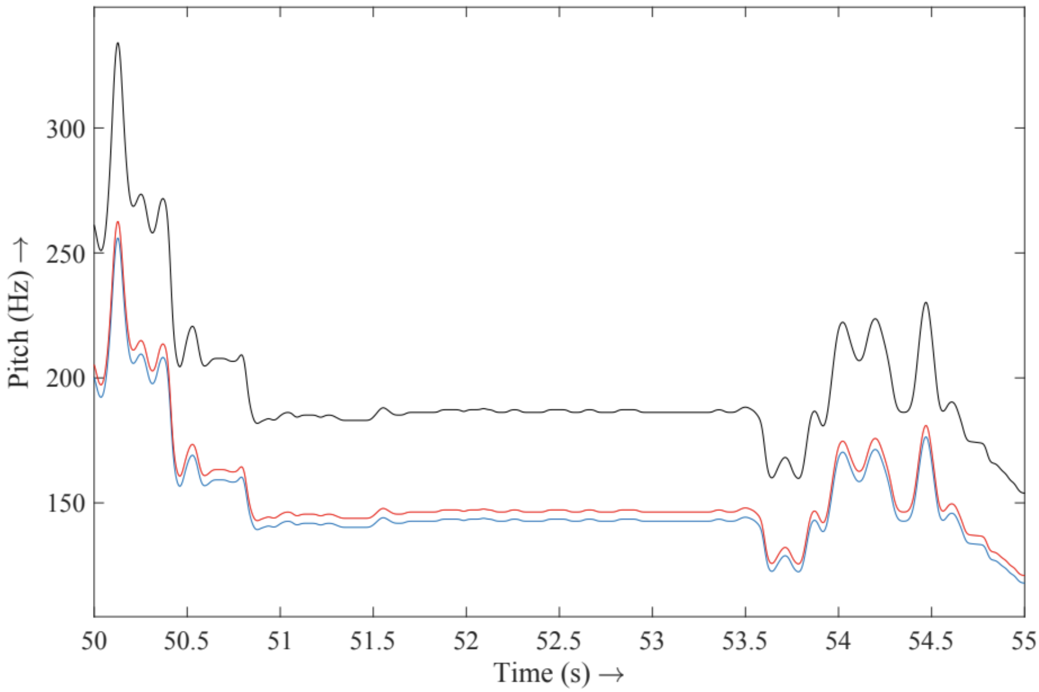


Figure 6.4: A sample visualization where a melodic motif rendered at different tonics will be identified as different without tonic normalization.

6.3.4.2 Challenges

Analysis of items in a concert requires the estimation of the tonic. The same melody can be performed at different tonics as shown in Figure 6.4 . Matching the melodies requires tonic normalization. In earlier works like [**Tansen**], the tonic was manually identified for analysis. Ideally tonic must be estimated for every rendition, especially for applications that require organisation of compositions based on raga. The tonic pitch of vocalists (male and female) spans more than one octave, and is typically between 100-260 Hz. Thus, the tonic frequency is dependent on the singer and instrument, and is performance specific. The most frequent pitch frequency also does not necessarily correspond to tonic frequency owing to raga specific saliency of the swaras in CM.

6.3.4.3 Approaches

There are various techniques proposed in the literature for tonic frequency estimation in CM [RAS11], [SGS12], [GSS12], [Bel+12], [BM13b], [BM13a]. The approaches can be broadly considered to contain 3 steps - pitch extraction, pitch distribution and candidate selection. The candidate selection step can be broadly categorized to be supervised or unsupervised. The unsupervised approaches mainly rely on knowledge based inputs in the problem formulation stage while supervised approaches intend to discover the “rules” through the ML approach.

In the work of [RAS11], the audio is down-mixed to a mono channel prior to obtaining the f_0 contours. The f_0 contour of the pre-dominant melody is estimated using the [Boe93] method. The proposed tonic frequency estimator, fixed time step of 10 ms and a frequency range of 80-800 Hz are considered. A Parzen window estimator with a Gaussian kernel is used to compute a pitch density function. A suitable choice of kernel function controls the smoothness of the estimated density and aids in a smoother peak picking process. J peaks of the density function are considered for candidate tonic frequencies and are chosen in the range suitable for the analysis. For every candidate, a semi-continuous GMM is fit such that the means correspond to the fixed frequency ratios defined in Table 1, while the weights and variances are estimated. The candidate which maximizes the ratio of weights to variances of candidate and its octaves along with its fifth is considered as the tonic frequency estimate. The method makes use of the CM domain knowledge that tonic frequency and its fifth (along with their octaves) do not deviate from the intended ratio.

The work of [SGS12], the predominant melody (f0) extraction algorithm of [SG12a] is considered for estimating f_0 from the audio signal. The multi-pitch estimates clearly capture the tonic frequency (along with Pa-Sa of lower octaves) of the drone instrument and the pitch trajectory of the voice. A histogram of peaks of the salience function is considered and top P peaks of the histogram are considered as candidate tonic frequencies. Each peak is ranked based on peak magnitude and is associated with a relative distance with the highest ranked peak. A decision tree based classification approach is considered, which learns the best set of rules for selecting the tonic, based on the above two features.

In the work of [GSS12], a two stage approach is considered. The first stage is akin to the work of [SGS12], however, the classification is considered for the tonic pitch class alone. The second stage considers identifying the octave of the tonic frequency. Histogram of f_0 from predominant melody is considered. A 25-component feature is extracted which corresponds to the 25 equally spaced bins of the histogram. The class-label for training at this stage corresponds to the presence of the correct octave of the candidate or otherwise.

The approaches in [Bel+12], the f_0 contour is extracted using Yin algorithm [De 02]. Histograms of f_0 estimates are post-processed using group delay functions to produce a group delay histogram. For identifying the tonic frequency from the candidates, three strategies are proposed - the concert technique, the template matching technique and segmented approach. For the concert technique, the GD histogram of each piece of the concert is aggregated. This is motivated by the fact that tonic frequency is the same throughout a concert while the ragas can change across individual pieces. In the template matching technique, the histogram is multiplied with a set of ratios of shadja and panchama and their octaves. This leads to the estimate of that candidate with the highest peak after the use of the template as a mask. With the segmented approach, piecewise histograms are considered in order to counter the issues such as absence of panchama in certain ragasor tonic frequency not corresponding to a dominant peak even in GD histogram.

The power of GD histogram is furthered in the work of [BM13b] where the GD histogram is detailed and applied for tonic frequency in a slightly adapted manner. The features now comprise histogram bins normalized by GD histogram. This in turn accentuates the potential candidates. The Sa-Pa-Sa template in the middle and lower octave is used. The candidate with the highest peak is estimated as the tonic frequency.

Cepstrum based pitch estimation technique is proposed in [BM13a]. It is argued that peaks in cepstrum due to tanpura/tamboora being of relatively low energy may not be captured in the heterophonic signal. Hence, the range for peak picking is reduced to the lower octave of the tonic. The desired tonic is then twice the estimated tonic frequency. Instead of processing all the frames, alternately, low energy frames are considered for pitch estimation, histogram and peak-picking. This assumes the presence of tanpura and has low volume. As a third approach, NMF is proposed as an alternative to histogram. The bag of low energy frames are decomposed to a single basis spectral vector which in turn must have captured the lower octave tonic frequency, and hence it is possible to obtain the peak of cepstral estimate as the tonic frequency.

6.3.4.4 Summary

Consolidating the approaches, the techniques estimate f_0 contours from the audio signal as the starting point, and progress towards the distribution of f_0. From here, the techniques can be seen to branch out in terms of utilizing the domain knowledge to create templates, handcrafting features, or selecting frames to inculcate the observations from the audio signal at hand. It is note-worthy that none of these techniques make any prior assumptions on the raga of the audio signal. This in turn makes the techniques valuable as a starting point for melodic analysis. The following datasets have been used in the literature for tonic estimation - CompMusic Tonic dataset, IITM concert dataset and IISc tonic estimation database.

Table 6.3: Summary of approaches proposed for tonic frequency estimation

Method	f_0 estimation	Features	Model
[RAS11]	Mono-pitch [Boe93]	Pitch histogram	SC-GMM
[SGS12]	Multi-pitch [SG12a]	Peak rank, relative peak distance	Decision tree
[GSS12]	Multi-pitch [SG12a]	Peak rank, relative peak distance,	2 stage decision tree
[**Ashwin_Istanbul2012**]	Mono-pitch [De 02]	Group delay histogram	Concert, template and segmented
[BM13a]	Cepstrum based pitch estimation [**rajanphdthesis**]	Histogram	Peak picking, NMF

A comprehensive comparison between a few of these methods is presented in [Gul+14a]. More details can be found in the corresponding works and in the chapters of theses arising out of these - [Gul12], [Bel14], [Gul17] and [Ran20].

6.3.5 Raga analysis

6.3.5.1 Raga

Raga forms the melodic framework of CM. Since there is no simple counterpart of raga in most cultures, for MIR purpose, we define it as the interplay of three layers of melodic structures: (i) swaras (notes and their positions) (ii) gamakas (iii) transitions between swaras. A raga has been interpreted to be more general than a specific melody but more specific than a scale and provides a grammar for the rendition of notes, melodic phrases and motifs.

Raga analysis has been the most popular yet challenging research problem in C-MIR owing to the ragabeing the main stay of interest in CM. The primary motivation for ragaanalysis stems from the melodic improvisation. The objective is understand and organise large music collections in terms of ragas, composers, and music lineage.

Raga analysis includes data-driven analysis of to characterise a raga; it is characterization of raga and is inclusive of (but not limited to) through use and dominance of notes, their sequences, arohanaavarohana, *gamaka* and intonation analysis, emotions invoked by a raga and scope for improvisations, perception studies and resolving raga into sub-components or symbolic notations through transcriptions. In this section, the scope of raga analysis includes all the above mentioned aspects of raga except transcription of melodic contour; we have isolated melodic transcriptions into a separate sub-section (c.f., sub-section Music transcription). Melodic transcription is a difficult aspect owing to the “gayaki” tradition of learning. Only skeletal notations are available for CM - there is no available information to retrieve the relevant gamakas for all the ragasṘecent work does show that it is indeed difficult to arrive at a score. This aspect is dealt with in the music transcription sub-section.

Within raga analysis, the classification, identification and/or verification of raga from melodic contours constitutes the problems that have been addressed. In turn, it is expected that this data driven analysis will result in a better understanding of the raga framework which can be correlated with musicological knowledge and in turn cater to technological and pedagogical advances in C-MIR.

6.3.5.2 Challenges

The challenge of raga analysis commences from the gap between the f_0 representation of melody and the musicological, abstract definition of a raga. Characterizing various aspects of ragas and mapping melodic contours to raga(s) requires understanding of ragas and estimating the same from f_0 contours. There are 2 ways of approaching the same:

- Observations derived data for a given raga and correlation of the same from a musicological perspective.
- Hypothesize certain constituents of raga based on musicological knowledge and use it for raga identification and perception

The scalability of musicological knowledge can be a limiting factor in such analysis due to two factors - there is no uniform representation of musicological knowledge. Gamakaswere first documented in Sangeetha Sampradaya Pradharshini by Dikshitar in [**dikshitar2004**]. T K Govinda Rao's book on different composers with compositions and notations has some skeletal representations. Nevertheless, a preliminary analysis showed that it is difficult to quantify gamakasusing the available notations. Many of the recent works have primarily focused on a data driven approach to ragaanalysis.

The work of [Vij07a] studies the analogies between languages and ragas. Just as languages are associated with sequences of "sounds" and have a corresponding grammar, ragas too can be seen to have sequences of "melodic events" and their corresponding grammar. Thus, in analogy with speech and language, the challenge of ARR (automatic raga recognition) is analogous to LID (language identification). The task of extracting "melodic events" and their grammar can be seen akin to the challenge of ASR (automatic speech recognition) task, but with unknown vocabulary. Hence, ARR can be approached in either a discriminative approach (identifying raga label without worrying about the constituting melodic events and the associated grammar) or through a generative approach - modeling a perceptually acceptable vocabulary and grammar from the data for ARR. The task of automatic raga recognition is well defined. Hence, to a naive reader, it may appear to be another machine learning problem and hence can be limited to the discriminative approach of ARR. The nuance of the generative approach appears from the fact that just a raga recognition/identification task is not complete by itself. It must also accommodate and/or learn and have the ability to explain the perceptual similarities and dissimilarities that are invoked in musicians/connoisseurs in various aspects (eg., modelling perceptually acceptable vocabulary, gamakas, grammar, perceptual similarities, allied ragas, raga-rasa interplay) as can be listed through musicology. A survey of all available formulations till 2011, the associated issues in the assorted formulations and the challenges that need to be addressed for raga recognition is summarized in [KMS11].

6.3.5.3 Approaches

The approaches can be broadly grouped as those addressing ARR tasks and those that analyse the raga itself through inflexions and micro tonalities of the signal The former approaches are more machine learning oriented. The latter leans towards understanding the structures of ragas gamakas across ragas from the pitch contour, contrasting with musicological knowledge and hence calls for strong signal processing foundations. These also form the precursor for transcription tasks.

6.3.5.4 *Gamaka*analysis

Gamakasare pitch movements and inflexions which can result in perception of microtonality in Carnatic music. An in-depth analysis and understanding of gamakas can lead towards computationally modeling and relating them to other musical entities. Such analysis also plays a role in synthesis of ragas [Sub99]. Hence, it is imperative to analyze gamakas in various contexts.

One of the earliest detailed studies on gamakas is reported in [Sub02]. Ragas such as Mayamalavagowla and Saveri by various artists are analyzed. The study graphs of pitch contours show that variations of frequencies towards the end of gamakas is higher within renditions as well as across renditions. It was also observed that vocal renditions overshoot the theoretical frequencies than violinists' performances. One of the inferences drawn from the study is that overall perception of note may not be based on just the note frequency but also on the duration of the same. I think Venkat's paper has some very good observations about this. This perception of gamakas being a 2-dimensional (pitch and time) phenomenon is further argued and reinforced in the work of [Kri04a]. This in turn is the foundation of subsequent work [Kri04b] on melodic atoms, a set of 2D melodic units, to transcribe Carnatic music. The work proposes segmenting pitch contours of Carnatic music into 2 dimensional melodic building blocks. However, there has been no computational approach proposed for the same. A similar argument with an accompanying computational approach can be seen in the work [RPS17a] and [Ran+19b], where the gamakas are hypothesized to be 2D points in the pitch-time plane. The same is verified through synthesis and perception experiments.

Gamakas have also been represented with a stage and dance model [SWM12]. It is observed through analysis of Sahana raga that the amplitudes of gamakas can be estimated by the pitch of the stage component given local melodic context. The work, though limited to one raga has successfully been able to model a few known gamakas in terms of stage-dance component sequences. The constant-pitch and transient model of [VAM17b] is inspired from this work. Here, instead of stage-dance components, the pitch contour is divided into constant-pitch (CP) and transients based on the steadiness of the note - as determined empirically by minimum duration and variation of pitch. In order to verify the model, 30 second snippets split into CP only and transient only components for perception tests. It is inferred that while it is possible to recognize the raga through transients only, the context provided by CP has a significant bearing.

The work of [RVV18] proposes to learn both the presence and the type of *gamaka*in a data-driven manner using annotated symbolic music. The data consists of 25 songs digitized from Sangita Sampradaya Pradarshini and scoped out to be a few - Kalyani and its derived (consisting of only of a subset of the notes of the parent) ragas. The gamakas(8 in number) are considered for the task. Note-based features (36 notes) within a context of 5 notes (difference in frequency of the middle note in a context of 5 notes), along with the duration of the notes are considered. A Random Forest classifier is considered and it is shown that frequency and duration of neighbouring notes contribute among important features. It has been observed through experiments that adding phonetic features (such as vowel, fricative, stop, nasal, etc) and structural features (such as pallavi, anupallavi, charana, etc) degrade performance.

6.3.5.5 ARR

In the work of [SA09], arohana and avarohana (ascending and descending) patterns of the ragas are considered. Features corresponding to the number of swaras their combinations, and linear and non-linear pairs in arohana and avarohana are considered for a NN model. However, it is not clear how swaras are inferred from the audio signal.

The work of [SG09], the tonic frequency (misrepresented as fundamental frequency) is identified at the singer level (not at rendition level). The swaras are inferred at a sub-segment level. Each sub-segment is considered at the level of thattu, Veechu or count within a tala. The work overlooks foundational aspects of C-MIR such as extraction of f_0, methodology for identification of tonic frequency, tonic normalization and hence the subsequent steps too.

A study and analysis of relation between ragas and rasas is reported in the work [KI10]. The goal of the study was to verify the claim if ragas indeed evoke the rasas specifically attributed to them. Perception experiments were conducted allowing listeners to rate a clip of raga in terms of clusters of emotions associated with the *rasa*. It is observed that an array of emotions can be invoked based on the notes stressed in the composition. Further, a preliminary analysis for raga recognition using pitch class distribution (PCD) features folded into one octave shows that there exists strong intra-raga consistency and inter-raga variance.

In the following we summarize the work that first perform tonic normalization before any further analysis of ragacharacteristics.

In the work of [RAS11], the tonic frequency is identified automatically and the notes are inferred based on the unigram, bigram or trigram of each pitch estimate. A template matching approach from the pitch frequency distributions is proposed for raga verification. The template matching approach involves checking for presence/absence of particular swaras. Here, the evaluation is limited to ragas having all 7 notes i.e., sampoorna ragas.

Template based approaches are studied in more detail in the work of [Kod+12b]. Local slope of the pitch contours is used to identify stable pitch segments. These stable pitch segments are quantized to the nearest swara using just intonation scale. Pitch class distributions (PCD) are constructed from the swara mapping. Templates based on theoretical aspects of music and templates inferred by average of tonic-normalized histograms are considered. Evaluation shows that PCD is an effective feature for raga identification. The work is based on stable pitch segments and thus gamakas have not been considered.

Peak parametric approaches are considered in detail in [KSS12a]. Histograms (of bin size of 1 cent) are convolved with a Gaussian kernel; peaks are identified using prominence and neighbourhood parameters are parameterized by location, mean, variance, skew and kurtosis. Various classifiers such as naive Bayes, K-NN, SVM, logistic regression and random forest are considered for the raga classification task. A detailed comparison of allied ragas (such as Bhairavi, Mukhari and Manji) show clear separation in terms of kurtosis and skewness features. This work is further extended in [Kod+14b] by considering context-based modulations of pitch instances. Pitch contours are smoothed using mean/median of a rolling window. This in turn is observed to give better classification of ragas. These intonation descriptors (mean, standard deviation, skewness of the pitch distributions and their temporal and frequency contexts) are used in combination with the relation tuples extracted from unstructured textual sources using open information extraction systems for creation of ontology and knowledge base of Carnatic music [Kod14].

In [DSM15], raga verification based on similarities between pitch contours of *pallavi* lines of a rendition are considered. Various matching techniques such as Dynamic time warping (DTW), Rough Longest Common Subsequence (RLCS), Longest Common Segment Set (LCSS) are considered for the task. Cohorts of ragas are defined as those ragas which have similar movements. The verification proposed is based on similarity of test signal to the claimed raga and its cohorts.

The approach in [Gul+16b] proposed raga recognition based on mined phrase patterns. This is achieved using DTW on pseudo-note transcriptions obtained from stable regions of the query. In the feature extraction stage, a term frequency - inverse document frequency (TF-IDF) matrix based on the frequency of phrases within a rendition is used to build a classification model for raga recognition.

In another work of the authors [Gul+16c], a time-delayed melodic surface is proposed. Here, delayed bigrams of finely quantized pitch samples form the feature vectors. Raga recognition is performed using simple classifiers such as K-nearest neighbor and different distance measures such as Euclidean, Kullback–Leibler (KL) divergence and Bhattacharya distance.

In [PM17], the authors use pitch vectors within a 4 second window and use Locality Sensitive Hashing (LSH) algorithm to obtain top-N matches for the queried raga. The LSH matches are compared against the pre-defined cohort raga list for verification. This matching technique is raga identification with the *pallavi* line and alapana segments forming bins of hash tables.This work clearly shows that *pallavi* lines and early portions of alapana contain significant raga information, whereas the same probably gets diluted in the later portions of the rendition.

The works in [RS17] utilize available prescriptive transcription instead of the pitch contours. Only seven basic notes are considered so as to include noise in prescriptive transcription. Note patterns are captured using stochastic language models such as N-grams, skip grams and multi-grams with varied temporal contexts. Maximum likelihood framework is used for raga classification.

Similar stochastic language models are considered with estimated descriptive transcription in [RPS19]. The descriptive transcription is obtained by quantizing the critical points to most likely swara position. Raga is modelled akin to topics using the PLSA model. Raga identification is also extended to ragamalika renditions to validate the model.

In the recent work of [MC19], the raga recognition task is attempted with a deep learning framework - in specific, LSTM models are proposed. Post tonic normalization step, a random sampling of the pitch sequences are considered (random is indicative of the start of the smaller subsequences) and the length considered is determined empirically. A single layer LSTM is followed by an attention layer and 2 densely connected layers, finally culminating in a softmax layer for ARR purpose and categorical cross entropy is the loss function considered. For the inference stage, audio is split into numerous subsequences and voting is used to determine the majority class.

6.3.5.6 Summary

In this section, raga analysis is constrained to the C-MIR task of automatic classification/identifying and verification of raga. Almost all works (except [KI10]) address the raga analysis from the context of pitch contour or its transcripts (manual or automatically derived).

The approaches can be seen under three broad themes where the primary focus is on feature engineering rather than classification.

1. Pitch histogram parametric based features such as peak location, mean, variance, skew and kurtosis
2. Automatically mined motifs as features such as phrases detected with DTW, RLCS, LCSS, LSH or LSTM
3. Transcription based where templates are created using various features such as transcription of regularly sampled pitch point/critical point or PCD

An interested reader can dig into the corresponding works and in the chapters of theses arising out of these - [Kod16], [Pad16], [Gul17] and [Ran20] for more details.

6.3.6 Identification of melodic patterns

6.3.6.1 Melodic patterns

Many a time, performers and connoisseurs view raga as a collection of melodic expressions or motifs. The view point arises due to the use of motifs to identify ragas within a very short interval of time - mostly within a few seconds of the rendition. Different ragas could share notes, phrases and renditions are said to be replete with phrases. A few phrases are said to be characteristic and signature phrases or motifs of a raga and are observed across renditions of a raga [BIM12]. While characteristic phrases are based out of musicology, and require hand-holding by musicians for identification of the same, the large repertoire of data encourages discovery of such repeated patterns.

6.3.6.2 Challenges

From C-MIR perspective, there are 2 kinds of melodic patterns

- Characteristic phrases or motifs which are rooted in musicology/music theory and are useful for musicological studies
- Patterns discovered/ derived from data analysis which are based on assumptions and are more from an engineering perspective of the music signal

Characteristic phrases or motifs are established in musicology. Such phrases need not be consistent across various schools of music and hence there is no well established consensus on the set of characteristic phrases. Again, the same is reflected in various phrases across musicological literature. Also, the assumption of repetitive nature may not hold for such musicological phrases. It is possible to address the C-MIR phrase identification problem by considering a limited set of phrases. However, such an approach is limited in terms of scalability across various ragas and invariably requires domain knowledge of musicians. A scalable approach for C-MIR would seem to be to automatically discover phrases which must then be validated for musicological relevance. Such an approach is scalable and can be computationally expensive with increase in data.

For a data driven analysis, it is assumed that certain melodic patterns are repetitive in nature. These can be considered to be representative as they occur more often than others. Repeating melodic patterns are of significance to C-MIR in multitude ways; summarization, indexing, faster retrieval and ARR are some of the applications.

In the C-MIR literature, both the approaches have been explored - identification of musicologically identified characteristic phrases as well as mining phrases from pitch contours as well as from available transcriptions.

6.3.6.3 Approaches

In this subsection, the approaches are grouped based on the data source considered for identifying repeated patterns - pitch contours and transcription based approaches. Hybrid techniques, which infer transcription from pitch contours and then identify repeated patterns, are also proposed.

- **Repeated patterns from pitch contours**: The initial focus of the effort was primarily on identification and discovery of motifs of CM. The first work in this direction [BIM12] approached motifs in CM by compiling a set of signature phrases for five ragasthat were manually marked by professional musicians. Then, HMMs are trained for each motif using 80% of the marked data and tested on the remaining 20%. Results show that the approach shows promise. The invariance of the states of HMM (likely depicting sequence of notes) within a motif is showcased in [Rao+14]. However, there exists a possibility of a significant amount of misclassification of short phrases. In a subsequent work [Ish+13], motif spotting is approached using the saddle points of the pitch contour. A rough match between the query and reference is obtained through 2 stages of rough longest common subsequence (RLCS) - In the first stage, the stationary points (referred to as saddle points in the paper) are roughly matched, and in the second stage the entire pitch contour is matched to remove false positives. To combat the false positives, a modification of the RLCS is proposed in [DM14a], where two different modifications were made a) choosing longer motifs, and algorithmically modifying the cost which is dependent upon the goodness of fit.

  Motif discovery is attempted by considering the *pallavi* (first line of a composition) in the work [DM14b]. *Pallavi* is more likely to contain the characteristic attribute of a raga. The typical motifs are extracted with a modification to the RLCS algorithm where the slope and deviation of the linear trend of the saddle points are utilized to compute the distance between two subsequences - this in turn reduces the number of false positives. Discovery of melodic patterns is proposed through intra and inter recording in [Gul+14b]. Seed patterns are discovered from each rendition of a raga composition. These in turn are used for inter rendition pattern matching. Melodic similarity between two subsequences is based on DTW distance. The patterns are rank ordered. The number of computations have been found to be the order of a few trillion. It is also observed in a subsequent work [Gul+14c] that separability between melodically similar and dissimilar subsequences is poor across recordings. This is hypothesized to be due to the possibility of phrases of two allied ragas which could be differentiated based on subtle melodic nuances. Musically meaningful patterns are extracted by network analysis in [Gul+16a]- nodes constitute discovered patterns and edges constitute distance between them. Statistically significant deviations are retained through disparity filtering and melodic similarity threshold. Non-overlapping community detection on the undirected network is proposed, and each community is ranked for each raga. For each raga, top 10 communities are selected and a listening test is conducted using representative subsequences of each community. It is observed that 91% of the phrases were considered to be raga motifs by at least 7 out of 10 musicians.

  In another approach, Locality Sensitive Hashing (LSH) is used to hash similar entries to a bin through random projections [PM17]. Pitch vectors of 4 sec duration along with raga label form the tuple and are indexed using LSH. For any test query, the raga labels and the ranks of top N retrieved candidates are used for identifying raga cohort. It is found that usage of initial portions of alapana along with *pallavi* lines of a composition at the LSH indexing step can result in better retrieval of raga as well as its cohort class.

Table 6.4: Summary of approaches for identifying melodic patterns

Literature	Motif spotting	Discovering patterns	ARR	Transcription	Approach
[**prao2014jnmr_motif**; BIM12]	✓				HMM models to spot manually identified motifs
[Ish+13]	✓				RLCS of saddle points to spot manually identified motifs
[DM14a]	✓				Modified RLCS of saddle points to spot identified motifs
[DM14b]		✓	✓		Motif discovery from *pallavi*\ lines using modified RLCS
[Gul+14b; Gul+14c; Gul+16a]		✓			Melodic similarities using DTW for mining similar patterns within and across recordings
[PM17]		✓	✓		Pitch vectors indexed into hashes with LSH of to
[**Ranjani2019sadhana**; RS17]		✓	✓	✓	PT of compositions and DT of renditions for finding patterns from symbolic data using stochastic models

- Repeated patterns from symbolic sequences: CM notations/ transcripts are available as symbolic sequences representative of the notes. The utility of these for finding repeated patterns is explored in [RS17]. Symbolic sequences as transcribed by music connoisseurs are used and stochastic models such as N-grams, skip-grams and multi-grams are considered to understand various temporal contexts. Each subsequence is associated with probability of occurrence in a raga. These sub sequences are also compared with the note sequences listed by musicians. Although the overlap between such stochastic subsequences and musicologically valid note patterns is minimal, it is shown that the stochastic subsequence serves well for ARR tasks. In the subsequent work [RPS19], the transcription obtained from pitch contours are considered for a similar analysis for every rendition and further used for clustering of similar ragas.

6.3.6.4 Summary

Most of the proposed C-MIR approaches are based on pitch contours. This is because of lack of consistent transcripts for compositions. Supervised and unsupervised approaches are used where the motifs are first spotted by musicians or discovered from the data. A summary of different approaches is given Table 6.4.

6.3.7 Music transcription

6.3.7.1 Prescriptive and descriptive transcription

A transcription of a music implies documenting the pitch, onset time, duration, and source of each sound event of a rendition [HKD06].

The role of transcriptions and notations conveying the essence of the melody in CM has been debated. This is more pronounced due to the importance of gamakas in renditions which are seldom captured in transcriptions. In this subsection, we first summarize two types of transcription namely, prescriptive and descriptive. While the prescriptive notation is a minimal notation in that, one can not render by simply rendering the notation, a descriptive notation attempts to be an authentic representation of the audio.

Early major works of CM, during the 9th century such as *Matanga Muni's Bhrihaddesi* contain solfegenotation [Dor14]. However, this notation has not been widely adopted. In the 17th century, CM became codified the way we know it today through the work *Chaturdandi Prakashika* by *Venkatamakhi* [Sri04]. Although raga renditions are associated with florid dynamism across notes, the notations used are prescriptive in nature. These prescriptive transcriptions (PT) form the basic skeletal structure of a composition and are usually regarded as generic guidelines. Many musicians are of the view that CM cannot be notated as it follows an aural tradition, and even if notated would not serve much purpose as much of the melodic information is in the gamakas. Hence, PT is presumed to be limited in terms of information and has been insufficiently studied.

The descriptive transcription (DT), similar to that followed in western classical music, for CM should offer a precise rule, wherein the note sequences follow those in the rendition very closely. Therefore DT must capture all the nuances of the renditions including those that are replete with gamakas. Although there have been attempts for obtaining DT from musicians, it is observed to be a very cumbersome and time consuming process.

Availability of such transcriptions can reveal the structures of CM in accordance with the theory that discrete elements carry the structural information of musical forms while expressions are realized as continuous variations [S89].

6.3.7.2 Challenges

The challenges in the study of CM through notations for C-MIR tasks can be summarized as:

- Prescriptive transcription (PT)
  - Lack of *gamaka* structures in the notation
  - Lack of availability of PT for extempore renditions
  - Inconsistent notations for a given composition based on school of learning
- Descriptive transcription (DT)
  - Not well-defined in terms of granularity
  - Non-availability of DT for any rendition
  - Cumbersome and time-consuming for transcribing a rendition

6.3.7.3 Approaches

Manual transcriptions of various renditions of Mayamalavagowla and Saveri are considered for analysis in [Sub02]. Representations for gamakas are studied in greater detail in [Kri04a; Kri04b; Kri04c] resulting in the definition of multi-dimensional musical atoms in the pitch time space. Three types of gamakas are represented - X+, X-, X*Y where X is representative of the note on which the *gamaka* is anchored and Y is representative of the adjacent note to which the ornamentation swings towards. Although, there have been no computational approaches to detect these melodic atoms, it is shown that it is possible to simulate the perception using synthetic contours that are generated to model these inflexions. The work in [SWM11b] models the gamakas as movement between 2 focal pitch values through a 4 component representation - focal pitch, attack, sustain and release (PASR). Manual transcription is considered and the representation is critically examined through perception of synthesized version with speed modifications. In [SWM12] the stage and dance components of the gamakas are represented through PASR components. PT of the composition is considered and the PASR representation of trigram of notes is considered for representation of gamakas. These approaches are not completely automated and hence may not be scalable, and the number of examples chosen is rather small.

Audio-score alignment method for CM is explored in the work of [SKX16]. A varnam dataset is created and associated with PT. The proposed computational approach aligns the pitch content with the PT in a hierarchical manner - (i) at metrical cycle level and (ii) at note-level. The transcription accuracy is evaluated for the cycle-level alignment, as note-level alignments can not be concurred with PT.

HMM based transcription is proposed on varnam dataset in [Raj+17b]. The HMMs are raga specific and use tala information for initializing the HMMs. DT is used as ground truth for the transcription task. It is observed that it is necessary to have a “ rich transcription for ground truth” which implies the need for DT for ground truth.

A scalable computational approach for obtaining DT of CM using critical points of the pitch contour is proposed in [RPS17a]. The critical points are quantized using SC-GMM to their most likely note position.

The evaluation of the representation is through perception tests as well as comparison of the accuracies of transcription with manually obtained DT as well as PT and is found to be a more accurate representation of the DT - this is one of the first attempts at resynthesizing automatically transcribed CM. The work has been extended in [Ran+19b] to obtain transcription on a pitch-time grid as well by making use of estimated the akshara pulse period for the time component [SS14b]. It is found through perception experiments that PT quantized contours could be more preferable to pitch only quantized versions in CM.

Pitch contours have been represented using constant pitch notes (CPN) and stationary points of transients (STA) in the work proposed in [VAM17b]. This model in turn is used to measure precision of notes in a rendition [Vir+20d; VMA18]. In this first reported work on precision of notes in CM, the target pitch-values of CPNs are obtained statistically as the mean-values of a pitch class. Only those STAs adjacent to a CPN are considered for measurement. It is observed that the CPN have deviations with respect to the target of order of 10-20 cents while STA ranges of the order of 45-85 cents. It is inferred that CM accommodates certain flexibility in terms of grammar. A transcription technique in terms of CPNs, STAs, and state information decoded by a Viterbi algorithm is proposed in [Vir+20e]. Pitch contour transition between 4 states - silence, CPN (anchor and non-anchor) and STA is explored. Perception tests are conducted to evaluate the transcriptions.

6.3.7.4 Summary

Although musicians are of the opinion that transcription for CM may not be of interest, they are definitely valuable in terms of C-MIR tasks such as music learning applications, transcribing large repositories of music and for compact representation of the music form. It can be observed that while initial studies concentrated on domain knowledge centric manual transcription approaches, the efforts of the past 5 years have been towards computational approaches for transcription while including the domain knowledge into the formulation; the former brings in nuances of CM while the latter approaches from a scalability perspective. Table 5 summarizes the works compactly. These works have resulted in the following theses works - [Kod16], [Sub13] and [Ran20]. Succinctly, while was largely believed that it is not possible to transcribe gamakas as they are considered to be unique, computational analysis have shown through analysis by synthesis that it is possible to transcribe gamakas, scale the transcription approach across ragas while retaining the structure of raga largely.

6.3.8 Music generation and synthesis

No C-MIR summary is complete without references to the possibility of a perceptually acceptable machine synthesized melody. At this stage, there have been a few models that address computer generated CM. All the methods are dependent on availability of explicit transcriptions.

One of the earliest works is reported in [Sub99]. A system where the user enters notations is developed. It is observed in this work that it is straightforward to synthesize DT into pitch and duration, whereas the complexity arises for gamakasṄotations indicative of gamakas have been introduced through symbols such as hyphen, colon, semicolon, comma, back slash, forward slash and less than and greater than. A subsequent report by the same author in [Sub12] describes the capabilities of “Gaayaka” and “AddGamaka” softwares which have been developed for synthesizing ragas from DT by explicitly collating gamakas in a “ragaspecific *gamaka* definition file”. These are in-turn used to replace the original note (of DT) with the appropriate *gamaka*.

Similarly, in [Sub13], DT along with the proposed *gamaka* representatives of CM are used for synthesis of acceptable raga renditions. The synthesis interface allows for users to vary the tempo of the synthesized CM.

The work presented in [GS19b] proposed a layered state space graph for Carnātic music where each layer is modeled with a Markov chain. The model is in line with the 2-component stage-dance model proposed in [**srikumar2012compmusic**]. State transition matrices are empirically calculated using note sequences as obtained from [Kal] and [Rao95].

All the music generated is validated through listening tests.

Table 6.5: Summary of approaches for CM transcription

	Manual/ Automated	PT/DT	Metric information	Approach	Evaluation
[Sub02]	M	PT	-		
[Kri04a; Kri04b; Kri04c]	M	DT	-	Melodic atoms to synthesize the perceptual effect of gamakas	Simulate the perception using synthetic contours
[SWM11b]	M	DT	-	Gamakas\as movement between 2 focal pitch values through a 4 component representation - focal pitch, attack, sustain and release (PASR)	Perception of synthesized version with speed modifications
[SWM12]	A	DT		PASR representation of stage and dance component	
[SKX16]	A	PT	✓	PT at metrical cycle level and at note-level.	Evaluated for the cycle-level alignment
[Raj+17b]	A	PT	✓	HMM based transcription, tala\information for initializing HMM	Evaluated with PT of the varnam\dataset
[RPS17a]	A	DT	✗	Quantize critical points of pitch as notes	Evaluated with perception tests as well as DT ground truths
[Ran+19b]	A	DT	✓	Quantize critical points of pitch and time	Evaluated with perception tests
[VAM17b]	A	DT	✗	Components of pitch contour representation as CPN and STA	Evaluated with perception tests
[Vir+20d; VMA18]	A	DT	✗	Quantized CPN and STA for precision of notes	Evaluated with perception tests
[Vir+20e]	A	DT	✓	Transcription technique in terms of CPNs, STAs, and state information decoded by a Viterbi algorithm	Evaluated with perception tests

6.3.9 Road ahead

The C-MIR work till date can be seen to be bottom-up work and has focused on analysis towards obtaining features, approaches to enable C-MIR technology tools (such as classification/verification, transcription, segmentation and recommendation systems). It is worth highlighting that the approaches enlisted in the above sections till date are based on analyses of huge amounts of data, across various ragas, encompassing various singers. The insights and results from various techniques are at a stage where it is possible to build applications to further enable listeners and learners. Prominent applications that have been built to enable end users with such tools are:

- Riyaz (url) an android application which facilitates music learning for beginner to intermediate level music students by making their practice (riyaz in Hindi) more efficient. Riyaz offers hundreds of exercises along with singing lessons in Hindustani music and Carnatic music. It is also a personal vocal trainer that listens to the singing user to provide precise feedback.
- Saraga (url) is an android app which proposes an enriched listening approach on a collection of Carnatic and Hindustani music. It includes visualizations and inter and intra-song navigation patterns that present musically relevant information to the user. These synchronized visualizations of musical facets such as melodic patterns, samas locations, and sections, provides a user with better understanding and appreciation of these music traditions.
- Patantara (url) is a web-application to create and share high quality Carnatic music notation. It provides a learner to have audio tracks integrated with music notation. It also includes slowing down the track being played to enable ease of learning as well as pitch shifting to enable association with a comfortable tonic frequency.

It is worth noting that the research on melody till date is based on uni-dimensional pitch featurization. While it is imperative to dig deeper to bring in more insights along the pitch features, there is also a need to analyse melody vis-a-vis the entire concert. With this perspective, the future work can be broadly grouped under 3 categories:

1. Low-level features - All of the melodic analysis till date considers pitch and time as the dominant features in order to map towards raga structures. It is also equally important to understand the perceptual impact of other attributes such as intensity, timbre alongside pitch in a raga rendition. Such a multi-modal analysis can result in a complete description of the music form.
2. Higher-level abstraction - Some of the wish list entries seen as future work can be summarized as:
    a) Tonic identification in complex structures such as graha-bedha
    b) Quantification of ragaspecific grammar, and study of similarity of ragas
    c) Better understanding of prescriptive notations of CM
    d) Quantifying differences across teacher-student lineages
    e) Multi-modal analysis and synthesis by connecting rhythmic, timbral aspects alongside melodic aspects
3. Technology - For the research outcomes to be adopted as technology, there must be an effort to ensure that the end users (CM learners, connoisseur and performers) are included to understand the effectiveness of the technologies in delivering their needs. In this regard, the technology must not be limited in terms of mobile apps; it is desirable that the technology is integrated as a part of the listening experience and hence the bigger aim must be towards culturally aware platform independent music players.

6.4 Rhythm and percussion analysis

6.4.1 Context

Rhythm exists in multiple forms in human lives in the form of regular events, biophysical activities and movement. In a music rendition, the musical events are arranged along time. These patterns along time are referred to as rhythm. Hence, rhythm forms a fundamental aspect of music. Tala is the rhythmic framework of CM and HM; it provides a broad structure for rendition and repetition of melodic and rhythmic phrases, motifs and improvisations [**ajay2016phdthesis**]. The different artists in a concert perform within the rhythmic framework provided by the tala of the music piece, with sufficient improvisations allowed within the framework of the tala. The lyrical phrases/passages by the lead artist, or the percussion strokes by the lead percussion instrument mridanga, and other supporting percussion instruments such as ghata, kanjira, thavilor morsing are arranged around the hierarchical time cycles. Tala comprises of hierarchical time cycles that provide a metrical structure for different rhythmic and melodic events to occur within a music piece. In a CM rendition, a connoisseur or an expert can focus on multiple rhythmic aspects and identify components of a tala such as avarthana, *laghu*, dhrutam anudhrutam, kala, jaati, nade, akshara, sama, edupu, along with the varied strokes of the mridanga. In this context, rhythmic analysis of C-MIR is a broad inclusive term to automatically capture the instances, identify and/or discover the patterns and map these to the cognitive abstractions of CM from the audio signals.

6.4.2 Overview

Rhythmic analysis in CM aims to extract different rhythmic events and components of metrical structure of an audio music recording. It ranges from (but not limited to) estimating the onsets as a mid-level features from a recording or performance to mapping them onto higher level attributes such as aksharas, nade, avarthana, tala, sama, edupu, solkattu. Rhythm analysis in this context hence refers to identification of low level rhythmic

Table 6.6: Summary of C-MIR rhythmic analysis

Author / Report	Onset detection	Tala\analysis	Analysis of taniavarthana	Percussion music synthesis/ generation
Ajay Srinivasamurthy, PhD thesis	✓	✓	✓	
Jilt Sebastian, PhD thesis	✓			
Jom Kuriakose, [**kuriakose2015ncc**], [**dawalatabad2018interspeech**]			✓	
Carlos Guedes				✓

Table 6.7: Summary of datasets available for C-MIR rhythmic analysis

Database	Task	Availability
Charsur dataset	Instrument diarization	Available on registration and for research purposes
Sangeethapriya	Artist identification	Available on registration at http://www.sangeethapriya.org
Carnatic music rhythm dataset	Tala\analysis	Available under Creative Commons licenses: https://compmusic.upf.edu/datasets
Mridangam stroke dataset (Akshay)	Stroke transcription	https://compmusic.upf.edu/mridangam-stroke-dataset
Mridangam solo dataset (UKS)	Onset detection, percussion transcription, rhythm and percussion pattern analysis, and mridangam stroke modeling	https://compmusic.upf.edu/mridangam-tani-dataset
Saraga Carnatic collections	Tala\analysis, and rhythm based structural segmentation	Creative Commons licenses: https://mtg.github.io/saraga/

events from audio signals as well as mapping those events to different components and aspects of the tala framework. In addition, the rhythmic events could also be used to identify rhythmic patterns within the tala framework, helping in defining different rhythm similarity metrics. In this respect, the C-MIR for rhythm can be seen under the following few broad headings [**ajay2016phdthesis**]:

- Onset detection
- Automatic meter and rhythm analysis (estimation of tempo in terms of aksharas per minute, tala recognition, estimating the nade structure and edupu (offset) estimation, tala and sama tracking of the rendition)
- Percussion pattern transcription (solkattu) and discovery of the rhythmic structures
- Rhythmic analysis of symbolic scores
- Rhythm similarities

The following sections summarize the sub-problems, the challenges associated with them and the existing approaches. Table 1 summarizes the different sub-problems addressed. Most of the works are collaborative efforts; this table is not intended to downplay the inputs and efforts of other contributors, but to provide a consolidated reference for an interested reader. We refer to specific works in detail in subsequent subsections. A summary of the datasets used/available for each of the sub-problems is detailed in Table 2.

6.4.3 Onset detection

6.4.3.1 Onset

The instant chosen to mark the start of transient of note or stroke is referred to as onset. Transient itself is defined as the interval of time over which the signal rapidly changes[**bello2005tsap**]. Automatically detecting onsets of musical note or stroke is the first step for rhythm analysis and is irrespective of the cultural form.

6.4.3.2 Challenges

The challenge in onset detection of any music culture is in identifying indicative features of an offset. In some recordings, at any given time, there are multiple active instruments; this results in varied “attack” patterns.

Hence, detecting onsets is dependent on the transient nature of the instrument. Again, these challenges are not specific to CM and have been identified across music cultures [**bello2005tasp**; **dixon2006**; **klapuri1999**].

One of the challenges pertaining to onset detection in CM can arise due to the harmonic nature of the percussion instruments such as mridangam [**cvraman1920nature**]. This necessitates identifying models and features to accommodate this nature of the signal. Also, some of the percussion instruments are tuned to the tonic frequency specific to the concert. The algorithms must normalize the variation of tonic frequencies across concerts.

Another challenge pertaining to onset detection in C-MIR pertains to annotated datasets of CM. This is important to evaluate and analyse the algorithms and improve on the existing techniques.

6.4.3.3 Approaches

Onset detection has been well studied, with a detailed tutorial in [**bello2005tasp**]. The approaches can be considered to contain 3 steps - (i) signal pre-processing to enhance the relevant bands of the audio signal through a filter bank approach (ii) identify indicative features such as spectral energy and flux to map to an onset detection function (iii) picking peaks of the onset detection function based on thresholds along with the associated intensity to measure strength of the onset.

In the case of polyphonic and heterophonic signals, the onset detection is approached from a signal separation perspective; models that separate signals into either constituent instruments or into melodic/percussion components forms the pre-processing step [2]. This is followed by onset detection of either the separated individual instrument component [**tian2014icassp**] or percussion component only [SS14b].

There are many onset detection approaches in the literature. At this point, we concentrate on approaches that either propose algorithms and/or evaluate for CM.

The work of [SS14b] is one of the earliest towards rhythmic analysis of CM. The method suppresses the harmonic components by obtaining residual post suppression of an estimated f_0 track. As the next step, two frame level spectral-flux based onset detection functions are proposed (i) spectral flux of the whole frequency range of the spectrogram (ii) spectral flux only in the range of 0-120 Hz to capture the low frequency onsets of the left(bass)-side of the mridanga. The onset detected functions are then used for tracking akshara period and sama tracking.

In [**manoj2015ncc**], detecting onsets in CM by considering 5 major percussion instruments - mridangam, ghatam, kanjira, morsing and thavil is studied. The percussive strokes in Carnatic music are modeled as an AM-FM signal. Such a model includes the presence of harmonic structures of the percussion instrument. The music signal is differentiated to emphasize the location of all onsets. It is seen that the information about an onset is contained in the envelope of the analytic signal (estimated using the Hilbert transform). This envelope forms the onset detection function from which onsets are identified using group delay function. The work is evaluated on taniavarthana segments of CM and evaluation is in terms of presence of detected onset within ±50ms of the ground truth. This proposed signal processing approach is contrasted with a CNN (convolutional neural network) based approach.

The above work is further extended in [**jilt2017ismir**]. A percusssive / non-percussive separation step precedes onset detection technique of [**manoj2015ncc**]. A deep recurrent model (DRNN) is trained with alapana and taniavarthanam to separate the percussion from the voice by generating a time-frequency percussive mask. The mask in turn separates the percussive portions from the mixture signal. The separated signal is used for onset detection in accordance with [**manoj2015ncc**].

[2] This pre-processing step itself can be addressed in multiple ways - through HPSS (Harmonic Percussive Source Separation) [**thoshkhana2011ismir**], by suppressing f_0 track and thus obtaining the residual [SS14b], suppressing singing voice [**zapata2013icassp**]

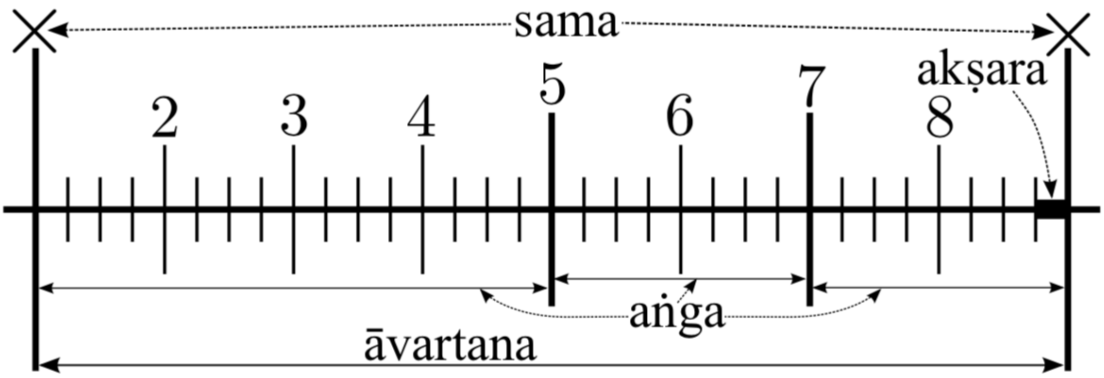


Figure 6.5: Components of a talacycle. Source: [**ajay2016phdthesis**]

6.4.3.4 Summary

Onset detection is a generic MIR task. The additional challenge of presence of harmonic nature of the instruments is addressed either by suppression of harmonics or by including it in the model and considering only the envelope. Alternatively, just the onset detection function is estimated (but not evaluated) and forms the input for the subsequent task. Instrument specific onset detection tasks are still unaddressed for CM.

6.4.4 Talaanalysis (or Meter analysis)

6.4.4.1 Tala

The rhythmic counterpart of raga analysis is the tala analysis. Tala provides the basic framework to track the rhythmic structure of a rendition and hence its identification is crucial for any rendition. A complete analysis of tala of a music piece involves an estimation of the hierarchical metrical structure of the music piece and hence the hierarchy of events in a cycle of the tala, aligning the audio recording with different events and rhythm related metadata specific to the music piece. Specifically, it involves the estimation of different components of the tala at various levels - tempo, akshara, laya, sama, anga, edupu. The relation between these components are as depicted in Figure 1. It is important to track the various components of the talafor most other C-MIR tasks such as rhythm segmentation and extraction of rhythmic patterns to define similarity across renditions. In specific, estimation of sama, the time instants where a talacycle begins in an audio recording is crucial for any further structural analysis of the music piece, since most melodic and rhythmic events are aligned with the start of a cycle.

In the context of C-MIR, the name of the tala for a music recording is often available as a part of metadata of the CM recordings. From the name of the tala, we can infer the underlying metrical structure of the piece, and hence time-align and tag a music recording with tala related events. However, untagged, live concerts may not have tala information.

Since a common hierarchical structure of taladefines three levels - akshara (sub-divisions), beats and sections (anga) within each time cycle, in the context of tala analysis, C-MIR tasks involve recognizing the tala, estimating the tempo, tracking beat locations, estimating akshara periods, tracking sama locations, inferring anga boundaries. Tala tracking without any prior information is referred to as tala inference and tala tracking with knowing the tala is referred to as tala tracking; in informed tala tracking variant, it is assumed that some amount of information such as tempo range or availability of annotations of few instances of beats, down-beats can be provided in addition to the tala name. These can be either manual or automatically annotated [**ajay2016phdthesis**].

6.4.4.2 Challenges

The theoretical frameworks in Carnatic music allow for cyclical structures at time-spans different from the avarthana. Also, metrical structure might also vary within a piece (4 akshara and 3 beats in rupaka tala can also be played as 3 akshara and 4 beats within the same tala framework during improvisatory passages). Disambiguating between talas having the same number of aksharas in a cycle is another challenge.

Owing to the above challenges, tala analysis is a complex task and involves several subtasks such as cycle length estimation, nade estimation, identifying number of aksharas estimating akshara period, sama detection and edupu estimation. Some of the challenges in tala tracking arise due to presence of expressive timing and varying tempo (attributed to absence of metronome) in CM. Tala recognition itself requires estimation and tracking of components of a tala. Dataset challenges such as absence of ground truth for akshara location must also be considered.

6.4.4.3 Approaches

One of the earliest approaches towards rhythm analysis in CM is [**ajay2012compmusic**]. Onsets detected from spectral flux are used to obtain inter-onset-interval (IOI) histogram. Tempo is estimated for the entire song (with the assumption of a constant tempo) and a dynamic programming approach is used for beat tracking. A beat similarity matrix is computed, and used for long-term periodicity estimation. Sub-beat structure is estimated from the IOI histogram with the help of a comb filter. The metrical level ambiguity limitation of beat tracking approaches exists and can translate as an integer multiple increase/decrease of sub-beat and long-term periodicity estimates.

The work of [SS14b], two frame level onset detection functions are considered. This is used to obtain a Fourier transform based tempogram where the most predominant curve is indicative of the akshara pulse track. The optimal path is estimated using dynamic programming with a local continuity constraint. The akshara pulse locations are peaks of the onset detection functions. The best candidates of the peaks of onset detection function are selected so as to account for periodicity of akshara pulse period and onset strength. For tracking sama, a self-similarity matrix is constructed using frame level MFCC features. Peaks of the novelty function are indicative of significant timbral change. Two methods are proposed for sama detection (i) Most prominent peaks of each segment (ii) Super posing tala cycle periodicity on to the candidates. These with left bass onsets are used to obtain samaestimates. The shortcoming of this approach is that related components such as tempo, akshara and sama, are tracked separately.

A Bayesian model to jointly estimate tempo, downbeats and rhythmic class is proposed in [**holzapfel2014ismir**]. It uses a bar pointer model based on a Dynamic Bayesian Network (DBN), where the bar pointer is assumed to traverse a bar. The bar pointer at each audio frame is represented by three hidden variables - tempo, rhythmic pattern, and position inside a bar. Regularity of the outputs is enforced by incorporating a simple HMM. A 2-dimensional onset function based on spectral flux forms the observation signal. The observation probabilities are modelled by GMM. The optimal state sequence is obtained using the Viterbi algorithm by maximizing the posterior probability of the hidden states to obtain beat, downbeat, and rhythmic class estimates.

Particle filter based approach proposed in [**ajay2015ismir**] gives a faster inference in a DBN. Some of the modifications to the earlier proposed DBN are the state space of bar position and tempo is now represented with 2 continuous hidden variables. Faster inference is obtained by using approximate inference techniques. The resulting inference schemes are observed to remain accurate even for long and complete renditions of CM. This model is further improved upon in [**ajay2016icassp**] where a section pointer model is proposed for analysis of renditions with long metrical cycles. Musically meaningful segments, anga or sections (sometimes of unequal lengths) form the basic unit for tracking, and hence the resultant model is referred to as section pointer model. Uninformed and informed (tempo class information is known) tracking tasks are evaluated with the model. Significant improvements are observed with section pointer models and with informed tracking.

In the subsequent work [**ajay2017ismir**], informed meter tracking is explored with 3 different levels of prior information - sama, tempo and sama with tempo. The bar pointer model of [**ajay2015ismir**] is considered and the prior information is suitably incorporated into the hidden variables in initialization phase - tempo variable is initialized to the known value, while sama instance is used to initialize bar the position to zero. It has been found that informed meter tracking has advantageous performance to that of tracking which in turn has shown to be advantageous with respect to inferencing tasks.

Given the tala information and alignment with a music piece for a large corpus of music, we can use it for a statistical analysis of rhythmic elaboration in Carnatic music. Analysis of a large data corpus provides us insights into aggregate behaviour of rhythmic events such as beats, sama, inter-sama interval, inter-akshara interval and their histograms, and intra-cycle rhythm patterns as reported in [**ajay2019frsm**], which is an example of data-driven statistical approach to musicology.

6.4.4.4 Summary

Approaches towards meter inference, meter tracking (model extensions), tempo-informed tracking, and tempo-sama-informed tracking have been proposed. Initial approaches estimated each of the components of tala. Joint estimation of meter, tempo estimation and sama detection are proposed with a probabilistic model. Informed meter tracking results inculcate prior information and provide more promising performance. Melody also can be used to track the progression through a tala with several melodic and lyrical markers indicating the sama. Incorporating melody and lyrics based features into the observation model can additionally help to improve meter analysis performance.

6.4.5 Analysis of taniavarthana

6.4.5.1 Taniavarthana

In a CM concert, taniavarthana refers to the solo percussion-based performance which exhibits creative and technical skills of the percussionist(s). It, generally, follows the main performance in a concert. It is an elaborate rhythmic improvisation within the framework of the tala, but with much improvisation on the percussion patterns. The taniavarthana strives to showcase the tala with a variety of percussion and rhythmic patterns that can be played within the framework of the tala. In the presence of multiple percussionists, the performance can be viewed as a duel and complement each other through sequential alternate performances, with all instruments coming together to a cadential end. The patterns played can last longer than one avartana, but stay within the framework of the tala. An oral recitation of these syllables, called konnakol, is often a part of a taniavarthana.

In the context of rhythmic analysis of C-MIR, the taniavarthana is a rich representation of the strokes of various percussion instruments and artists' playing styles. This provides opportunities in terms of the stroke transcription, analysis of grammatical structures leading towards possibilities of generation of percussion music.

6.4.5.2 Challenges

The first step towards analysis of taniavarthana is in identifying taniavarthana segments in a concert. Given such a concert segment, the presence of multiple instruments, the difference in their attack patterns and timbre, tuning of the instruments to the concert-specific tonic frequencies, unique mapping of strokes to konakkol syllables, instrument independent transcription of the strokes, analysis of similar syllables/patterns constitute some of the challenges in C-MIR. Other challenges include capturing the grammatical structures of the percussion performance.

6.4.5.3 Approaches

The approaches can be roughly grouped into 3 aspects - (i) identification of taniavarthanasegment in a concert (ii) instrument based analysis (iii) stroke transcription. We detail the approaches proposed below:

Identification of taniavarthanasegments

The works of [**sarala2013ismir**] and [**ranjani2013ismir**] provide a general framework towards segmenting concerts into various items/forms/renditions. The [**sarala2013ismir**; **saralaphdthesis**] uses applause as landmark events, along with CFCC features and pitch histograms for segmenting concerts. In [**ranjani2013ismir**; **ranjaniphdthesis**], presence of meter, percussion instrument and strength of onsets are the features considered. The same has also been captured through modulation domain features of STE. These are captured in detail in Section 4 as a part of structural analysis.

Instrument based analysis

Instrument diarization has been attempted in [**Dawalatabad2018interspeech**]; the problem is formulated along similar lines as speaker diarization. Information bottleneck (IB) based approach is proposed; fixed length audio segments are clustered iteratively in a bottom-up manner. However, fixed audio length poses problems as the stroke rate varies across segments. To counter this, stroke rate is estimated using onset detection as proposed in [**manoj2015ncc**] and the same is used towards a variable length IB approach such that the number of strokes is constant in each segment. A KL-HMM realignment is used to refine the boundaries of diarization.

Stroke transcription

The earliest report on the acoustic properties of mridanga and tabla is that of [**cvraman1920nature**; **cvraman1934pias**] and is close to a century ago. The construction of the drumhead, position of excitation and the resulting modes of vibration are detailed using sand figures. In the study reported in [**bsramakrishna1954jasa**], measured values of the frequencies of the first nine modes are shown to confirm harmonicity indicated by the theory. These are some of the earliest works on characterization of strokes.

From the C-MIR perspective, modal analysis of strokes was analyzed from data using NMF in [**akshay2013icassp**]. In the first step, for every recorded mode, a single spectral vector is estimated using NMF decomposition. A concatenated dictionary of such spectral basis forms the modes on to which the strokes are projected. The activations of the modes are used to determine the identity of the stroke played. Hidden Markov Models are adopted to learn the modal activations for each of the strokes. The method is dependent on prior knowledge about the specific modes of the instrument. A data-driven, tonic independent transcription of the strokes of the mridanga is proposed in [**akshay2014jasa**]. Magnitude of the Fourier transform of CQT representation of the signal is found to be tonic invariant and forms the features for NMF decomposition. Statistics of frames between 2 onsets of the activation functions form the feature vectors and SVM is the classifier used.

The thesis [**ajay2016phdthesis**] explores a modeling of timbre and sequential syllables. Timbre is captured with MFCC with velocity and acceleration coefficients. A seven state L-R HMM, with a 3 component GMM per state. Viterbi algorithm is used to obtain transcription of strokes. Further, an approximate search of query syllable sequence is proposed using a variant of RLCS algorithm. Enhancements to RLCS are proposed to reduce insertions, deletions and substitutions (i) warped density RLCS (penalize low density values more than the higher ones) (ii) RLCS based on timbral distance. A similar work [**kuriakose2015ncc**], explores the use of CFCC features with a three-state HMM trained for each stroke, a language model is learnt from the data with Viterbi decoding for stroke transcription.

A combination of acoustic and semantic approaches for the contextual transcription of mridangam strokes is considered in the work [**kaustuv2020timbre**]. Transcription based on 5 frequency bands is compared with konakkol, and it is observed that there is a many-to-one mapping from konakkol syllables to the mridanga strokes. It is hypothesized that this could be to accommodate physiological and language constraints of speech production.

6.4.5.4 Summary

Analysis of taniavarthana section of a concert is an important C-MIR task, and commences with the task of automatic segmentation of CM concert. Given a taniavarthana section, identifying the constituent instruments and transcription of the strokes (of mridanga) are some of the tasks that have been addressed. The proposed approaches are largely motivated from similar formulations in the speech processing realm. It is important to note that there have been limited studies. There is a need for more focused attention towards study of taniavarthana, possibility of extending the transcription task towards other instruments, timbral similarity measures and improving syllable boundaries of phrases.

6.4.6 Percussion music synthesis and generation

6.4.6.1 Synthesis

Generative music involves an algorithmic approach to include the knowledge, strategies of the artist, the grammatical information of the music form along with perceptually pleasing phrases as match the musical abilities of an expert. A better generative model must lead to a perceptually acceptable machine synthesized percussion pattern.

6.4.6.2 Challenges

The challenges in percussion music synthesis/generation is to capture the strategies of percussionists in terms of grouping (both short and long term phrase) and thus categorize the strokes into phrases and learn the music grammar. Hence, it is important to define “naturalness” of computer generated music - be it in terms of dynamic accent, grouping, or even a better language model [**ganguli2019tee**].

6.4.6.3 Approaches

The work of [**konstantinos2016smc**] studies the rhythmical organization of CM. Percussion solo performances of adi tala are considered. Each stroke is represented as low/ mid/ high frequency bands (110-190 Hz for low, 190-600 Hz for mid and 600- 1200 Hz for high strokes), inter-onset interval (IOI) between strokes and normalized velocity value (computed using strength of normalized onset detection function. N-gram analysis is used to model the stroke sequence representation. For a given stroke initialization, the stroke events based on the past $N-1$ strokes are considered. The generated stroke is based on that event which has the highest probability among those that start with the last $N-1$ stroke events. The generated sequences are evaluated by experts on the basis of presence of recognizable CM rhythmic patterns, patterns in appropriate metrical positions and long term evolution and progression. Higher grams were found to be more successful in capturing longer term structures.

In the subsequent work [**carlos2018smc**], strokes are grouped together in two ways - (i) pre-recorded phrase variations which form a dictionary (ii) segmentation approach if they are nearest neighbors to each other. The idea here is to obtain grouping of strokes akin to how it is grouped in solkattu and thus have more grammatical long and short term phrases. K-means is used for clustering of similar groups of strokes. In order to generate talas of a certain arithmetic partition, groupings from the same cluster are concatenated. However, longer term phrases were not captured. In order to understand the reason, a recent study [**ganguli2019tee**] approaches rhythmic phrases as a gestalt. A recorded phrase is played at 2x, 4x speed sounds and compared and contrasted with the same phrase played by the musician at 2x, 4x speed. It is observed that at such high speeds, the individual segments are not so important and may not even be amenable to segmentation (as seen through RMSE and spectral flux).

6.4.6.4 Summary

The studies with respect to percussion synthesis is recent, and there is a need for more concentrated efforts towards such an analysis. It is to be noted that the studies are right now at initial stages and addresses solo percussion recordings. Short term patterns have been models using N-grams. It is important to capture long term grammatical rules, understand improvisations, and capture interplay of melody, lyrics, and rhythm in a concert.

6.4.7 Future work

Tala analysis tasks such as meter inference and tracking have been addressed in detail in the work of [**ajay2016phdthesis**]. Apart from improving on performance for these tasks, there is a need to explore other features for rhythmic analysis. Formulations which can study rhythm, alongside melody and lyrics is inevitable. There have been limited studies analysing taniavarthana segments of concerts. Percussion pattern discovery has been addressed, albeit to a lesser extent, and is limited to studies with mridangam instrument. It is required to extend such analysis to other percussion instruments of CM. This in turn will open up problems such as understanding onsets across instruments, stroke similarity across timbral variations, the mapping of strokes to solkattu and their effects on learning rhythm. Evaluating commonalities and differences across various playing styles, schools of learning are not only challenging but necessary in rhythmic analysis of C-MIR. Finally, unlike melody for which there exists at least a prescriptive notation that is more or less universal, the notation for percussion varies from school to school. Comparison across different lineages thus becomes difficult. A universal representation is perhaps an MIR task that can be easily addressed.

6.5 Timbral/structural analysis of CM

Music is inherently of repetitive nature and hence often viewed as a concatenation of segments comprising phrases and sections. Such a viewpoint results in a segmentation approach to music summarization and structural analysis. Therefore, structural analysis refers to identifying and characterizing repetitive structures in any form of music. This formulation is independent of the cultural form of music.

6.5.1 Carnatic music structures

CM has a long, rich and illustrious concert tradition. The presentation is rich in a variety of music pieces, typically can last for 2-3 hours, can be led by a vocalist or a lead instrumentalist, along with drone accompaniment and rhythmic accompaniments. The musical ideas of the performer are presented in various forms such as alapana, kriti, viruttam, taniavartana, tanam, tillana, swarakalpana, neraval and so on. The structures of CM can be seen as those across renditions (inter-rendition structures) and those within renditions (intra-rendition structures). These are detailed as follows:

6.5.1.1 Inter-rendition structures (at concert level)

The concert is a sequence of multiple independent musical ideas showcasing the performers' flair in CM whilst also catering to the interests of the connoisseurs. The ideas are showcased in various formats; some examples are listed below:

1. Choice of lead and accompanying instruments

Table 6.8: Summary of C-MIR structural analysis

Author / Report	Inter rendition structures	Intra rendition structures	Instrument identification
Sarala, PhD thesis	✓		✓
Sridharan, Masters thesis		✓	
Ranjani, PhD thesis	✓		✓

Table 6.9: Summary of datasets for C-MIR structural analysis tasks

Dataset	**Tasks**	**Availability**
IITM 50 personal live recordings	Inter and intra rendition structural analysis	Available on request to authors of [**sarala2018sadhana**]
Charsur live concert dataset	Inter and intra rendition structural analysis	Live concerts licensed for research purposes
ARKAY recordings	Inter and intra rendition structural analysis, instrument identification	Available on request to authors of
Sangeethapriya dataset	inter and intra rendition structural analysis	Available on registration at http://www.sangeethapriya.org (No registration required to access this site)

2. Choice of ragas in a rendition within the larger context of a concert

3. Manodharma sangita (extempore), kalpita (those already composed) and a hybrid approach

4. Different music forms such as alapana RTP, swara kalpanas, taniavarthana thillana ...

5. Melodic dialogue between the lead vocalist and violinist

6. Rhythmic dialogue between rhythmic accompanying artists, or between rhythm accompanying artist and lead performer

Such structures are independent, vary across renditions, and are referred to as inter-rendition structures.

6.5.1.2 Intra-rendition structures

The structures within a given rendition are referred to as intra-rendition structures. Different forms of music are associated with varied structures:

1. Krithis - *Pallavi*, *Anupallavi* and Charanas usually with the sequence - *pallavi* *anupallavi* charana1 *pallavi* charana2 *pallavi* ... charanaN *pallavi* *anupallavi*

2. Varnam - *Pallavi* *Anupallavi* charana chittai1 , charana chittai2 , charana ... chittaiN , charana *Pallavi* *anupallavi*

3. RTP - ragam (alapana), thanam *pallavi* *niraval* kalpanaswara

4. Thillana - Jathiswaram for *pallavi*, *Anupallavi* lyrics for charana

5. Taniavarthana - avarthana at various speeds (1x, 2x, 4x), faran, mohara, teermanam

Such structures within a rendition are dependent on the form of music, and are referred to as intra-rendition structures.

The interactions between the melodic, rhythmic and timbral aspects in CM gives rise to nuances of the C-MIR problem. Automatic segmentation of music into such structures at both inter and intra rendition levels is of importance for archival application.

6.5.2 Overview

The subproblems in C-MIR for structural analysis are presented in Table 1 and datasets available for structural analysis are presented in Table 2.

6.5.3 Challenges

Some of the challenges in segmenting CM at inter and intra rendition are listed below:

- Inter rendition structures - The structures are present at the following levels - forms of music, instruments used, extempore or composition variety. Hence, segmentation of music can be conceptualized at various levels of patterns and is driven by the application at hand
  - Segmentation based on form of music
    - The inter rendition forms are independent of choice of tonic frequency, raga, tala, singer/instruments
    - Identifying manodharma and kalpita forms require knowledge of numerous compositions
    - Features which can separate such structures are not evident and may need signals of longer duration (atleast 5-10 s)
  - Segmentation based on instruments used
    - Simultaneous activity of some of the instruments within a segment with same melody line (heterophonic form of music)
    - Presence of harmonics in percussion instruments
    - Wide variety of instruments - vocal, whistling, vina, violin, flute, mandolin, saxophone, guitar, jalatarang, nadaswara, tamboori, electronic shruti box, mridanga, ghata, kanjira, thavil, morsing
  - Segmentation based on improvised or composition
    - Detecting improvised vs composition sections in CM is non-obvious
- Intra rendition structures - The organizational elements and their repetitive patterns are different and are dependent on the form of rendition. Hence it is important to know the form of rendition, so as to be able to identify the elemental units and the repetitive patterns. Additional knowledge of lyrics, vocalizing syllables, metric cycle, identity of instruments and such inputs can aid in identifying repetitive structures. The challenges of extracting repetitions in improvisational music form is worth noting.

6.5.4 Approaches

The earliest work on segmenting Carnatic concerts is [**sarala2012compmusic**]. The work proposes utilization of applause as landmark elements of a concert. The location of applause is estimated by processing the audio signal using features such as spectral flux, spectral entropy. The cumulative sum of the features is used to determine strength of the applause. Subsequent work in [**sarala2013cmmr**], the inter-applause segments are classified into 4 segments - vocal solo, violin solo, composition and taniavarthanam segments. The use of Cent Filter-bank based Cepstral Coefficients (CFCC) that is tonic invariant is proposed thus normalizing the tonic frequency variation across concerts. CFCC is similar to MFCC if seen as a bank of filters; the differentiating factor is that the Cent filter bank, the filters are uniformly spaced in the tonic normalised cent scale. GMM models with CFCC features are proposed for each of the 4 segments. In the follow up work [**sarala2013ismir**], the segments are labelled into different “items” based on similarity measures computed on the warped pitch histograms of the adjacent composition segments - this in essence is detecting raga change, and is based on the fact that adjacent items of a concert will be of different ragas. Dynamic programming is used to calculate the similarity between adjacent segments by warping the cent axis. Empirical thresholds are used to determine similarity-dissimilarity boundaries. Evaluation of applause detection with additional features such as zero crossing rate, spectral energy and use of ML techniques such as SVM, GMM to detect applauses is reported in [**sarala2018sadhana**; **sarala2015phdthesis**].

The work of [**ranjani2013ismir**] proposes hand crafted energy and timbral based features to identify any given 10 sec clip of CM as one of the few considered forms - alapana, krithi, viruttam, taniavarthana, thanam, thillana. The work proposed the use of features such as presence and absence of periodicity in rhythmogram, onset strength to determine presence and absence of percussion instruments, formant distribution to model usage of spread of syllables, offset strength to determine cessation of energy. Simple threshold based classifiers are proposed. Further experiments reported in [**ranjani2020phdthesis**] show that the hand-crafted features can be generalized to modulation domain features of multi-band representation of STE; these modulation domain features are used with a powerful classifier such as SVM for identifying 10 sec clip. It is worth noting that here the structural analysis is posed as a classification problem.

The work of [**sridharan2015cmmr**] proposes to segment a rendering of a composition using a partial *pallavi* query as the template. CFCC energies are computed for both the query and the composition. Slope of cent filterbank along frequency is proposed to counter the effect of mridangam strokes which can potentially reduce the correlation between query and candidate. In the same work, it is proposed to automatically obtain the *pallavi* query - it is based on the assumption that a composition commences with the *pallavi* - this is achieved by checking for correlations of segments of varying lengths ranging from 0.5 seconds to 3 seconds. In order to ensure that the composition part of a musical item is separated from alapana, the work of [**pvkrishnaraj2016ncc**] considers segmenting an identified musical item into alapana and krithi segments. Symmetric KL divergence between normalized CFCC features of adjacent segments is proposed. The underlying intuition is the metric can capture the timbral change from alapana to kriti due to the presence of percussion instruments. The threshold change points are potential candidates; the identification of true boundaries is proposed with GMM models for alapana and krithi classes.

Identifying the presence of simultaneous instruments in polyphonic signals using monophonic models as proposed in [**ranjani2015icassp**] has been utilized for CM in [**ranjani2020phdthesis**]. The monophonic models are combined through use of latent variables and the temporal and simultaneous structures are captured through monophonic HMM models and factorial models of the same. The approaches are evaluated on multi-track concert recordings of CM.

The work of [**sarala2017transactions**] does not directly address content based structural analysis of CM. The work proposes a framework - a category theory [3] framework to merge CM ontology with MIR tasks. The framework not only proposes to map meta-data of MIR encompassing fundamental aspects of Carnatic music to category-theoretic ontology, but also includes details of analysis of Carnatic music with experimental evaluation. Towards this end, a three layered structure is proposed - (i) First layer contains a categorical representation of the characteristics of Carnatic music, details the MIR task, features and approaches for the particular analysis (ii) Second layer in the generic model is a database schema which mediates the category theoretic ontology and the query processing system (iii) Final layer supports query analysis for information retrieval purposes.

6.5.5 Summary

Structures of CM concerts and recordings is quite rich; most of the above described works intended for identification and segmentation of inter and intra renditions have been consolidated in the following theses - [**sarala2015phdthesis**; **sridharan2017mastersthesis**; **ranjani2020phdthesis**]. In contrast to the melodic and rhythmic analysis, there has been limited efforts towards this topic. There is a need for more concentrated work in order to be able to segment the audio recordings and label into desirable structures.

6.5.6 Future work

The unattempted topics are as following:

- Manodharma sangita vs kalpita sangita
- Discovering the constituent instruments of any concert

[3] A category is a mathematical structure that appears much like a directed graph: it consists of objects (often drawn as nodes or dots, but here drawn as boxes) and arrows between them denoting the relations between objects [**spivak2014mitpress**]

- Discovering repetitive sectional structures such as ABABCACAC akin to the western structural form, where sections A, B, C are dependent on the CM form

6.6 Conclusions

In this Chapter, an attempt has been to document the various efforts on MIR for Carnatic Music. Three important aspects of Carnatic Music have been documented. There are about 200 melodies seen commonly in Carnatic Music concerts today. Identification of melodies, characterisation of melodies and transcription of melodies is first studied. Rhythm in Carnatic music is based on tala, and there are about 175 different talas. Identifying the beat cycle, and mapping to the different angas of a talais reviewed. Percussion instruments in Carnatic music also have their own improvisations which can be studied in terms of stroke patterns. Machine learning is used to identify these patterns. Finally the structure of a concert, segmentation of a concert into items, and the structure of each item are also studied.

While the work done is quite comprehensive, it is still nascent. There are many other aspects in terms of timbre, lineage, the interplay between melody, percussion, and lyrics of the composition. The improvisational prowess of an artist are yet to be studied. While there are huge collections of CM available on the Internet/youtube, a categorisation using rich meta-data will go a long way in enhancing the musical experience.

References

[AB15] Arora, V. and Behera, L. “Multiple F0 estimation and source clustering of polyphonic music audio using PLCA and HMRFs”. In: *IEEE Transactions on Audio, Speech, and Language Processing (TASLP)* vol. 23, no. 2 (2015), pp. 278–287.

[Bel+12] Bellur, A. et al. “A Knowledge Based Signal Processing Approach to Tonic Identification in Indian Classical Music”. In: *Proceedings of the 2nd CompMusic Workshop* (2012).

[Bel14] Bellur, A. “Automatic Tonic Identification In Indian Classical Music”. MA thesis. Indian Institute of Technology, Madras (IITM), 2014.

[BIM12] Bellur, A., Ishwar, V., and Murthy, H. “Motivic analysis and its relevance to raga identification in carnatic music”. In: *Proceedings of the 2nd CompMusic Workshop* (2012), pp. 153–157.

[BM13a] Bellur, A. and Murthy, H. “A cepstrum based approach for identifying tonic pitch in Indian classical music”. In: *National Conference on Communications (NCC)* (2013), pp. 1–5.

[BM13b] Bellur, A. and Murthy, H. “A novel application of group delay function for identifying tonic in Carnatic music.” In: *European Signal Processing Conference (EUSIPCO)* (2013), pp. 1–5.

[Boe93] Boersma, P. “Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound”. In: *Proceedings of the institute of phonetic sciences* vol. 17 (1993), pp. 97–110.

[De 02] De Cheveigné A. & Kawahara, H. “Yin, a fundamental frequency estimator for speech and music.” In: *The Journal of the Acoustical Society of America (JASA)* vol. 111, no. 4 (2002), pp. 1917–1930.

[DM14a] Dutta, S. and Murthy, H. “A modified rough longest common subsequence algorithm for motif spotting in an alapana of carnatic music”. In: *National Conference on Communications (NCC)* (2014), pp. 1–6.

[DM14b] Dutta, S. and Murthy, H. “Discovering Typical Motifs of a Raga from One-Liners of Songs in Carnatic Music”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2014), pp. 397–402.

[Dor14] Doraiswamy, G. *Sound of Indian Music*. Berklee College of Music, 2014.

[DSM15] Dutta, S., Sekhar, P. V. K., and Murthy, H. “Raga Verification in Carnatic Music Using Longest Common Segment Set”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2015), pp. 605–611.

[GS19b] Garani, S. S. and Seshadri., H. “An algorithmic approach to South Indian classical music”. In: *Journal of Mathematics and Music* vol. 13, no. 2 (2019), pp. 107–134.

[GSS12] Gulati, S., Salamon, J., and Serra, X. “A two-stage approach for tonic identification in indian art music”. In: *Proceedings of the 2nd CompMusic Workshop* (2012), pp. 119–127.

[Gul+14b] Gulati, S. et al. “Melodic pattern extraction in large collections of music recordings using time series mining techniques”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2014).

[Gul+14c] Gulati, S. et al. “Mining melodic patterns in large audio collections of indian art music”. In: *IEEE International Conference on Signal-Image Technology and Internet-Based Systems* (2014), pp. 264–271.

[Gul+16a] Gulati, S. et al. “Discovering rāga motifs by characterizing communities in networks of melodic patterns”. In: *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)* (2016), pp. 286–290.

[Gul+16b] Gulati, S. et al. “Phrase-based rāga recognition using vector space modeling”. In: *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)* (2016), pp. 66–70.

[Gul+16c] Gulati, S. et al. “Time-delayed melody surfaces for rāga recognition”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2016), pp. 751–757.

[Gul12] Gulati, S. “A tonic identification approach for Indian art music”. MA thesis. Universitat Pompeu Fabra, 2012.

[Gul17] Gulati, S. “Computational approaches for melodic description in Indian art music corpora”. PhD thesis. Universitat Pompeu Fabra, 2017.

[HKD06] Herrera-Boyer, P., Klapuri, A., and Davy, M. “Automatic classification of pitched musical instrument sounds”. In: *Signal processing methods for music transcription* (2006), pp. 163–200.

[Ish+13] Ishwar, V. et al. “Motif Spotting in an Alapana in Carnatic Music”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2013), pp. 499–504.

[Ish14] Ishwar, V. “Pitch estimation of the predominant vocal melody from heterophonic music audio recordings”. MA thesis. Universitat Pompeu Fabra, Spain, 2014.

[Kal] Kalyanaraman, S. *http://www.shivkumar.org/music/*.

[KI10] Koduri, G. K. and Indurkhya, B. “A behavioral study of emotions in south indian classical music and its implications in music recommendation systems”. In: *ACM workshop on Social, adaptive and personalized multimedia interaction and access* (2010), pp. 55–60.

[KI12b] Krishna, T. and Ishwar, V. “Carnatic music: Svara, gamaka, motif and raga identity”. In: *Proceedings of the 2nd CompMusic Workshop* (2012).

[KMS11] Koduri, G. K., M. Miron Serrà Julià, J., and Serra, X. “Computational approaches for the understanding of melody in carnatic music”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2011).

[Kod+12b] Koduri, G. K. et al. “Rāga recognition based on pitch distribution methods”. In: *Journal of New Music Research (JNMR)* vol. 41, no. 4 (2012), pp. 337–350.

[Kod+14b] Koduri, G. K. et al. “Intonation analysis of rāgas in Carnatic music”. In: *Journal of New Music Research (JNMR)* vol. 43, no. 1 (2014), pp. 72–93.

[Kod14] Koduri, G. K. “Culture-aware approaches to modeling and description of intonation using multimodal data”. In: *nternational Conference on Knowledge Engineering and Knowledge Management* (2014), pp. 209–217.

[Kod16] Koduri, G. K. “Towards a multimodal knowledge base for Indian art music: A case study with melodic intonation”. PhD thesis. Universitat Pompeu Fabra, Spain, 2016.

[Kri03a] Krishnaswamy, A. “Application of pitch tracking to South Indian classical music”. In: *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)* (2003), p. 389.

[Kri03b] Krishnaswamy, A. “Pitch measurements versus perception of south indian classical music”. In: *Stockholm Music Acoustics Conference (SMAC)* (2003), pp. 627–630.

[Kri04a] Krishnaswamy, A. “Inflexions and microtonality in south indian classical music”. In: *Frontiers of Research on Speech and Music (FRSM)* (2004).

[KSS12a] Koduri, G. K., Serrà Julià, J., and Serra, X. “Characterization of intonation in carnatic music by parametrizing pitch histograms”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2012).

[MC19] Madhusudhan, S. T. and Chowdhary, G. “Deepsrgm-Sequence classification and ranking in Indian classical music with deep learning”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2019), pp. 533–540.

[MSu09] M.Subramanian. “Carnatic Music - Automatic Computer Synthesis of Gamakams”. In: *Journal of the Sangeet Natak Akademi (Academy),* vol. 43, no. 3 (2009), pp. 28–36.

[Pad16] Padmasundari. “Music information retrieval: Raga identification using machine learning”. MA thesis. Indian Institute of Technology, Madras (IITM), India, 2016.

[PM17] Padmasundari and Murthy, H. “Raga Identification using locality sensitive hashing”. In: *National Conference on Communications (NCC)*, no. 1–6 (2017).

[Pol+07] Poliner, G. et al. “Melody transcription from music audio: Approaches and evaluation”. In: *IEEE Transactions on Audio, Speech, and Language Processing (TASLP)* vol. 15, no. 4 (2007), pp. 1247–1256.

[Raj+17b] Raj, P. V. K. et al. “An Approach to Transcription of Varnams in Carnatic Music Using Hidden Markov Models”. In: *National Conference on Communications (NCC)* (2017), pp. 1–6.

[Ran+19b] Ranjani, H. et al. “A compact pitch and time representation for melodic contours in Indian art music”. In: *The Journal of the Acoustical Society of America (JASA)* vol. 145, no. 1 (2019), pp. 597–603.

[Ran20] Ranjani, H. G. “Computational Approaches to Structural Analysis of Indian Art Music Signals”. MA thesis. Indian Institute of Science (IISc), Bangalore, India, Jan. 2020.

[Rao+14] Rao, P. et al. “Melodic motivic analysis of indian music”. In: *Journal of New Music Research (JNMR)* vol. 43, no. 1 (2014), pp. 115–130.

[Rao95] Rao, T. K. G. *Compositions of Tyagaraja*. Chennai: Ganamandir, 1995.

[RAS11] Ranjani, H. G., Arthi, S., and Sreenivas, T. V. “Carnatic music analysis: Shadja, swara identification and raga verification in alapana using stochastic models”. In: *Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)* (2011), pp. 29–32.

[RM13] Rajan, R. and Murthy, H. “Group delay based melody mono pitch extraction from music”. In: *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)* (2013), pp. 186–190.

[RPS17a] Ranjani, H., Paramashivan, D., and Sreenivas, T. “Quantized Melodic Contours in Indian Art Music Perception: Application to Transcription”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2017), pp. 174–180.

[RPS19] Ranjani, H. G., Paramashivan, D., and Sreenivas, T. V. “Discovering structural similarities among rāgas in Indian Art Music: a computational approach”. In: *Sādhanā* vol. 44, no. 5 (2019), p. 120.

[RS17] Ranjani, H. G. and Sreenivas, T. V. “Raga identification using repetitive note patterns from prescriptive notations of Carnatic music”. In: *arXiv* (2017), p. 1711.

[RVV18] Ragesh Rajan, M., Vijayakumar, A., and Vijayasenan, D. “Prediction of aesthetic elements in Karnatic Music: A machine learning approach”. In: *Proceedings of Interspeech* (2018), pp. 2042–2046.

[S89] S., M. “Psychological constraints on form-bearing dimensions in music”. In: *Contemporary Music Review* vol. 4, no. 1 (1989), pp. 181–198.

[SA09] Shetty, S. and Achary, K. “Raga mining of Indian music by extracting arohana-avarohana pattern.” In: *nternational Journal of Recent Trends in Engineering* (2009), p. 362.

[Sam98a] Sambamoorthy, P. *South Indian Music*. The Karnatic Music Book Centre, Chennai, 1998.

[San17] Sankaran, S. “Segmentation of Music Items: A Computational Perspective”. MA thesis. Indian Institute of Technology, Madras (IITM), India, 2017.

[SG09] Sridhar, R. and Geetha, T. “Raga identification of carnatic music for music information retrieval”. In: *nternational Journal of recent trends in Engineering,* (2009), p. 571.

[SG12a] Salamon, J. and Gómez, E. “Melody extraction from polyphonic music signals using pitch contour characteristics”. In: *IEEE Transactions on Audio, Speech, and Language Processing (TASLP)* vol. 20, no. 6 (2012), pp. 1759–1770.

[SGS12] Salamon, J., Gulati, S., and Serra, X. “A multipitch approach to tonic identification in Indian classical music”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2012), pp. 499–504.

[SKX16] Sentürk, S., Koduri, G. K., and Xavier, S. “A score-informed computational description of svaras using a statistical model”. In: *Sound & Music Computing Conference* (2016), pp. 427–433.

[Sri04] Sriram, V. *Carnatic Summer: Lives of Twenty Great Exponents*. EastWest Books (Madras), 2004.

[SS14b] Srinivasamurthy, A. and Serra, X. “A supervised approach to hierarchical metrical cycle tracking from audio music recordings”. In: *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)* (2014), pp. 5217–5221.

[Sub02] Subramanium, M. “An Analysis of Gamakams using the computer”. In: *Sangeet Natak, Journal of the Sangeet Natak Akademi,* vol. 37 (2002), pp. 26–47.

[Sub12] Subramanian, M. “Generating Computer Music from Skeletal Notations for Carnatic Music Compositions”. In: *Proceedings of the 2nd CompMusic Workshop* (2012).

[Sub13] Subramanian, S. K. “Modelling gamakas of Carnatic music as a synthesizer for sparse prescriptive notation”. PhD thesis. National University of Singapore, Singapore, 2013.

[Sub99] Subramanium, M. “Synthesizing Carnatic Music with a Computer”. In: *Journal of the Sangeet Natak Akademi (Academy),* vol. 133 (1999), pp. 16–24.

[SWM11b] Subramanian, S. K., Wyse, L., and McGee, K. “Modeling speed doubling in carnatic music”. In: *International Computer Music Conference* (2011), pp. 478–485.

[SWM12] Subramanian, S. K., Wyse, L., and McGee, K. “A two-component representation for modeling gamakas of carnatic music”. In: *Proceedings of the 2nd CompMusic Workshop* (2012).

[VAM17b] Viraraghavan, V., Aravind, R., and Murthy, H. “A Statistical Analysis of Gamakas in Carnatic Music”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2017), pp. 243–249.

[Vij07a] Vijayakrishnan, K. G. *The grammar of Carnatic musi*. De Gruyter Mouton, 2007.

[Vir+20d] Viraraghavan, V. et al. “Data driven measurement of precision of components of pitch curves in Carnatic music”. In: *The Journal of the Acoustical Society of America (JASA)* vol. 147, no. 5 (2020), pp. 3657–3666.

[Vir+20e] Viraraghavan, V. et al. “State-based Transcription Of Components Of Carnatic Music”. In: *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)* (2020), pp. 811–815.

[VMA18] Viraraghavan, V., Murthy, H., and Aravind, R. “Precision of sung notes in Carnatic music”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2018), pp. 499–505.

[ZCY20] Zhang, W., Chen, Z., and Yin, F. “Multi-Pitch Estimation of Polyphonic Music Based on Pseudo Two-Dimensional Spectrum”. In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)* vol. 28 (2020), pp. 2095–2108.

Part II

The Second Part

Chapter 7

Unique aspects and sound synthesis of stringed Indian musical instruments

Gary P. Scavone[1], Dhiraj V. Patil[2], SR Mahadeva Prasanna[3]

[1]Computational Acoustic Modeling Laboratory, Center for Interdisciplinary Research in Music Media and Technology, Schulich School of Music, McGill University, Montreal, Quebec, H3A 1E3, Canada
[2]Department of Mechanical, Materials and Aerospace Engineering, Indian Institute of Technology Dharwad, Karnataka, 580011, India.
[3]Department of Electrical Engineering, Indian Institute of Technology Dharwad, Karnataka, 580011, India

7.1 Abstract

This article provides an acoustic review of a set of common Indian stringed musical instruments, including the Sarangi, Sarod, Sitar, Veena and Tanpura, with a focus on those aspects of their construction and performance that most distinguish them in comparison to Western stringed instruments such as the violin and guitar. Other than the sarangi, which is played by bowing, these are plucked-string instruments that are excited either using a pluctrum or the fingers. The structural features that most differentiate this group of musical instruments include the use of drone and sympathetic strings, fret designs that readily support pitch bending, and bridge shapes and mountings that produce more harmonically rich sounds.

Acoustic analyses that have been conducted on Indian stringed instruments are summarized in this article. These studies have primarily been focused on the Tanpura and the Sitar, and in particular on the bridge-string interactions that lead to time-varying timbral qualities that are strongly harmonic and spectrally dense. Relevant analyses concerning traditional Western music instruments that could be applied to Indian instruments are discussed, including aspects such as the effect of bridges being mounted on membranes and possible interactions between sympathetic or drone strings and primary played strings through bridge-body coupling.

Finally, approaches for computer modeling and sound synthesis of this group of instruments are outlined, including examples of refined though computationally expensive models of the tanpura string-bridge interaction, as well as physically-informed approaches that seek to capture the essential characteristics of the sound in efficient ways. Techniques to enable accurate body modeling based on transfer function measurements are described, as well as the steps required to properly account for coupling between strings on a common bridge or coupling between different bridges on the same instrument body. Directions for future acoustic analyses and sound synthesis research are discussed.

7.2 Introduction

Indian music has many unique characteristics, including its form, melodic contours, and harmonies, to name a few. But the instruments used in creating Indian music also contribute to the music's distinctiveness. In comparison with Western music, there is significantly less concern with equal-tempered tuning, a strong emphasis on drone-like harmonies and melodic contours that tend to "slide" in pitch from one note to another.

One might be led to ask whether the instruments were adapted to the music and/or the music evolved in response to the particularities of the instruments.

Compared to the standard Western orchestral instruments, the most unique Indian instruments are the various plucked- and bowed-string instruments, as well as the tuned drums. Instruments such as the Murali, a transverse flute made of bamboo, and the Sahnai, a double-reed conical instrument, are more distinguished in terms of playing style, rather than construction, from similar Western counterparts.

The tuned drums include the Mrdangam, Pakhavaj and the more well-known Tabla and each makes use of a paste carefully applied to a stretched membrane surface to tune several of the membrane vibrational modes to harmonic or nearly-harmonic intervals. This also enables a wider variation of the sound characteristics when striking at different locations across the drum surface, sometimes achieving a metallic-like quality. There have been a number of previous research investigations into Indian drums, including analyses by the famous Indian physicist C. V. Raman [Ram35; RK20], as well as more recent investigations on the Mridangam [MP20], Idakka [KG18], Tabla [TG17], and more generally, bifacial drums [GSG19].

7.3 Common Indian Stringed Instruments

Figure 7.1: A sarangi (source: web).

This paper is focused on stringed instruments, including the Sarangi, Sarod, Sitar, Veena and Tanpura. Each of these instruments has unique structural and sounding qualities but they also share certain features that distinguish them from Western stringed instruments. The violin is widely used in Indian music but it has already been the subject of extensive analysis by acousticians.

7.3.1 The Sarangi

The sarangi, shown in Fig. 7.1, is excited by bowing. Its body consists of three hollow chambers carved from a single block of cedar, typically measuring about 2 feet long and 6 inches wide. The lower chamber is covered with goat skin, on top of which sits a bridge over which the three main gut strings pass, as well as an additional 35-37 sympathetic steel or brass strings. Only the three main strings are bowed. The sympathetic strings resonate in response to vibrational coupling through the bridge and body.

The most unique aspects of the sarangi, particularly in comparison to the Western viols, are the sympathetic strings and the bridge supported on top of a membrane. That said, there are many examples of instruments in other cultures with sympathetic strings, such as the Norwegian hardanger fiddle, as well as instruments such as the Chinese sanxian and American banjo that have bridges sitting atop membranes.

While there has been a large quantity of scientific research devoted to bowed stringed instruments that would also be relevant to the sarangi [MWS16b; Woo14], no acoustic analyses specific to the sarangi are known to these authors.

7.3.2 The Sarod

Figure 7.2: A sarod, with zoomed views of important parts.

The sarod is a plucked string instrument consisting of 4 or 5 main melody strings which run from the bottom of the resonator, over the bridge and are affixed to wooden tuning pegs (also referred to as 'kunti') at the top of the fingerboard. A few artists use 6 or 8 main strings, e.g. the Indian classical sarod player Ustad Amjad Ali Khan. The sarod also typically has 1 or 2 drone strings, two "chikari" strings and 9–11 (or even up to fifteen) sympathetic strings. The strings are generally steel or bronze. Like the sarangi, it has a skin-covered resonator that supports the bridge. The body is made of teak and the unfretted fingerboard is made of stainless steel, either polished or plated with chrome or nickel. The sarod is thus the only Indian string instrument that makes use of skin, metal, and wood. The strings are excited with a pluctrum and stopped on the fingerboard either with the edge of a finger nail or a combination of the nail and finger.

While some research has been presented specific to the sound of "chikari" strings [Sin13], no acoustic or structural analyses of the instrument itself have been found. The sarod is believed to have evolved from the Afghan rabab or the veena.

7.3.3 The Sitar

The sitar is one of the most widely known Indian instruments, due in part to (limited) use by a number of Western popular music groups in the 1960s and 1970s. The instrument is plucked using a pluctrum (as shown in Fig. 7.3), with 6–7 played strings and 12–13 sympathetic strings. It is common that the main or 'chikari' string(s) ('ma') be of steel and the 'Jora' or second string ('sa') made of phosphor bronze. The played strings pass over curved, moveable frets, while the sympathetic strings run below the frets. The body is typically made of teak or mahogony for the neck and faceplate, while the resonating chambers on either end are calabash gourds. The instrument may have upper 'tumba' made of gourds, which are used to sustain the base sound. There are two bridges, a larger one for the played strings and a smaller one for the sympathetic strings. A unique feature of the sitar, as well as the veena and tanpura, is that the bridge is curved and the strings effectively change length over time as they "wrap" and "unwrap" smoothly along the bridge over the course of a single

Figure 7.3: A sitar, with zoomed views of important parts.

period of oscillation (see Fig. 7.6 for an illustration). This moving boundary introduces nonlinearities that accentuate the harmonic content of the sound. Another distinctive aspect of sitar performance involves tension modulation, which is especially facilitated by the shape of the frets.

There have been quite a few analyses of the sitar string-bridge interaction, going as far back as [Ram21] (see [MW15] for a fairly complete list of references). More generally, there have been a number of recent papers concerning the numerical simulation of strings colliding with fretboards or bridges [BT14; BTC12; BW17; KS03; VBM09].

7.3.4 The Veena

Figure 7.4: A veena (source: web).

The veena is considered to be a precursor to the sitar. It features a hollow wooden body made of jack wood with two gourd resonators, one at each end. There are four main melodic strings and three auxiliary drone strings. The design of the frets is a bit different from the sitar but the playing technique also makes use of a significant amount of tension modulation by pulling a string across a fret, however, the pulling is not sustained in the veena. Like the sitar, the bridge is curved.

In addition to the research investigations referred to earlier with respect to the string-bridge interaction, some acoustic analyses of the veena have been conducted by [Ram21], [Sun+16] and [CSV21].

Figure 7.5: A tanpura, with zoomed views of important parts.

7.3.5 The Tanpura

The tanpura is a long-necked plucked string instrument that is used to accompany other melodic instruments, providing a harmonically-rich set of drone pitches. There are four or five metal strings which are plucked consecutively in a regular pattern using the finger tips. The strings are plucked at their natural lengths and thus, there are no frets. The body is made of a light, hollow wood with a resonator made either from a gourd or fabricated from wood.

The tanpura has a curved bridge like that of the veena and sitar. Among the Indian stringed instruments, however, the tanpura is unique in that it also makes use of a thin cotton thread, the "jivari" (also referred to as "javaari," "jiwari" and "juari" in different texts), placed between the string and the bridge. This imposes an additional modulation in the effective string length, creating a more abrupt change in the free length of the string with a resulting "buzzing," overtone-rich quality to the sound. Without "jivari" the pitch heard is generally near to the string fundamental and the partials are seen to fall in magnitude quickly. However, with "jivari" the pitch is heard at one/two octaves above the base pitch and higher partials have great power and modulate over time [CJV88]. The acoustics of the tanpura have been considered in depth by [Dat+19]. The string-bridge interaction has been studied by several other researchers, including [PG18; Ram21; Sen+83; Val+91].

7.4 Distinctive Instrument Features

The distinctiveness of the Indian string instruments highlighted in the previous section can be roughly grouped in the following three ways: 1. All include extra, non-melodic drone or resonant strings to provide harmonic support; 2. Other than the tanpura, all support pitch bending or modulation within and between notes; and 3. Several have bridge structures that lead to harmonically-rich sounds in comparison to Western plucked strings.

7.4.1 Drone and Sympathetic Strings

All of the instruments briefly discussed in the previous section include drone and/or sympathetic strings.

Sympathetic strings are tuned to particular pitches and are indirectly excited by the transfer of vibrational energy from plucked or bowed strings through common structural supports (the bridge or other end termination). Any harmonic resonances of a sympathetic string can be excited through coupling, though the strongest components will normally be the fundamental and first octave. Thin, light-weight metal strings are typically used because they have longer decay times and more easily couple with body vibrations because of their lower impedance. There are many examples of the use of sympathetic strings in instruments from various cultures. Indeed, even the piano supports the coupling of string vibrations into multiple other strings when the damping pedal is raised. Likewise, sympathetic vibrations are possible in other multi-string Western instruments such as the guitar, lutes and harps. That said, the style of music most typically performed on these Western instruments does not make significant use of sympathetic vibrations. Further, one might argue that sympathetic strings are more common with instruments played alone or with just a few other instruments, as the sympathetic vibrations provide a harmonic basis over which melodic material can be improvised. In larger ensemble contexts, these sympathetic vibrations, which are generally of lower amplitude, would likely be harder to hear and might not fuse as well.

In contrast with sympathetic strings (which are only indirectly excited), drone strings are directly excited, normally through plucking or strumming. Some examples of Western instruments making use of drones include bagpipes, the hurdy-gurdy and the banjo, which are most often used in folk-like contexts. Drone strings are inherently louder than sympathetic strings because they are directly excited and they tend to sound a smaller set of pitches than sympathetic strings (though their sounds may be harmonically rich).

7.4.2 Pitch Modulation

The bending or sliding of pitch is a quintessential aspect of the music played on Indian stringed instruments. This continuous frequency modulation can be achieved in two different ways. For instruments with frets, bending is produced by stretching of the string, which increases its tension and thus the speed of wave propagation (wave speed is proportional to the square root of tension divided by mass density). Tension modulation always produces frequencies above the string fundamental in its rest position (given the fixed nature of the string terminations). Because pitch bending is such an integral aspect of Indian music, the frets used on many of these instruments are designed in a way that accentuates the ability to stretch the strings during playing. For example, the sitar in particular has rather large, curved frets that allow significant displacement of the string away from rest position. On the other hand, the sarod and sarangi have no frets, and thus pitch bending is achieved by a continuous motion of the finger tip along the string. Fret-less instruments support pitch modulations in any direction, whereas the fretted instruments can only bend notes above the rest-position frequency. Some pitch modulation is common in Western instruments, particularly to produce vibrato, though the effect is taken to a much greater extreme in Indian music.

Though not discussed in this paper, the fretless Western violin is also commonly used in Indian music and one might argue the reason it has gained acceptance in this tradition is in part due to its inherent flexibility of pitch.

7.4.3 Bridge Curvature, Jivari and Mounting

The aspect of Indian stringed instruments that has received the most interest among acousticians is the curvature of the bridge and the use of jivari. As noted by Raman (1921), the sounds produced by the strings of the sitar, veena and tanpura do not exhibit the expected behaviour for stretched strings, whereby a string plucked at a specific point along its length will lack frequency components having a displacement minima (or node) at that point. For example, a stretched string plucked or struck 1/4 of the distance from a termination will lack energy at the 4th, 8th, 12th, … partials. Raman discussed a simple experiment in which he noted “On trying the same experiment with the ‘Veena’ or the ‘Tanpura’, it will be found that the overtone having a node at the plucked point sings out powerfully. In fact the position of the plucked point hardly appears to make a difference in regard to the intensity of the overtones in the ‘Tanpura’.”

As previously mentioned, the veena, sitar and tanpura all have curved bridges, as illustrated in Fig. 7.6. This is in contrast to the bridges of most Western string instruments, which have sharp edges such that the point of

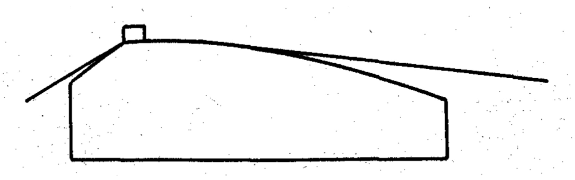


Figure 7.6: Illustration of the bridge of the sitar, veena and tanpura (from [Ram21]).

contact between the string and bridge remains relatively constant. With a curved bridge, the interaction between string and bridge changes over time depending on the vibrational phase and amplitude of the string. Thus, the effective length of the string is modulated both over the course of a single period of oscillation and also more gradually over time as the string energy decays and its displacement amplitude decreases. This also produces a modulation of the forces that are transmitted through the bridge and into the air. This moving boundary problem (as the string wraps and unwraps itself over the curved boundary) has been investigated by [MW15] and others. There is general agreement that the nonlinearity introduced by the curved bridge contributes to a transfer of energy to higher string partials, though the analysis of [MW15] was not able to explain the improved harmonicity of these instruments compared to more normally terminated plucked strings (where dispersion results in some inharmonicity). The curved bridge is also agreed to produce constant phase fluctuations between the fundamental and harmonic components over time [Mod70].

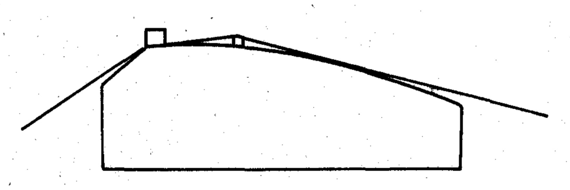


Figure 7.7: Illustration of the bridge of the tanpura (from [Ram21]).

The tanpura, in contrast to the veena and sitar, makes use of a thin thread (the jivari) inserted between the string and the curved bridge (see Fig. 7.7). Thus, rather than wrapping and unwrapping over the length of the curved bridge in a generally smooth way as discussed for the veena and sitar, instead the string-bridge interaction is more impulsive in nature and produces a sound often referred to as "buzz-like." As well, the strength of the frequency components fluctuates significantly over the course of a given pluck. [Dat+19] note a general trend of higher fluctuation rates with increasing harmonic number and varying decay rates for different harmonic components.

Another interesting aspect of the bridges in several Indian instruments is that they are mounted on membranes, rather than on more solid surfaces. This is similar to the case of the Western banjo, which has not garnered a significant level of interest among acousticians [Rae10]. However, recent investigations [Woo21] point out that because of the lower impedance seen by the strings at the bridge, the strings couple more strongly to the body. This results in a loud initial sound that decays relatively quickly and a larger contribution of body modes in the initial sound of each pluck event when compared to an instrument such as the guitar. As well, the banjo player can easily make adjustments to several key aspects of the instrument, including the membrane tension and the bridge mass, and these have important contributions to the sound quality. These conclusions are likely equally valid for the sarod and probably have some influence in the sarangi as well.

7.5 Modeling and Sound Synthesis Considerations

Computer modeling of sound producing vibrating structures is a fairly common pursuit in modern times that can be used both as a means of analysis of the physics involved, as well as for sound synthesis. Generally speaking, models intended for sound synthesis and in particular, real-time sound synthesis with parametric control, often need to be simplified through various approximations in order to be computationally efficient. That said, improvements in computational processing power has enabled real-time synthesis systems with very high levels of accuracy.

7.5.1 Body Modeling

The modeling of plucked string instruments has achieved very high levels of refinement in recent years [MWS16a; Woo04]. The instrument body is normally assumed to be fixed in shape and modeled as linear and time-invariant. This assumption might not be completely accurate for bridges mounted on membranes, though motion of the bridge can probably be ignored to first approximation. Refinements have included multiple dimensions of wave propagation along the string (two transverse polarisations, as well as longitudinal and torsional motion), string bending stiffness, more accurate damping properties of strings, and more accurate coupling characteristics through the bridge. Refinements in bowed string modeling have included all the former aspects and in addition, representations for the modal properties of the bow stick, bow hair and most importantly, the very complex bow-string interaction and frictional characteristics.

The bodies of the Indian stringed instruments considered here tend to be more bulky than Western viols, which have been refined over centuries to maximize sound projection and particular timbral qualities. On the other hand, the Indian string instruments tend to be used in small ensemble contexts and, at least in modern times, are often amplified. Also, because these instruments are made by hand and likely in lower quantities, we may expect that there is more variation between instruments. One might conjecture that there is perhaps less emphasis placed on the vibrational characteristics of the instrument bodies than for Western viols and instead, the refinement of Indian instruments is more focused on adjustments to bridges, frets and membranes to support coupling with sympathetic strings.

A high-quality approach to body modeling of string instruments involves measuring the input admittance and radiation transfer functions of an instrument. A common technique to measure input admittance involves damping the strings, tapping at the bridge in the plane of predominant string excitation and in proximity to the strings using a miniature impact hammer, and measuring the resulting velocity at the same location using a laser-doppler vibrometer or a low-mass accelerometer. For sound radiation transfer functions, the response to the impact hammer excitation can be measured using one or more pressure transducers located at some distance away from the instrument. These measurements provide linear representations of the body response as seen by the strings, as well as sound radiation characteristics when using external pressure transducers. From the measurements, digital filters can be designing to reproduce the body response when combined with a string model [MSS17; MWS16a; Woo04]. The approach of [MSS17] is particularly efficient in the context of digital waveguide synthesis [Smi92] because it allows joint realization of bridge reflectance and sound radiativity using a single shared bank of resonant filters.

One complicating feature of some Indian string instruments is the existence of multiple bridges and perhaps less common mechanisms for the transfer and coupling of energy between strings. That said, transfer function measurements and corresponding digital filter structures should be able to reproduce such effects with good quality.

7.5.2 String Models and Collisions

There has been a long history of digital modeling of strings, with common current approaches including digital waveguide, finite difference and modal schemes. Complications relevant to this paper involve collisions of the string with fixed structures, such as a curved bridge and/or frets. There has been considerable interest over the last 10 years or so in the modeling of collisions in musical instruments, such as a clarinet reed beating against a mouthpiece lay, the piano hammer striking a string, and string interactions with fixed boundaries [BT14; BTC12; BW17; KS03]. Much of the emphasis in these investigations is on numerical stability and energy conservation, rather than analyses of the particularities of the string-bridge interaction with respect to geometry modifications and/or resultant qualities of the sounds produced in comparison with measurements. The model of [BW17] compares the influence of tension modulation and a point- versus distributed- bridge model on the resulting time-varying 'jvari' sound. The computations involve the solution of a nonlinear vector equation at each time step. The combination of tension modulation and a distributed bridge is reported to produce the "liveliest" sound and spectrograms demonstrate a clear high-frequency formant structure that evolves after an initial pluck. The online sound examples (see [BW17]) demonstrate high-quality reproduction of the essential/unique aspects of the sound of the tanpura.

7.5.3 Physically-Informed Synthesis Approaches

As previously mentioned, realtime physics-based sound synthesis approaches often require simplifications, sometimes referred to as "physically-informed" modeling [Coo97], in order to achieve computational efficiency. The tanpura model of [BW17] accurately accounts for tension modulation and collision modeling but takes roughly 131 seconds to compute 1 second of sound. An efficient, physically-informed approach to the synthesis of the tanpura sound is represented by the `Drone` C++ class by Cook included in the "Ragamatic" project in the Synthesis ToolKit [CS99]. The string model itself is a simple 1D, extremely efficient digital waveguide structure, as diagrammed in Fig. 7.8. The upper and lower delay lines simulate displacement traveling wave propagation in two directions along the string. The right end of the string is assumed to be rigidly terminated and implemented with a reflection coefficient of -1. At the left end, a loop gain g_l is applied, then a loop reflectance filter $H_l(z)$ that commutes losses and reflection properties at the bridge, followed by an allpass fractional delay filter $H_a(z)$ to properly account for non-integer delay lengths. This is combined with an excitation signal $x(0,nT)$ and fed back into the upper delay line. The output $y(\xi,nT)$ is taken at an arbitrary pluck position ξ. To achieve the characteristic "sizzle-like" quality, the model is excited with a long duration enveloped noise sequence that spans many periods of oscillation, which is a variation of the well-known Karplus-Strong algorithm [JS83; KS83]. This simulates the non-periodic collision behaviour between the string and bridge and produces a high-bandwidth, time-varying energy profile. The long decay time of the drone is implemented with a loop gain very close to one, such that the string "rings" for a long time. An implementation of this model is provided in the `drone.m` and `dronetest.m` Matlab scripts shown in Appendix A. The tanpura bridge response applied in the `dronetest.m` script is available from [WBM16].

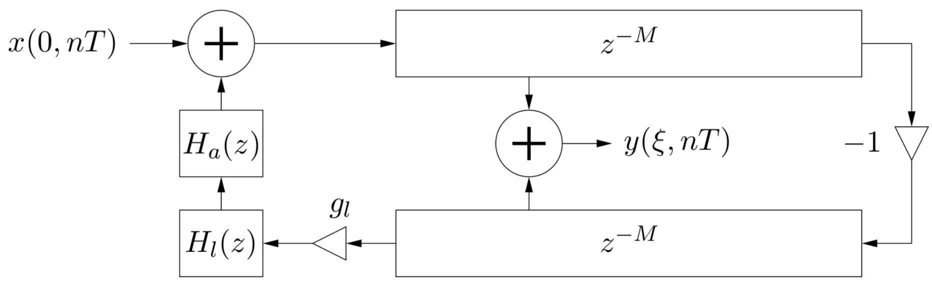


Figure 7.8: Digital waveguide model of a string.

Tension modulation is also easily and efficiently incorporated into a physically-informed approach by quickly varying the effective string length at the beginning of a pluck onset. This is demonstrated in the `sitar.m` script in Appendix A, which is based on the `Sitar` C++ class in the Synthesis ToolKit. In this example, the two delay lines shown in Fig. 7.8 are combined into a single delayline and the nut reflectance is commuted with the loop filter $H_l(z)$ to simplify the dynamic variation of the delay line length over time. An initial delay line length is randomly set above or below the target value and during each step of the iterative calculations, the current length exponentially progresses toward the target length. The excitation signal in this case is significantly shorter than for the tanpura.

7.6 Future Directions

It is clear that Indian stringed music instruments have generally received significantly less attention in terms of acoustical analysis and modeling compared to Western counterparts. In a way, this is good news for the community, as there remain a wealth of incredibly interesting instruments to investigate in the coming years. And while there are many unique aspects to these instruments, the methodologies used to measure, analyze and synthesize Western instruments can likely be applied with similar success.

7.7 Acknowledgements

We would like to thank Mr. Vijaykumar Patil (vocalist), Mr. Praveen Hugar (Sitarist), and Mr. Mallesh Hugar (Sarod and Tabla artist) from the City of Dharwad, Karnataka for allowing us to take photographs of the musical instruments. We would also like to thank ‘The Scheme for Promotion of Academic and Research Collaboration (SPARC)’ program of the Government of India which enabled the visit of Prof. Gary Scavone to the Indian Institute of Technology Dharwad, India.

7.8 Drone Matlab Scripts

This section has to be moved to the Appendix at the end.

7.8.1 drone.m

```
function y = drone( duration, f0, fs, pluckPoint )
% DRONE: 'duration' is in seconds, 'fs' is sample rate in samples per
% second, 'f0' is the fundamental frequency in Hertz to be synthesized
% and 'pluckPoint' is the position along the string where output is
% taken from.
%
% This Matlab function creates a simple waveguide plucked string drone
% model, similar to that implemented by Perry R. Cook in the Synthesis
% Toolkit in C++. No body filtering is performed
% here but the pluck position of the string is accounted for.
%
% by Gary Scavone, McGill University, 2021.

N = round(duration*fs);        % samples to compute
delta = mod(fs./f0-0.5, 2);    % fractional delay (compensate FIR 1/2 delay)
a = (1 - delta) / (1 + delta); % fractional delay allpass coefficient
D = (fs./f0 - delta - 0.5)/2;  % length of each delayline in samples
loopGain = min( 0.997 + f0 * 0.0000002, 0.99999 );

dlines = zeros(2, D); % initialize delay lines
y = zeros(1, N);      % initialize output vector

% Noise input
bpx = [0 2000 2500 duration*1000];   % time breakpoints in milliseconds
```

```
bpy = [0.0 1.0 0.0 0.0];             % corresponding breakpoint amplitudes
env = interp1(bpx, bpy, 1000*(0:N-1)/fs);
x = 0.005 * env.*(2 * rand(1, N) - 1);

ptr = 1;
uptr = round((1-pluckPoint)*(D-1))+1;
dptr = D-uptr+1;
loopz = 0; % FIR filter state
apz = 0;   % allpass filter state

for n = 1:N
  temp = loopGain * dlines(2, ptr); % scaled output from lower delay
  y(n) = dlines(1, uptr) + dlines(2, dptr); % output at pluck point

  dlines(2, ptr) = -dlines(1, ptr); % rigid reflection at nut

  dlines(1, ptr) = -0.5*temp - 0.5*loopz; % 1st-order FIR loop filter
  loopz = temp;
   % fractional delay filter
  [dlines(1, ptr), apz] = filter([a 1], [1 a], dlines(1, ptr), apz);

  dlines(1, ptr) = dlines(1, ptr) + x(n); % input to upper delay

  % Increment pointers & check limits
  ptr = mod(ptr, D) + 1;
  uptr = mod(uptr, D) + 1;
  dptr = mod(dptr, D) + 1;
end
end
```

7.8.2 **dronetest.m**

```
% dronetest.m
%
% This Matlab script creates a simple looping sequence of 3 plucked
% string notes that might be played on a drone-like instrument.
%
% by Gary P. Scavone, Mcgill University, 2021.

% Synthesis parameters
fs = 44100;               % sampling rate
duration = 10.0;          % seconds
f0 = 65*2.^([7 12 0]/12); % note frequencies in hertz
doBody = true;            % apply body filter
pluckPoint = 0.1;         % normalized pluck position (0-1)
nCycles = 1;              % number of times to repeat sequence
spb = 1.2;                % seconds per beat
```

```
% Synthesize separate signals for each note
y1 = drone( duration, f0(1), fs, pluckPoint );
y2 = drone( duration, f0(2), fs, pluckPoint );
y3 = drone( duration, f0(3), fs, pluckPoint );

% Overlap note signals into a sequence
beatSamples = round(spb * fs);
zs = zeros(1, beatSamples);
zs4 = zeros(1, 4*beatSamples);
preZs = [];
postZs = repmat( zs4, 1, nCycles-1 );
y = [preZs y1 zs zs zs postZs] + [preZs zs y2 zs zs postZs] + ...
    [preZs zs zs y3 zs postZs];
for n = 1:nCycles
  preZs = repmat( zs4, 1, n-1 );
  postZs = repmat( zs4, 1, nCycles-n );
  y = y + [preZs y1 zs zs zs postZs] + [preZs zs y2 zs zs postZs] + ...
      [preZs zs zs y3 zs postZs];
end

% Scale output if necessary
if max(abs(y)) > 0.95
  y = y./(max(abs(y))+0.1);
end

if doBody % Post-apply body filter
  [ir, fs] = audioread('tanpurabridgeir.wav');
  y = filter(ir, 1, y);
end

% Play result
sound(y, fs);
```

7.8.3 sitar.m

```
function y = sitar( duration, f0, fs )
% SITAR: 'duration' is in seconds, 'fs' is sample rate in samples per
% second and 'f0' is the fundamental frequency in Hertz to be synthesized.
%
% This Matlab function creates a simple waveguide plucked string drone
% model, similar to that implemented by Perry R. Cook in the Synthesis
% Toolkit in C++. No body filtering is performed.
%
% by Gary Scavone, McGill University, 2021.

N = round(duration*fs);               % samples to compute

targetD = fs./f0;  % length of single delayline in samples
```

```
currentD = targetD * (1.0 + 0.1 * (2*rand(1, 1)-1));
M = floor(currentD);
delta = currentD - M; % fractional part
a = (1 - delta) / (1 + delta); % fractional delay allpass coefficient
loopGain = min( 0.995 + f0 * 0.005, 0.9995 );

dlines = zeros(1, round(max(targetD, currentD))); % initialize delay line
y = zeros(1, N);      % initialize output vector

% Noise input
bpx = [0 1 41 duration*1000];   % time breakpoints in milliseconds
bpy = [0.0 0.1 0.0 0.0];        % corresponding breakpoint amplitudes
env = interp1(bpx, bpy, 1000*(0:N-1)/fs);
x = env.*(2 * rand(1, N) - 1);

ptr = 1;
loopz = 0; % FIR filter state
apz = 0;   % allpass filter state

for n = 1:N
  temp = loopGain * dlines(1, ptr); % scaled output from delay

  dlines(1, ptr) = 0.9901*temp - 0.0099*loopz; % 1st-order FIR loop filter
  loopz = temp;
  % fractional delay filter
  [dlines(1, ptr), apz] = filter([a 1], [1 a], dlines(1, ptr), apz);
  y(n) = dlines(1, ptr);
  dlines(1, ptr) = dlines(1, ptr) + x(n); % input to upper delay

  % Update delay length
  if abs(currentD - targetD) > 0.001
    if targetD < currentD
      currentD = currentD * 0.99999;
    else
      currentD = currentD * 1.00001;
    end
    M = floor(currentD);
    delta = currentD - M;
    a = (1 - delta) / (1 + delta);
  end

  % Increment pointer & check limits
  ptr = mod(ptr, M) + 1;
end
end
```

7.8.4 sitartest.m

```
% sitartest.m
%
% This Matlab script creates a simple sequence of plucked
% string notes that might be played on a sitar-like instrument.
%
% by Gary P. Scavone, Mcgill University, 2021.

clear all;

% Synthesis parameters
fs = 44100;                   % sampling rate
duration = 1.0;               % note duration in seconds
f0 = 220*2.^([0 2 4 7 9]/12); % possible note frequencies in hertz
nNotes = 10;                  % number of randomly chosen notes
spb = 0.3;                    % seconds between notes

N = round(fs*(nNotes*spb + duration)); % total sound duration in samples
y = zeros(1, N);
rCounter = 0;  % rest counter
nCounter = 1;  % note counter
nDur = round(duration * fs); % note duration in samples
for n = 1:N
  if rCounter == 0 % play new note
    tmp = sitar( duration, f0(ceil(length(f0)*rand(1, 1))), fs );
    y(n:n+nDur-1) = y(n:n+nDur-1) + tmp;
    rCounter = round(spb * fs);
    nCounter = nCounter + 1;
    if nCounter > nNotes, break; end
  end
  rCounter = rCounter - 1;
end

% Scale output if necessary
if max(abs(y)) > 0.95
  y = y./(max(abs(y))+0.1);
end

% Play result
sound(y, fs);
```

Chapter 8

A Comparison of Percussion Instruments in Hindustani and Carnatic Music

Gowriprasad R[(a)] and Ajay Srinivasamurthy[(b,1)]

[(a)]Indian Institute of Technology Madras, Chennai, India
[(b)]Music Technology Group, Universitat Pompeu Fabra (UPF), Barcelona.

Abstract

Hindustani and Carnatic music are India's two predominant art music traditions. While these two music traditions have clear distinctions, in theory, performance, and sociocultural aspects, they also share many similarities. These similarities and distinctions extend to all musical concepts, including melody, rhythm, and timbre. This article focuses on the rhythm and percussion in the two music traditions, comparing and contrasting them across sociocultural, technical, and timbral dimensions. The sociocultural aspects include discussions around the origins, history, current practice, teaching, and pedagogy. Since mridangam and the tabla form the primary percussion instruments in Carnatic and Hindustani music, respectively, our timbral analysis focuses mainly on these instruments and their practice. We primarily focus on the percussion techniques of the mridangam and tabla in both accompaniment and solo performances, providing an overview of both generalities and nuances of percussion. The article is intended to help technology enthusiasts, engineers, and researchers build common and culture specific computational tools for analyzing percussion in Indian art music.

8.1 Introduction

India is culturally diverse due to its vast and varied geography, multiple languages, and multiple long-existing civilizations. The cultural diversity of India is also reflected in its well-established predominant art music cultures, *Hindustani* Music (HM) and *Carnatic* Music (CM) traditions. These two music traditions are culturally diverse and are influenced by geography, sociocultural practices, and languages (e.g., compositions in multiple different languages of India). Both music traditions are well studied with significant musicological literature and have a large audience. These music traditions have many similarities and distinctions in literature, teaching and learning, performance, and sociocultural aspects. These similarities and distinctions extend to all musical concepts, including melody, rhythm, and timbre. This article aims to provide an overview of the similarities and differences between HM and CM, focusing primarily on the percussion instruments used in these music cultures.

Percussion instruments are an integral part of Indian Art Music (IAM) traditions, where they serve multiple roles in music performance and practice. They are often the primary device for tracking rhythm in music performance, helping to maintain the tempo and track progression through different metrical structures. They also act as primary rhythmic accompaniments, helping with the exposition of rhythmic structures of the music piece. Fundamentally, percussion instruments are used as accompaniment instruments in vocal and instrumental music performances. However, there are solo performances of percussion instruments within a concert as well

[(1)]Ajay Srinivasamurthy is currently with Amazon Alexa, India and contributed to the work described in the chapter during his PhD at Music Technology Group, UPF, Barcelona before joining Amazon.

as percussion solo concerts. The properties of the instruments such as construction, timbre, sound, resonance characteristics, sustenance are evolved and highly influenced by the sociocultural aspects of the music tradition.

This article aims to provide cross-cultural analysis focusing on the tabla and mridangam and their practices, as these two are the primary percussion instruments in HM and CM. Such an effort to compare long-existing music cultures could be undertaken from multiple perspectives and for different audiences. However, within the scope of the article, our analysis is intended to help technology enthusiasts, engineers, and researchers build common and culture-specific computational tools to analyze percussion in Indian art music. The sociocultural aspects influence these percussion instruments' evolution, performance roles, and practices. The playing technique would depend on the roles and practices, and influences both the rhythm and timbre of the percussion music performance. We hence believe that our aims of the article to help build computational analysis tools for percussion need us to compare and contrast the percussion in both accompaniment and solo performances on the sociocultural, technical, and timbral dimensions. Given the broad scope of the subject, the article is only intended as a brief primer on the topics, and we cannot make any claims about being comprehensive and exhaustive in our analysis. However, we aim to provide a light introduction and comparative analysis that is useful to further computational analysis of percussion in IAM. An in-depth treatment of either sociocultural, technical, or timbral aspects of percussion is beyond the scope of this article.

The article is broadly divided into the following sections. We first describe the historical and sociocultural background of mridangam and tabla, followed by a description of the timbre of these percussion instruments. We then focus on the percussion techniques of mridangam and tabla as an accompaniment, followed by a section that elaborates on their solo performances. In each section, we first describe the concerned dimension of mridangam and tabla individually and then follow it up to compare and contrast the features across the two music cultures. The final section summarizes the article with key takeaways and discussion. A glossary is included for culture-specific terminology. We first start with a basic introduction to percussion instruments that also cover the sociocultural aspects.

8.2 Percussion in Indian art music

Mridangam, tabla, and pakhavaj (also called mridang) are the predominant percussion instruments used in Carnatic, Hindustani khayāl, and Dhrupad music cultures, respectively. They are called the tāla vādya, avanaddha vādya or membranophones - percussion instruments, since they have two stretched loaded membranes on which the strokes are played. Unlike most percussion instruments in western music, which are not tuned to any tonic, Indian percussion instruments are tuned to the specific tonic set by the lead artist in the performance.

Percussion instruments are fundamentally used as an accompaniment to keep track of rhythm. The properties of the instruments such as timbre, sound, resonance characteristics, sustenance are evolved and highly influenced by the very nature of the music traditions they accompany. The evolution of these instruments' repertoire and playing techniques is also based on the sociocultural identities of HM and CM. The sociocultural aspects are important for understanding the current performance roles, practices, techniques, and timbre that could be studied under multiple heads. While we describe the historical and music learning aspects in this section, we elaborate on the performance roles and timbre in separate sections that follow. A basic understanding of the cultural history of HM and CM and the evolution of the repertoire and playing techniques of mridangam and tabla helps us to understand the percussion in two traditions by putting them in their sociocultural contexts. However, given the long existence of IAM, many aspects of its cultural history are based on expert opinions and hence quite subjective. Several unsubstantiated claims exist, and hence we focus only on the implications of the history and not on the accuracy therein. This helps us focus on forward-looking computational musicology issues without academically relevant debates on the accuracy of opinions on historical evolution.

8.2.1 Cultural History

Indian art music is an integral part of the art and culture of the Indian sub-continent and carries a legacy of two millennia. The roots of Indian music are often attributed back to Bharata's nātya shāstra (dated approximately between 200 BCE and 200 CE), which deals with the basics of music, dance, and drama. The 12th century Sanskrit text sangīta-ratnākara of sārangadēva is regarded as the definitive text by both the Hindustani music and the Carnatic music traditions [SS+78]. It was later that the distinctions of Hindustani and Carnatic emerged. Hence, the essential foundational elements of IAM such as rāga, svara and tāla have continued to exist in both cultures, albeit with differences that have accumulated as the two music cultures evolved in their respective sociocultural contexts.

8.2.1.1 Tabla in Hindustani music

The foreign invasions into the Indian subcontinent during the 10th to 14th centuries are slated to have influenced music culture. These influences served as a prologue for the development of khayāl form of Hindustani music. It also started the vocal forms such as Thumri, Dadra, Ghazal, and Khilwada. Thus arose the necessity for a percussion instrument that is well suited for the very sweet and pleasant khayāl music, necessitating the invention of the tabla. There are many opinions on how the tabla was invented without a clear consensus [**tabla_Visharad**; Pra11]. Some leading hypotheses of its origin trace it to a Persian percussion instrument that is called ”tabla”, or that it was constructed by splitting the pakhavaj (or mridang, a double-barrel drum used to accompany Dhrupad music) into two. Few believe that Amir Khusro, a musician in the Delhi sultan court, invented tabla [**tabla_Vidwat**].

Tabla, in its current performance practice, has evolved to be the primary rhythm accompaniment that is capable of producing different stroke timbres that help to build rhythmic patterns for both functional time-keeping and aesthetic enhancement of khayāl music. Tabla has also been adapted to popular and light music in the Indian subcontinent (such as Bollywood music), along with the adaptation of other elements of Hindustani music [Pra11]. As Hindustani music evolved into different stylistic schools called the gharānā, tabla also evolved into its own set of gharānā. Tabla continues to evolve and is currently used as accompaniment in multiple forms of Hindustani music such as khayāl, thumri, dadra, bhajan, ghazal, while also used extensively in film and popular music.

8.2.1.2 Mridangam in Carnatic music

The foreign invasions in the Indian sub-continent had minimal effect on the south, and hence it is claimed that south India's music culture retained its ancient tradition. It is considered that Saint Purandara Dasa is the CM tradition's founder. The Maratha empire (16th to 18th century) contributed to the advancement of CM [Sun21]. But the CM developed extensively by the contribution of three musicians Shyama Shastri, Tyagaraja, and Muthuswami Dikshitar. They are called the triratnas (trinity) of CM. Their period stands from 1750 to 1850 CE. Most Carnatic music training currently revolves around kritis (compositions) composed by these great composers, who are revered as saints. It it believed that the pakhavaj ”mridang” has been brought to South India (Thanjavur) by Maratha kirtanakars during this period [Sun21]. Narayanaswami Appa, the founder of modern-day mridangam was a great kirtanakar. Thanjavur, under the Maratha regin, evolved as the fountainhead of mridangam [Sun21].

Mridangam is the primary rhythm accompaniment in Carnatic music, though Carnatic music has developed a set of additional rhythmic accompaniments such as ghaṭam (clay pot), khanjira (Indian tambourine) and morsing (Indian jaw harp). Mridangam is a combination of two words, ”mrit + anga = mridanga,” which essentially means the part made out of clay [PE05]. In recent centuries, the hollow cylinder made out of clay was replaced by wood for convenience. The tempo of music performances in Carnatic music is usually more restricted than Hindustani music, ranging between 60-120 beats per minute. The mridangam construction and playing techniques hence evolved to accompany the relatively faster rhythm movements. This made mridangam's strokes less resonant and sustained compared to the pakhavaj. Mridangam is relatively less adapted to popular music in South India, with its current use largely restricted to Carnatic music and popular music that is heavily influenced by Carnatic music.

Based on the historical facts of tabla and mridangam, it is evident that there are several similarities and parallel concepts in the percussion systems of Hindustani and Carnatic music. However, we wish to highlight what we consider and some primary distinctions between the two systems, which would help us build culture-specific tools or extend generic tools to these music cultures for percussion analysis.

8.2.2 Pedagogy in Indian percussion

Like many music cultures that follow an oral transmission of knowledge through speech, vocables, oral mnemonics, etc, the repertoire and playing techniques of mridangam and tabla are also transmitted with a system based on onomatopoeic oral mnemonic syllables [Pra11]. These mnemonic syllables are associated with individual strokes corresponding to different timbres that can be produced on the mridangam and tabla are called *solkaṭṭu* and *bōl*, respectively [Sun21]. The oral mnemonic syllables are a well-established representation of the timbres/strokes of mridangam and tabla. The initial lessons taught are called the *bālapātha*, the basics of fingering postures and strokes. The student is always made to recite the individual stroke patterns with timing and then to play on the instrument.

The system of representation for percussion strokes using oral mnemonic syllables provides several advantages in pedagogy. The strokes of the instruments can be perceived to be analogous to different letters in the alphabet of a language, allowing students to learn them as such. However, a single stroke (timbre) on the instrument is represented by several solkaṭṭu/bōl depending upon the context in which it is being played, and the syllable notation varies marginally across different schools of learning. These two aspects lead to a one-to-many mapping of stroke timbres to syllables, which adds to some additional learning challenges analogous to similar language learning challenges.

8.2.3 Schools of Indian percussion

The sophisticated vocabulary of the Indian percussion instruments has helped evolve a rich and varied repertoire for solo performances. The repertoire which exists today is vast and represents the cumulative contributions of many generations of players. The repertoire and technique are orally transmitted from generation to generation by guru-shishya (teacher-student) parampara [Bag98b]. The guru-shishya lineages exploring different innumerable aesthetic possibilities have given rise to distinct playing styles, loosely grouped into percussion schools.

Gottlieb [Got93] mentions three factors for comparing the similarities and differences in the playing style: (1) Sound production, that is, quality and the technique used, (2) Repertoires, and (3) Rhythmic practices. The technique and the rhythmic practices differ from artiste to artiste. Compositions bearing the stylistic school distinctions are based on the repertoires from each school. A few essential requirements for a style to be called a school are [Sun21]:

- Defined playing methodologies
- Standard characteristics
- Clear, distinct patterns
- Distinct playing techniques
- Large number of artists following the style

An effective description of the percussion schools needs the use of culture-specific terminology. A few of the culture-specific terminologies used in this section are defined in the glossary and described in greater detail in the subsequent sections.

8.2.3.1 Mridangam bāni

The terms bāni, schools, styles are often used interchangeably. But the differences can be thought of as follows. A bāni can be considered as a large tree with multiple branches. The various branches can be considered as multiple schools. If a player follows a certain school and adapts certain characteristics and features from another school or bāni and develops his own method of playing, it can be called a style. It is hard to get the number of bāni in mridangam playing as there are no written sources and documents that present a clear classification [Sun21]. The two standard and popular bāni are Thanjavur bāni and Pudukottai bāni. We can map and trace back the playing techniques of different mridangam maestros to these two bāni.

Thanjavur Narayanaswami Appa (19th century CE) is considered the originator of modern mridangam. Kumbakonam Azhaganambi Pillai (1863-1939) was also a very senior *vidvān* (maestro) who influenced many on the mridangam playing technique. The playing methodologies of Narayanaswami Appa evolved due to several years of collaboration and exchange of ideas, thus emerging in the two big bāni, Thanjavur and Pudukottai. Thanjavur Vaidyanatha Iyer (1897-1947) is considered the founder of Thanjavur bāni. A few well-known among his disciples are Palghat T S Mani Iyer, Padma Vibhushan Dr. Umayalapuram K Sivaraman, Dr. T K Murthy. There are two major schools under Thanjavur bāni - the Palghat Mani Iyer school, and Dr. T.K Murthy school [Sun21]. Pudukottai Manpoondia Pillai (1859-1922) is considered a torchbearer of Pudukottai bāni. His major disciples are Pudukottai Dakshinamoorthi Pillai, Palani Muthiah Pillai (father of Palani Subramanya Pillai), Ramanathapuram Chitsabai Servai (father of C S Muruga Bhoopathi). These three major disciples created their own school of playing with their expertise.

As mentioned earlier in the same section, leading masters under a particular school get influenced by other schools and develop their own playing methodology, resulting in a new style. Sometimes, the style created emerges as a big tree and is called a bāni by itself because of its uniqueness and its significant number of followers. Palghat Mani Iyer got highly influenced by Pudukottai bāni, especially by Dakshinamoorthi Pillai, and created his own style. This emerged as bāni by itself due to a large number of followers over the year. Another example is the case of Guru Karaikudi R Mani. He was first trained under Karaikudi Rangu Iyengar (Pudukottai bāni), and later from T R Harihara Sharma and K M Vaidyanathan (Thanjavur bāni). Further, he was highly influenced by the thavil (a horizontal barrel shaped twin-faced drum from South India, played with a stick and fingers) and came up with a unique style that later emerged as a bāni with many followers. These are just two examples of the many, hence the exact number of bāni-s is not documented clearly. However, it is more certain and well accepted that almost all successful masters such as Palghat R Raghu, Umayalapuram K Sivaraman, Trichy Shankaran have developed their own signature styles by getting influenced by other schools and also by their self improvisation and innovation [Sun21].

To summarize, the answer to the question ”Can one create a new bāni?” would be NO. But, one can have their own style. The style then needs to evolve and be accepted by many followers to emerge as a tree with lots of branches and leaves to organically become a bāni in due course.

8.2.3.2 Tabla Gharānā(s)

The tabla was popularized during the Delhi Sultanate’s period (14-16th century CE) and evolved to accompany the khayāl music. Hence, the playing techniques were initially kept simple with limited strokes and variations to aid and not disturb the leading artist. Later on, the tabla repertoire was developed over the years to portray tabla as an independent instrument with solo performances. Thus many compositions were created, increasing the repertoire considerably. Usage of different strokes and techniques evolved. This created different styles and lineages.

There are two major playing styles (*baj*) in tabla - *bandh* (closed) baj, and *khulā* (open) baj [Sax06]. In bandh baj, the tabla is played more on the drum face's outer border annular area, leading to controlled or subdued stroke resonance or sound. In the closed style of playing, importance is given to the sound of tabla and speedy progressions. In khulā baj, the tabla is played more in the central portion of the drum face with open strokes using full palm and fingers. In this style, the tonal richness of the strokes with resonance is given importance rather than speed. Based on these two playing styles, six stylistic schools of the tabla (called the gharānā) developed - *Dilli* (Delhi), *Ajrada*, *Lucknow* (Purab), *Farrukhabad*, *Banaras* and *Punjab*. The word gharānā literally means house and implies the house of the teacher [**tabla_Visharad**; Got93].

Shri Siddhar Khan Dadhi (approx. 1700 CE) is considered as the founder and promoter of Delhi gharānā (the oldest gharānā of tabla) [Sax06]. He modified and converted the pakhavaj bōl and played on tabla. Delhi gharānā baj is known as "*do ungli baj*", the two-finger style. Most of the strokes are played only with the middle finger and forefinger. Tabla is played on the border of the membrane in this gharānā. Hence it is also called "*kinare ki baj*", the corner style. Delhi gharānā is known for its compositions such as pēśkār, kāyadā, and rēla.

Ustad Kallukhan and Ustad Meerukhan, from the town Ajrada, learned tabla in Delhi, went back to their place, and started playing according to their zest and interest. Thus a new Ajrada gharānā emerged. This gharānā has a special way of playing the bass drum and many āḍi lay compositions are composed. It is also special for its compositions in tiśra naḍe. The usage of the ring finger is also seen in this gharānā. Delhi and Ajrada gharānā are the forms of bandh baj.

Tabla players from Delhi gharānā, such as Ustad Modu Khan and Bhakshu Khan, came to Lucknow and were influenced by the music culture. Pakhavaj and Kathak dance forms made an impact on their the style of playing tabla. Thus the Lucknow baj got separated from Delhi baj. Tabla is played more in the central area of the membrane accompanying dance. Most of the dance compositions like gaṭ are played on tabla. This is one of the forms of khulā baj.

Lucknow gharānā got extended and developed separately to form Farukhabad gharānā. This gharānā mainly showcases the Lucknow gharānā compositions' extension, also introducing tipalli and chaupallis [**tabla_Visharad**]. Farrukhabad repertoire exploited entire vocabulary of the instrument [Pra11]. The renowned tabla player Ustad Ahmad Jaan Thirakwa and Pandit Jnan Prakash Ghosh are from this gharānā. Most of the tabla players found in the current time belong to Farukhabad gharānā.

Pandit Ram Sahai learned tabla from Ustad Modu Khan and came to Banaras and developed Banaras gharānā. Unlike the Delhi gharānā, this gharānā makes use of all the fingers and palm while playing. Banaras gharānā compositions got greatly improved to play laggis and ladis for thumri [Got93]. The specialty of this gharānā is that the tabla solo performance usually starts with uṭhān and is also famous for the compositions like stuti paraṇ.

Apart from these above-mentioned gharānā, Punjab gharānā developed independently by the transfusion of pakhavaj bōl onto tabla by Ustad Fakir Bhaksh. Thus it had a different kind of development. This gharānā is known for its complex compositions, such as gaṭ-paraṇ, which are vivid rhythmic. All the tāl are extended and played, which is not found in any other gharānā. The famous tabla players like Ustad Alla Rakha, Ustad Zakir Hussain, and Pandit Yogesh Samsi are the Punjab gharānā exponents.

To summarize the stylistic schools of percussion in IAM, the gharānā of tabla have a clear parallel in Hindustani music gharānā, hence are more well defined. The lack of such a system in Carnatic music implies that though different styles of mridangam playing exist in practice, they are less well-defined without formally identified names.

8.3 Timbre of percussion instruments of IAM

Both the tabla and mridangam are loaded with stretched membranophones. Mridangam is a double-sided barrel-shaped single hollow drum, whereas the tabla consists of two drums with a single drum head each.

Hence the sound characteristics of these two instruments have many similarities and contrasts. The fundamental sounds of the percussion instruments are their strokes. In its simplest form, a sequence of strokes makes a percussion pattern.

As described earlier, the strokes of both mridangam and tabla and learned and represented through onomatopoeic oral mnemonic spoken syllables. The strokes are mapped to syllables essentially based on the timbre of the strokes. Hence, the instruments' timbre is an important factor to consider while developing the computational techniques to analyze the percussion patterns and styles. The system of representation for percussion strokes using oral mnemonic syllables provides multiple opportunities for computational research on percussion in IAM. It helps us have a musically meaningful representation of the percussion strokes played by capturing articulation, timbre, and dynamics subtleties. The syllable to stroke mapping is not exactly one-to-one and may vary while playing different combinations of patterns and contexts where the specific strokes are played. The same syllable is played slightly differently or even with different strokes and vice-versa. Due to this, the timbre is also influenced by the previous and the following strokes as well, with the context playing an important role. This gives a clear analogy to speech, due to which we could possibly extend the speech signal processing methods for the percussion audio analysis [Sri+17a].

Unlike the common percussion instruments in western music, which are not tuned to any tonic, tabla and mridangam are tuned to a specific tonic set by the main artist in the performance [Cla08]. The load and the tension of the membrane are helpful in tuning. Typically, the tabla tuning will be one octave higher than the mridangam. Due to this, mridangam will have a much bolder sound compared to tabla [CC]. We can classify almost all the strokes played on either side of the mridangam and tabla as (1) open and closed strokes, (2) resonant and damped strokes, (3) individual or compound strokes.

The left (bass) drum head typically has one closed stroke played with full palm, one open, and one stroke with modulation effect played with fingers. The left drum of the tabla has more surface area, and it is a bit more comfortable to play the modulating stroke compared to the mridangam. The right drum head has a few open strokes played on the border area, closed strokes in the center, which are damped, and open strokes in the middle and slightly away from the center that have rich resonance characteristics. The left and right drums are naturally interchanged when being played by left-handed musicians. Few strokes often seen in tabla like "Dhi Re Dhi Re" played with full palm via right hand are seen rarely in mridangam. Tabla has a very pronounced fundamental (harmonic) and a long sustain. The sustenance duration of the mridangam strokes is less than tabla's and has a much more complex harmonic spectrum [CC]. The strokes played on one drum head affect the other drum head in the mridangam due to the single resonance column [Bha91], whereas it is not the case with tabla. This coupling leads to some differences in playing technique, where one drum head is intentionally damped while the strokes from the other drum head need to be emphasized.

8.4 Percussion technique in IAM

The artists' *manodharma* (improvisation), nuances, and skills are essentially showcased based on their approach and techniques. These percussion techniques are useful to understand the underlying metrical structure of the tāla and its components such as vibhāg, forming an essential part of rhythm and percussion pattern analysis. The percussion patterns are very much necessary while analyzing the stroke accents, improvisations, nuances, etc.

Both tabla and mridangam developed as accompanying instruments but also established themselves as solo instruments. We have four major idioms of percussion instruments - (1) accompanying vocal/instrumental music, (2) accompanying dance forms, (3) solo performance, (4) percussion ensemble, typically called a *tāla vādya kacēri* (concert). Hence it is useful to analyze the instruments' percussion techniques with respect to these aspects individually, then compare and contrast their functionality.

8.4.1 Accompaniment to Vocal Music

The percussion instruments originally evolved to accompany vocal music. Both tabla and mridangam act as the primary rhythm accompaniment in IAM performances, where their accompaniment role is to enhance the experience with music with rhythmic exposition while also playing a role as time-keepers of the metrical structures.

8.4.1.1 Carnatic music

In CM tradition, the mridangam is played according to its melody and lyrics. While singing, the vocal artist also uses visual hand gestures to indicate the progression through the metrical time cycles (tāla) and hence implicitly keeping track of the tempo [PE05]. This frees the mridangam from a timekeeping and rhythm tracking role, enabling more complex improvisation and exposition of rhythmic patterns during a concert performance. The mridangam is played with naḍe, jati, uruṭu, and tirmāna (muktāya), which includes different kōrvai with different time signatures according to the lyrics and melody following the lead artist. In this system, the mridangam player follows the lead artist's manodharma (vocal or instrumental), aiding them in maintaining the tempo.

According to the artists' compositions and temperaments, the mridangam player starts playing in madhya laya or dṛt laya. The mridangam rendition is according to the musical compositions or sarvalaghu maintaining the tempo by playing different naḍe with respect to the melodic variations. Small muktāya are played at the end of each section, such as pallavi, anupallavi, charanam to indicate the ending of that particular section and the beginning of the next section. Apart from this, kārve (pauses/gaps) are also showcased to match the compositions. Mridangam artists anticipate and reproduce the lyrics and melodic phrasing almost verbatim.

In a CM concert, the lead artist will perform "neraval" in certain compositional parts during kriti, kīrtane and pallavi. This is followed by "svara kalpana", which is based on the manodharma that may or may not have kōravai, which is based on precise calculations that help to compose a variety of rhythmic patterns within the framework of the tāla of the music piece. The mridangam player even tries anticipating the rhythm-bound melodic phrases and plays percussion patterns that match them at points. The continuous rhythmic improvisation is sometimes aligned with the main melodic flow and sometimes via cross-rhythms. There is a separate portion called a "tani" exclusively for solo performances of percussionists after all these sections [PE05]. This is the section where the mridangam players showcase different nuances, which are based on the precise calculations on tāla. CM tradition has different forms of compositions such as svarajati, varnam, kīrtane, pallavi and tillāna. It also has semi-classical renditions such as ugābhōga, devaranāma, vacana, etc. Mridangam artists adapt and play according to the aesthetics of the rendition.

A senior vocalist Smt. Sakuntala Narasimhan says "the role of the tāla vādya (mridangam, ghaṭam, kanjira) is mainly to follow and embellish the rhythmic patterns with matching phrases on percussion. A percussionist can do this well only if he (or she) is conversant with a large repertoire of kṛti (songs) which he (or she) can then accompany well, anticipating the sequence of sangati (musical variations), or chittasvara (pre-set svara patterns), etc.)." [Nar94]

8.4.1.2 Hindustani music

The vocal and instrumental recital in HM is called a "khayāl". This form of music tradition is performed with composed music pieces with lyrics and melody set to specific tāl, but the tabla accompaniment techniques are not the same as the CM tradition. There is no practice followed of the lead artist using hand gestures to indicate the progression through the tāl. Here the tabla bears the timekeeping responsibility by maintaining the tempo and indicating the tāla positions [Nar94]. Since the lead musician will continuously refer to the tabla, it is essential that a universally understood conventionally established pattern is used to indicate metrical progression. This pattern is called ṭhēkā [Cla08], which can be interpreted as a canonical rhythmic pattern of a tāl that is well understood and played to indicate the progression through a tāl cycle. The meeting of the leading artist and the tabla ṭhēkā at the instances of "sam" itself is important in this tradition.

The first essential rhythmic idea that a tabla player needs to be comfortable with is playing a ṭhēkā at a steady tempo [Bag98b]. The calculations are not very well organized compared to the CM tradition. The importance is given to the "sam" and "husi". The next step involves the embellishment of the basic ṭhēkā. Thus maintaining the main ṭhēkā, the tabla player will make use of the base drum (bāyan) vividly and plays small variations such as small mukudas and tihāi according to the flow of the melody. Still, the fundamental responsibility of timekeeping remains [CC]. Undoubtedly, the ornamentation of the ṭhēkā is far more subdued as compared to the solo passages that intersperse the melodic elaboration.

The khayāl performances begin with an alāp being a prelude to the main melodic framework. The lead artist introduces the composition after the alāp. The rendition starts with vilambit lay and is followed by madhya lay and dṛt lay progressions. The tabla player has to be adept at keeping time at all tempi. At points, the tabla player may even anticipate the melodic passages, particularly those that are rhythm-bound, and will proceed to play percussive passages along with those that the vocalist is presenting. This is called the tabla sāth-sangat [Pra11]. Tabla player also needs to respond to the harmonium or sārangi players in a vocal concert, both of whom provide melodic accompaniment. Various tihāi are played in a single concert to mark the conclusion of a melodic idea. There are different types of tihāi - some that are short, others that span over several āvartan (tāl cycles). The tabla players only play cross-rhythms and individual improvisations if the lead artist permits a short solo during the concert.

Few semi-art music renditions in Hindustani are thumris, dhun, thappa. Apart from these, one can see dadra, kawwali, ghazals and bhajans. The tāl used for these renditions are usually different from those that are employed in khayāl. The tabla player introduces laggi, laggi-chanti, or laggi-nada in the thumri-dadra and allied forms. Laggi refers to the section, which allows the tabla player to launch into kaherva or dādra based variations when the vocalist changes the tāl of the theme to accommodate these patterns [Pra11].

8.4.2 Accompaniment to instrumental music

There are basically two prominent styles of instrumental performances - (1) tantrakāri (plucked string) style, (2) gāyaki (vocal) style. Tantrakāri style is most often observed in Hindustani tradition, during the gāyaki style in Carnatic tradition. The gāyaki style of instrumental playing is developed by imitating the singing voice, makes extensive use of the mind (glide), while the tantrakāri style is characterized by melodic leaps. The choice of style influences the instrumentalist's repertoire and concert structure as well as the actual rendering of the raga in terms of the realization of phrases and ornamentation [VR]. These two styles demand a different approach from percussionists.

In carnatic music, the gāyaki style closely follows the structure of vocal rendition. The instrumental concert contain all the attributes that are present in vocal concert such as pallavi, anupallavi, charanam, svara kalpana etc. Thus the mridangam accompaniment to gāyaki style is similar to that of vocal music. The mridangam is played with naḍe, jati, uruṭu, and tirmāna which includes different kōrvai with different time signatures according to the lyrics and melody following the lead artist.

Conventionally, Hindustani instrumental concerts begin with an alāp to establish the mood of the rāg. The performance begins at a slow pace and then works its way to a climax at an accelerated tempo. Instrumental music also demands a controlled ṭhēkā from the tabla player. There exists a section called the savāl-javāb (question-answer). The instrumentalist chooses a series of melodic-rhythmic phrases that shorten over every few āvartan of the ṭhēkā. The tabla player reproduces each of these phrases, and the entire section culminates with a tihāi. The other section is generally termed as sāth-sangat, literally meaning "being together or in the company of". In the sāth-sangat section, the tabla player is required to anticipate the melodic-rhythmic movement and play simultaneously rather than reproduce later. This section is more challenging and is in many ways more engaging for the initiated listeners. The savāl-javāb is more appealing to uninitiated listeners. Thus even though tabla and mridangam play the role of timekeeping, the approach and techniques are different, meeting the demands of their respective cultures.

8.4.3 Accompaniment to dance forms

Rhythm plays a vital role in both music and dance. The rhythm accompaniment to nṛtya (dance) is very distinct from what is heard in vocal or instrumental performance. The pure dance elements are almost completely based on preconceived rhythmic compositions incorporating gesture, movement, and footwork. The percussionist needs to accompany the vocal and also needs to observe the footwork and play [Pra11]. Hence the job is much more complex compared to the normal accompaniment. The primary duty is to establish the context, theme, mood, and aid in the audience joining the experience. Dance performances also include a special repertoire highlighting the abhinaya or expressive aspect. Hence the playing technique should be in a manner complementary to the expressions, emotions, and movements of the dancer [KHA].

The South Indian dance forms like Bharatanātyam, Kuchipuḍi and Mōhini Attam use mridangam for rhythmic accompaniment. Usually, other instruments like kanjira, manjira, nattuvangam (cymbals) are also used with mridangam. There is no coded tradition in playing mridangam for dance. The mridangist has all the liberty to improvise and embellish the dance with her/his improvised playing.

The North Indian dance forms like Kathak and Odissi either use tabla and pakhavaj for rhythmic accompaniment, without any supplementary percussion instruments. The compositions such as gaṭ and paraṇ are very popular for dance accompaniment. The laggi section was originally related to the footwork of the performers, who were also trained in kathak and would add gesture, movement, and footwork to heighten the emotional content of the vocal composition. The tabla player would reproduce rhythmic patterns that corresponded to the footwork, and at times, would anticipate and suggest other patterns.

8.4.4 Mridangam tani-āvartana and tabla solo

Though percussion instruments tabla and mridangam are used as accompaniments, they also developed as exclusive instruments for solo performances. Both HM and CM traditions have a separate music forms for showcasing the skills and nuances of the percussionist. In the case of CM, a section in the main concert is dedicated to percussion instruments. In HM, a separate solo concert form exists for percussion.

8.4.4.1 Tani-āvartana

In this section of the concert, the mridangam player showcases different tempo designs playing various nuances based on the precise calculations on tāla. Since neraval, svarakalpanā prastāra precedes the tani-āvartana, the percussionist usually starts the tani with the same tāla and naḍe as that of the svarakalpanā prastāra and the whole tani-āvartana is played at a single metrical tempo.

The mridangam artistes follow a standard structure while developing the tani-āvartana that is generally accepted. The tani-āvartana initially starts by showcasing the tāla structure which are called the sarvalaghu patterns. These structures gradually build in complexity hinting the more structured mathematical patterns to come in the future. These hints are called “aasu”. Further, tani-āvartana can be broadly sectioned into different abhiprāya (tirmānā). The abhiprāya are thematic development of rhythm structures in different naḍe such as chaturaśra and tiśra. Each abhiprāya ends with kōravai. Towards the end, the artiste starts playing faster (farans), building up the necessary momentum for playing mohara and longer kōravai. There are different kinds of kōravai such as kāla, mel-kāla, tri-kāla, jāti, naḍe kōravai. Each of these has a specific composition structure upon which the artiste builds. The above-described structure may change depending on the artists and concert.

If the additional percussion accompaniments such as ghaṭam, kanjera, and morsing are present, then the total time in the tani-āvartana will be divided accordingly to showcase the individual performances, e.g. if there are only mridangam and ghaṭam, the structural framework of the tani-āvartana will usually be as follows:

Table 8.1: An example for the mathematical progression in CM to create different phrase lengths. Each phrase has two parts that are progressively made shorter to achieve the required total length. The numbers in parentheses indicate the length of the phrase.

Part-1	Part-2	Total length (in aksharā)
Ta Ka | Dhi Mi (4)	Ta Ka | Ta Ki Ta (5)	4 + 5 = 9
Ta Ka | Dhi Mi (4)	Ta Ka | Jha Nu (4)	4 + 4 = 8
Ta Ka | Dhi Mi (4)	Ta Ki Ta (3)	4 + 3 = 7
Ta Ki Ta (3)	Ta Ki Ta (3)	3 + 3 = 6
Ta Ka (2)	Ta Ki Ta (3)	2 + 3 = 5
Ta Ka (2)	Dhi Mi (2)	2 + 2 = 4
Ta (1)	Ki Ta (2)	1 + 2 = 3
Ta Ka (2)		2
Ta (1)		1

The mridangam will always start playing at first the prescribed tāla building from basic strokes from some time, usually in a particular naḍe (chaturaśhra, for example). Ghaṭam follows, which tries to keep the same theme built by the mridangam in the first cycle by playing in the same naḍe [Gir20]. In the second cycle, the mridangist usually may change the naḍe (to tiśra, for example). The ghaṭam will usually play in the same naḍe or switch to some other naḍe (khanḍa). These may or may not continue for more than two cycles, usually owing to time constraints. The percussionists play different naḍe at various gati, showcases urutus, small kōravai improvisations that are spontaneously composed.

These abhiprāya are followed by the korapu, which is usually seen as a question-answer (similar to the savāl-javāb in HM) between mridangam and ghaṭam. Here it starts with multiple āvartana of rhythmic patterns by the mridangam followed by ghaṭam. Here, the ghaṭam player is bound to follow and play whatever mridangam has played. It can be translated as “rhythmic descent” or “step by step reduction”. The duration of the rhythm patterns in korapu keeps reducing progressively into a full cycle, half cycle, quarter cycle, till it reaches the length of a single beat. Both artists then start playing together to go towards the end by playing mohara and the final muktāya [Gir20].

To summarize, the sequence of such a tani-āvartana would broadly as: sarvalaghu patterns → abhiprāya in a specific naḍe → change of naḍe to a different naḍe → back to starting naḍe → korapu → farans → mohra → final muktāya.

The tani-āvartana rendition is essentially based on precise mathematical progressions. The intricacies are built according to the artist’s manodharma at that spur of the moment. The mathematical progressions are given more prominence in the performance. An example for mathematical progressions can be seen as the number of aksharā may start descending from 9, 8, 7, 6, towards 1, and start increasing from 1 to 9 as shown in Table 8.1, or it can be the progressions on odd numbers such as 1, 3, 5, 7, 9.

8.4.4.2 Tabla solo

Tabla solo is a different music form compared to the Hindustani vocal or other instrumental concerts. The tabla is the main instrument, and the instruments like sārangi or harmonium or sometimes violin is used for *lehra* accompaniment. The purpose of lehra is to act as background melody partly but is to be mainly responsible for timekeeping by indicating the metric tempo for reference [Bag98b]. Tabla players are free to play the compositions and do as many improvisations as possible by showcasing their skills. A tabla solo is intricate and elaborate, developing various pre-composed forms within the framework of the *tāl*.

Tabla solo consists of different compositions such as *thēkā, uṭhān, pēśkār, kāyadā, rēlā, rau, gaṭ, paraṇ, tukḍā, chakradhār* [Pra11]. Each composition has different functional and aesthetic roles in a solo performance. Although there are no set rules for the presentation of the tablā solo, there are generally accepted guidelines according to which a pēśkār and kāyadā come before the gaṭ, tukda, or chakradhār. The rēlā can be played anytime after the kāyadā and be used as a finale. However, for gaṭ, tukḍā, and chakradhār, there is no fixed sequence, but these are usually performed at the end of the recital. Therefore, the sequence of these pieces would broadly be as follows: pēśkār/uṭhān → kāyadā → rēlā → gaṭ → tukḍā/chakradhār.

The structural framework of a tabla solo in most of the cases will be as follows: The lehra accompaniment plays an alāp initially for a few minutes. Then the tabla player starts with the initial uṭhān or directly pēśkār. A pēśkār rendition can go from five minutes to more than 15-20 mins depending on the overall concert duration. Next to pēśkār, various kāyadā compositions from different gharānā are played one after the other. Each kāyadā will be having short-term endings called tihāi. Further, the compositions such as rēlā and rau are presented. Later on the metric tempo increases and the fixed compositions such as gaṭ, stuti paraṇ, tukḍā and mukudā are played. In the end, the chakradhār are played, which are essentially long-term muktayā. All the sections contain pre-composed compositions developed within the framework of the tāl.

It is necessary to point out the role that is played by the ṭhēkā in the context of solo tabla recitals. Except at the beginning of the performance, the basic ṭhēkā is usually played for a few cycles in between adjacent compositions, marking the start and at end of the item [Pra11]. These ṭhēkā are played at the metric tempo as indicated and maintained by lehra accompaniment (the barābar lay) on which the repertoire and the ensuing improvisation are based. After establishing this tempo, the player can gradually increase the tempo as the performance progresses [Bag98b].

Based on the item/compositions' structure and their position in the recital, one can find three major sections in the tabla solo concert. They are pēśkār, kāyadā, and gaṭ–tukḍā–chakradhār (GTC) sections. Pēśkār is an extempore rendition usually played in a slower metric tempo. Kāyadā are improvisatory compositions upon a theme. The GTC are examples of fixed compositions. There may not be a fixed repertoire, and the artistes are free to come up with their own compositions adhering to the tāla structure. Each composition has different functional and aesthetic roles in a solo performance. It is the compositions that are the heart of tabla's repertoire. The fingering techniques and playing methodology are fundamentally based on the compositions. Any extempore rendition will usually be built upon the basic compositional pattern. This is called vistār. The compositions are the central portion of the tabla repertoire. Hence tabla solo performance is fundamentally different from mridangam tani-āvartana. The details of different tabla compositions are elaborated further in Appendix-A.

8.4.5 A percussion ensemble (tāla vādya kacēri)

Due to technological advancements and better media communication, cross-cultural influences pervade most music cultures. Due to such influences, the scope of percussion instruments in both HM and CM systems has extended over the years. In more recent years, tabla players have been influenced by percussion instruments other than the pakhavaj, like the mridangam, and have added some of this content to their recitals [Pra11]. Few tabla players such as Suresh Talwalkar and Bikram Ghosh have undergone training in Carnatic rhythm. Similarly, mridangam player Anantha R Krishnan has also taken Hindustani rhythm training.

Cross-cultural influences have led to collaborative ventures and performances. We have instances where Vidvān Palghat Raghu and Vidvān Umayalapuram K Sivaraman accompanied the Hindustani sitar by Pandit Ravishankar on the mridangam alongside Ustad Alla Rakha on tabla. Likewise, the tabla has also been used to accompany Carnatic concerts alongside mridangam. The past few decades have seen a rise in joint performances by Hindustani and Carnatic music practitioners. Duets, trios, and more, consisting of Hindustani and Carnatic musicians, are seen in performance. The percussion ensemble is seen in such vocal and instrument performances and emerges as separate concerts where percussion occupies center stage. These special performances of percussion ensembles are popularly known as tāla vādya kacēri.

In Carnatic music, multiple percussion instruments also help set up interactive percussion performances and percussion ensemble concerts. The percussion ensembles may include the mridangam, tabla, ghaṭam, khanjira, morsing, konnakol (oral recitation of syllables, considered an art form under percussion) and possibly the tavil.

More recently, other instruments such as drum kits and drum pads are also included, in a not-so-aptly named ‘fusion’ concerts. In Hindustani-Carnatic percussion collaborations, there is cooperation and compromise among the practitioners of both systems regarding the proportional importance of mathematics and repertoire. In most cases, the dialogues start and end at a particular chosen tempo. Due to demands placed by unchanging tempo, traditional tabla repertoire composed in different tempi take a back seat during such dialogues. Consequently, the tabla players focus on the tonal quality of their instrument utilize various metrical patterns to introduce variety through individual and combined left and right-hand strokes and on the speed that the instrument lends itself to. Hindustani percussionists respond to Carnatic mathematical designs with the rēla patterns, using basic syllables as accents rather than as detailed compositional expositions. Similarly, the Carnatic percussionists face difficulties in executing speedy passages like rēla and find it alien to elaborate in a theme and variations pattern like in the gaṭ or the kāyadā forms.

Another easily noticeable dissimilarity between the two approaches is a lehara or nagma in the Hindustani percussion solos. This accompaniment is absent in the Carnatic tani-āvartanam or taal vādya kacēri presentation. The absence of melodic accompaniment in a Hindustani-Carnatic percussion dialogue is, to an extent, unsettling to practitioners and lovers of Hindustani percussion. Lately, Hindustani-Carnatic percussion dialogues have seen melodic instruments like the sitar and even western instruments like bass guitar and synthesizer. However, these instruments are not used for delineating a nagma, but to establish a melodic theme more in the Carnatic mold, to act as a starting point for the subsequent unaccompanied tani-āvartanam.

One of the tabla artists who perform jugalbandi with mridangam says, “we (tabla and mridangam player) discuss the kind of muktāya to be played. Both will develop from starting individually to arrive towards the muktāya. Whereas while playing with another tabla player, the compositions to be played will be discussed, and the ending tihāi will be played according to the compositions”. Thus, it becomes increasingly clear that the Hindustani and Carnatic rhythm and percussion perspectives need to be addressed if tāla vādya kacēri consisting of percussion instruments from both systems are to be experimented with.

8.5 Summary

The percussion instruments mridangam and tabla have a special position in Indian art music and fusion performances. This article is intended to help technology enthusiasts, engineers, and researchers build common and culture-specific computational tools for analyzing percussion in CM and HM. Hence, we wished to highlight what we consider are some primary distinctions between the two systems, which would help us build culture-specific tools or extend generic tools to these music cultures for percussion analysis. This article aimed to portray the similarities and differences between the mridangam and tabla based on their sociocultural contexts, performance practice/techniques, as accompaniments, and in solo performances.

After a brief discussion of the history and basic features of HM and CM, the article described the percussion instruments’ role in these music cultures. Though HM and CM cultures share a common origin, they evolved as different music cultures, thus sharing many similarities and contrasts. This extends to the accompanying percussion instruments as well. The evolution of mridangam and tabla in terms of sociocultural context and practices were explained while briefing out the stylistic schools of percussion, the bāni, and gharānā. Based on our discussion of sociocultural aspects of tabla and mridangam, it is evident that there are several similarities and parallel concepts in the percussion systems of Hindustani and Carnatic music.

Mridangam and tabla also evolved in their construction, technique, repertoire, and practices. The instruments’ structure and the music culture’s influence contrast with their timbre. The article described the roles of mridangam and tabla during accompaniment and compared the percussion techniques during solo concerts and during collaborations that showcase the similarities and contrasts between the two instruments. Mridangam usually follows the melodic variation of the main artist. Tabla always accompanies by playing essentially the conventionally established pattern “ṭhēkā” that can be universally understood. The repertoire

Table 8.2: Summary: a comparison between mridangam and tabla along different aspects discussed in the article

Characteristic	**Mridangam**	**Tabla**
	Construction and Timbre	
Construction	Double-sided single hollow drum	Twin drums with a single drum head each
Tuning	Relatively lower pitch	Relatively higher pitch
Resonant strokes	Shorter sustenance	Longer sustenance
	Sociocultural context	
Stylistic schools	Less well-defined bāni	Well-defined and documented gharānā
# schools	No specific number documented	Six distinct gharānā
	As an accompaniment	
Style	Closely follow melody and lyrics	Play the ṭhēkā at a steady tempo
Improvisation	More freedom to improvise and play variations	Freedom to improvise and play variations only in restricted sections
Tāla-keeper role	Not bound to play ṭhēkā	Bound to play the ṭhēkā indicating the tāl cycle
	Solo performance	
Performance	Part of the vocal/instrumental concert	Independent concert
Repertoire	Few compositions, largely extempore	Large repertoire of compositions
Emphasis	Emphasis on precise mathematical progressions	Emphasis on the rhythmic composition
Tempo	Single tempo for the whole solo performance	Different compositions can be played on different tempi
Tāl progression	Hand gestures by lead artist	Lehra
paraṇ	No such concept exists	Common in few renditions
Subdivisions	An akśhara cannot be sub-divided in theory	All kinds of subdivisions are accepted and are based on the compositions

and patterns of mridangam are primarily based on precise mathematical progressions, whereas for tabla, it's based on compositions with a poetic aesthetics language with the flow.

The article provided information about the similar basic qualities and contrast between percussion of CM and HM based on various factors. The article also provided an overview of different sections of CM and HM, briefing out the roles of mridangam and tabla in those sections. This helps us understand the percussion in two traditions by including them in their sociocultural contexts. These are significant factors to be considered while developing the computational techniques to analyze the percussion patterns and styles. We conclude the article by summarizing important takeaways from the article in Table 8.2. Audio/video examples of the performance clips corresponding to the different sections in this article and the links for further reading and references can be accessed on the companion website to this article [1].

[1] https://gowriprasadmysore.github.io/SPARC_Addons/

reference

Glossary

tani-āvartana The solo performance of a percussion ensemble.

akṣara Literally, a syllable in Indian languages, corresponds to the lowest metrical pulse (subdivision) in Carnatic music.

mātra The lowest defined metrical pulse in Hindustani music (equivalent to a beat)

solkaṭṭu The onomatopoeic oral percussion syllables in mridangam.

bōl The onomatopoeic oral percussion syllables of the tabla.

naḍe The subdivision structure within a beat in Carnatic music

jāti The subdivision structure within a beat in Hindustani music

caturaśra A naḍe with 2 or 4 akṣaras per beat.

tiśra A naḍe with 3 or 6 akṣaras per beat.

khanda A naḍe with 5 akṣaras per beat.

āvartana One complete cycle of a tāla.

tāla The rhythmic framework of IAM.

gharānā The stylistic schools of Hindustani music

bāni The established stylistic schools of mridangam

lay The tempo class in Hindustani Music, (laya) in Carnatic Music

druth Fast tempo class

madhya Medium tempo class

vilambit Slow tempo class

sam The first mātra of an āvartana

ṭhēkā The basic bōl pattern associated with a tāl

palaṭā A compositional form in tabla

khayāl A major form of Hindustani classical music

thumri A North Indian vocal form

Appendix - A

Description of different rhythmic compositional/improvisatory forms in Indian art music percussion.

pēśkār Pēśkār is rendered in a slow tempo with different combinations of strokes showcasing the nuances which are possibly extempore. It is derived from the work ”Pesh Karna” which means to showcase. The tabla players from Banaras gharānā will use the term ”alāp” for the initial slow rendition.

kāyadā Kāyadā are the significant portions of the solo concerts. It is derived from the Persian word ”kaidh” which means rules or tradition. Kāyadā are essentially the elaborate classic compositions with distinct bōl based on tāla, mātra, khanda, and peṭṭu, husi. Initially, the face patterns are played in first and second speed, then the paltā and prakār are embellished in fourth speed. These are the improvisatory compositions upon a face pattern. Usually, the kāyadā section will be a bigger chunk in the whole concert. These kāyadā are the compositions from different gharānā. Each kāyadā has a specific theme, jāti/naḍe.

gaṭ Gaṭ are the examples for fixed compositions. Gaṭ are the compositions initially composed to accompany kathak Dance in Lucknow gharānā. Gaṭ will be having attractive rhythmic bōl patterns indicating the peṭṭu and husi. Usually, tihāyi and palaṭā will not be there in gaṭ. There are compositions which are called gaṭ kāyadā wherein one can have the palaṭā and prakār for the gaṭ which includes tihāyi.

tihāyi The group of shorter bōl patterns which can start from any mātra of the tāla, which are played three times so that the last bōl comes precisely to the Sam are called tihāyi. These will be at max one to two āvartanas long. tihāyi can be seen as short term muktāya in CM tradition.

tukḍā Tukḍā are the compositions having distinct bōl that start and end at Sam.

mukudā Mukudā is the composition that can start at Sam or any mātra and ends at Sam which are played three times. Mukudā may or may not include tihāyi

stuti paraṇ These are the special bōl compositions played on pakhavaj. These are played with khulā and vigorous bōl, which are attractive that last for two to three āvartana. Most of the paraṇ are the poetic and rhythmic hymns in praise of lords recited at first and then played. paraṇ are played on tabla also. These renditions are most seen in Purab and Banaras gharānā.

uṭhān In Banaras gharānā, the tabla solo starts with ”paraṇ” or ”gaṭ” with attractive bōl pattern played before the ṭhēkā which is called uṭhān. It can last for one or two āvartana.

chakradhār Gaṭ inclusive of tihāi, which are composed to play three times, are called chakradhār. This will come to sam only if it is played three times. These can be seen as long term muktāya in CM tradition.

rēlā Rēlā are also elaborate compositions like kāyadā have closely packed bōl. TiReKiTa, DhiRe DhiRe, bōl are usually played extensively in rēlā. Rēlā are played especially in fourth and eighth speed as well. The word rēlā is taken from rail as it is played at a very high speed sounds like a fast-moving rail.

rau These are the compositions using the bōl from gaṭ, but played at greater speed, gives the flavor of rēlā for the listeners.

laggi These are the smallest compositions in the realm of tabla playing. Both are used mainly in accompanying the lighter forms of vocal music such as thumris, ghazals, and bhajans [Sax06]. It is commonly understood as a theme and variations development, which consists of an unbroken chain or string of strokes, often not numbering more than four to six. It is played at a high speed [Pra11].

neraval Neraval is the elaboration and improvisation of melody for a particular line. Usually, just one or two lines of text from the song are sung repeatedly, but with improvised elaborations. This elaboration remains within the framework of the original patters of duration (tāla).

farans These are brisk rhythmic patterns played in the faster tempo.

arudi Arudi is an ending pattern. This is a Tamil word meaning ”limit” or ”boundary”. Arudi is also construed to be an ending point of the first portion of a rāgam-tānam-oallavi (RTP) [Bal15].

tirmāna Tirmāna are short ending phrases; the rhythmic phrases played between pallavi and anupallai, pallavi and charanam is called tirmāna, if it observes the rule of not exceeding 3/4th of a tāla Cycle. [Bal15].

kōravai Kōravai is a Tamil word meaning ‘joining’. A kōravai can be defined as a rhythmic pattern set to a metre adhering to a structure.

mohara/mohra Mohara is a rhythmic form played 4 times just before the percussion solo ends. Various types of mohara can be formed by observing the following rules: 1) Mohara is generally rendered for 4 tāla cycles. 2) The second cycle is just a repetition of first cycle. 3) After mohara a pattern called tirmāna or kōravai is generally played. Some artistes call short ending patterns as short moharas and the final climax pattern as a long mohara [Bal15].

References

[Bag98b] Bagchee, S. *Nād: Understanding rāga music*. Eeshwar, 1998.

[Bal15] Balaji, M. J. “Rhythmic Syllables: Introduction, Analysis and Conceptual Approach in Carnatic Music of South India”. In: *Rupkatha Journal on Interdisciplinary Studies in Humanities* vol. 7, no. 3 (2015), pp. 149–161.

[Bha91] Bhat, R. “Acoustics of a cavity-backed membrane: The Indian musical drum”. In: *The Journal of the Acoustical Society of America* vol. 90, no. 3 (1991), pp. 1469–1474.

[CC] CHANDRA and COURTNEY, D. *Tabla v/s Mridangam*. URL: https://chandrakantha.com/.

[Gir20] Giridhar, U. *Description of Tani Avartanam*. 2020. URL: https://www.ghatamudupa.com/.

[Got93] Gottlieb, R. S. *Solo tabla drumming of North India: Its repertoire, styles, and performance practices*. Motilal Banarsidass Publishers, 1993.

[KHA] KHANNA, R. *Mridangam and Bharatanatyam*. URL: https://kalyanikalamandir.com/blogs/mridangam-and-bharatnatyam/.

[Nar94] Narasimhan, S. “THE ROLE OF TALA VADYAS IN CARNATIC AND HINDUSTANI MUSIC: A COMPARATIVE ASSESSMENT”. In: *Proceedings of The Talavadya Seminar No. 1* (1994).

[PE05] Primary, D. of and Education, S. *Karnataka Talavadya, Vidwat Purva Hanta (Per- Proficiency Grade)*. Government of Karnataka, 2005.

[Pra11] Pradhan, A. *Tabla: A Performer’s Perspective*. BookBaby, 2011.

[Sax06] Saxena, S. K. *The Art of Tablā Rhythm: Essentials, Tradition, and Creativity*. 8. Sangeet Natak Akademi, 2006.

[Sri+17a] Srinivasamurthy, A. et al. “A Data-driven bayesian approach to automatic rhythm analysis of indian art music”. PhD thesis. Universitat Pompeu Fabra, 2017.

[SS+78] Shringy, R. K., Sharma, P. L., et al. *Sagıta-ratnākara of Śārgadeva: Treatment of svara*. Vol. 1. Motilal Banarsidass, 1978.

[Sun21] Sunil, E. N. *Resounding Mridangam: The Majestic South Indian Drum*. Erickavu N Sunil, March 2021.

[VR] Verma, P. and Rao, P. “DISTINGUISHING MUSICAL INSTRUMENT PLAYING STYLES WITH ACOUSTIC SIGNAL ANALYSES”. In: ().

Chapter 9

Analytical Models for Vibro-Acoustics of a Mridangam

Harikanth Mohandas and Chandramouli Padmanabhan

9

Abstract

Sound is produced by percussion instruments due to the vibration of the membranes when struck. Normally the frequencies generated are not integer multiples of the fundamental frequency to which the drumhead has been tuned to. A unique characteristic of the Mridangam, studied in this paper, is the fact that one of its two drumheads, the *Karanai*, has natural frequencies which are integer multiples of the fundamental. Now there is a closed air cavity between the two drumheads in the case of a Mridangam (as well as Khol/Pakhwaj). When the drumhead is struck there is a possibility of some part of the vibration energy being transferred to the air cavity which in turn can produce sound in the cavity. In turn the acoustic energy can flow back to the drumhead modifying its vibrations. This is called *vibro-acoustic* coupling. The key question that this paper is trying to address is whether the *Karanai* vibration frequencies being harmonic multiples is due to the vibro-acoustic coupling and if so, what role does the shape of the cavity have on the sound generated by the Mridangam.

A very recent study has pointed out that the vibro-acoustic coupling between the acoustic cavity and the harmonic drum head may not be significant. A key question, that arises then, is why does the Mridangam cavity have a rather elaborate barrel-like shape. The primary objective of this research is to find out if changing the cavity geometry affects the membrane vibrations significantly. Analytical models are developed for the membranes and the air cavity for fulfilling this objective. Such models help to make rapid and efficient changes to the shape of the Mridangam.

A cavity-shape sensitivity analysis is carried out for three widely different shapes. These results are also validated using the commercial FE software ABAQUS. From the extensive investigation, it becomes clear that the shape of the Mridangam cavity has no significant effect on the harmonic drum head vibrations. Mridangams with such shapes will be made and tested in the future to confirm the numerical predictions.

9.1 Introduction

Mridangam is a familiar presence in the Carnatic *katcheris* or music concerts of South India. It is the primary rhythm accompaniment, an indispensable part of the traditional ensemble. This double-headed percussion instrument, believed to have its origins in the Vedic period, has been part and parcel of the region's music and dance cultures for ages. It is a common belief that Mridangam got its name because early drums used to have clay as its body (In Sanskrit *mrit* means mud and *anga* means limb). But [Rag53] explains that this is not the case. During Bharata Muni's time, *mrit* was the ingredient applied to the main drum-head, to give Mridangam its characteristic sound; it was he who named the instrument for this reason. Its construction has evolved over the centuries, and a familiar present-day Mridangam looks like one shown in Figure 9.1. The contemporary version has the shell made of wood, commonly from the jackfruit tree, and the drum heads made using hides from various animals. The drum head visible in the picture is the primary membrane karanai, or more commonly the *valanthalai*. The opposite end is called the thoppi or *edanthalai*.

Figure 9.1: Modern version of Mridangam

Sound is produced by percussion instruments due to the vibration of the membranes when struck. Normally the frequencies generated are not integer multiples of the fundamental frequency to which the drumhead has been tuned to. A unique characteristic of the Mridangam, studied in this paper, is the fact that one of its two drumheads has natural frequencies which are integer multiples of the fundamental. Now there is a closed air cavity between the two drumheads in the case of a Mridangam (as well as Khol/Pakhwaj). When the drumhead is struck there is a possibility of some part of the vibration energy being transferred to the air cavity which in turn can produce sound in the cavity. In turn the acoustic energy can flow back to the drumhead modifying its vibrations. This phenomenon, where there is a constant flow of energy from the drumhead to the air cavity and back, is called as *vibro-acoustic* coupling. Figure 9.2 shows a graphical representation of the possible coupling.

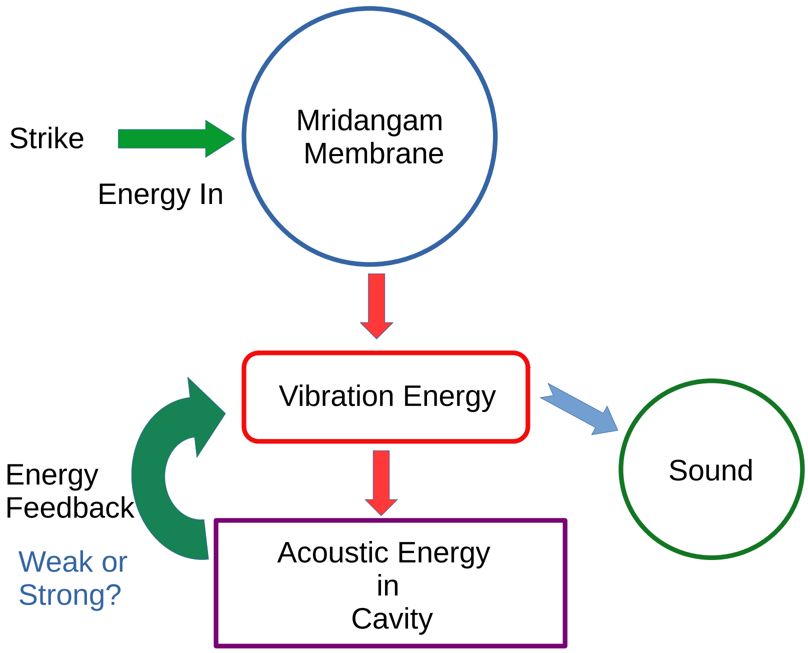


Figure 9.2: A graphical representation of vibro-acoustic energy flow in a Mridangam.

Mridangam has unique musical qualities; qualities that have aroused curiosity in the scientific community from the likes of [Ram22]. He was the first to notice the tonal nature of its main drum head. He observed that it produced overtones which had musical attributes like that of a stringed instrument. A conventional stretched membrane, like that of most Western percussion instruments, produce ‘noisy’ overtones and is not musically pleasing. This discernible character of Mridangam is also inherent to some of the other Indian drums like the Tabla, Pakhawaj and Khol. A Western counterpart having this tonal property is the kettledrum or timpani used in modern orchestras. But it achieves this with a different approach compared to Indian instruments. The air loading is attributed to making the overtones of the membrane harmonic in a kettledrum rossing1992. In contrast, the Indian drums achieve this with the addition of a central mass loading, seen as a black patch, on the leather. [Ram34] found out that the tabla membrane produces at least four overtones that are harmonic to the fundamental. A skilled musician, using different playing strokes, can activate these modes, thereby generating a variety of musical sounds.

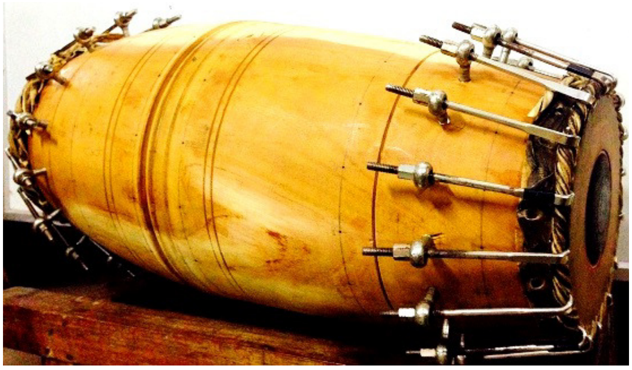
Not much work has been done on the Mridangam shell, the cavity it encloses, and how these interact with the membranes, until recently. [Soo18] was one of the first to venture into this area. He conducted experimental and numerical investigations on different parts of the drum and how these interact with each other. This paper is a continuation of his work. The dynamics between membranes and the cavity are carefully examined. Of more interest is how the Mridangam shape contributes to these interactions and consequently, the sound generated. A lot of research on musical instruments is carried out and reported in the literature. As a result, these instruments have evolved with modern methods and materials used for manufacturing. But very little research goes into the traditional 'vadyas' of India. Consequently, people still rely on old methods and practices for making them. Most artisans making these instruments pass on the craft as heirlooms and no documentation exists for the same. The industry, as well as the instruments, faces a challenge as fewer families follow their tradition and others move on to more lucrative jobs. Thus, it is imperative to achieve some standardisation in manufacturing so that new methods and materials can be employed. This task requires the scientific community to involve more in such kind of studies. This investigation is a contribution to this goal.

9.1.1 Construction of a Mridangam

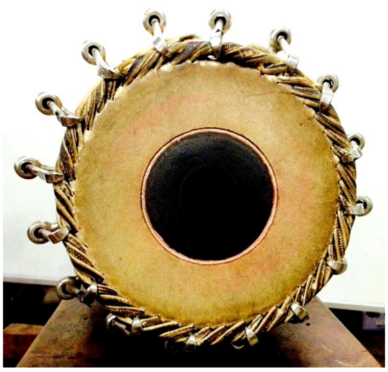
Even though the reason for naming is different, the predecessors of Mridangam might have had its body made of clay raghavan. From that, Mridangam has evolved over the centuries to take the present form. The shell is now predominantly made of wood. Jackwood is the preferred material of choice for its availability, lower density, medium hardness for machining and apparent sound radiation characteristics sivaraman. This Mridangam is heavy and could weigh more than 10 kgs. A few have attempted to incorporate lighter fibreglass body and a few others with open frame structures, to reduce weight. But the exponents are still to be convinced to use these newer innovations in the mainstream. The loading on the drum head has also evolved from clay during Bharata Muni's time to the contemporary paste made from minerals and starch. Recent innovations on the drum heads are the use of synthetic materials for membranes and the back patch. These materials have an advantage that they are less sensitive to temperature and humidity changes. This advantage eases the effort and frequency to tune the instrument. But these modifications have also not yet reached the mainstream.

Mridangam considered for investigation in this paper is a traditional variant made from wood and leathers (Figure 9.1). It is the typical kind in use today and hence is the choice for study. The *valanthalai* (head visible in Figure 9.1) and the *edanthalai* membranes in these Mridangams have leather hoops braided to them at their peripheries. A leather strap runs through these hoops (generally 16 times), weaving back and forth from one end to the other over the wooden shell. This strap secures the drum heads on to the shell and also applies tension to them. Artists fine-tune the tension by lightly tapping on the hoops. The pitch of the instrument depends on its size and can only be adjusted a few semitones up or down. Standard sizes made are 22, 23 and 24 inches long. The particular specimen used here is a 24-inch one adapted from [Soo18] (Figure 9.3a). It has a pitch equivalent to the Western note E3 (164.8 Hz). This Mridangam has metal hooks bolted to the shell instead of a leather strap to secure and apply tensions on the membranes. This configuration made it easy for him to disassemble the parts for experiments. Figures 9.3b and 9.3c show the dismantled drum heads *valanthalai* and *edanthalai*, respectively.

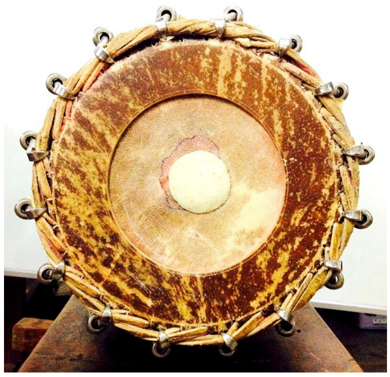
To better understand the details of the construction of Mridangam, consider Figure 9.4 from [SRN10], a schematic diagram illustrating its exploded view. The wooden shell has a barrel-like shape but is asymmetrical with the broadest portion slightly offset nearer to the *edanthalai* end (right side in the figure). The hollow cavity inside the shell has small cylindrical sections at the extremities, but the rest of the profile is curved. The right side membrane is larger compared to the left one. Hence the right part of the shell has broader cross-section compared to the left. The shell is made from a single block of wood by machining. Leathers from cow, goat and buffalo are used for making the drum heads, hoops and straps. The straps for securing the



(a) Modern Mridangam



(b) Valanthalai



(c) Edanthalai

Figure 9.3: Mridangam specimen and drum heads used for investigation (reproduced from [Soo18])

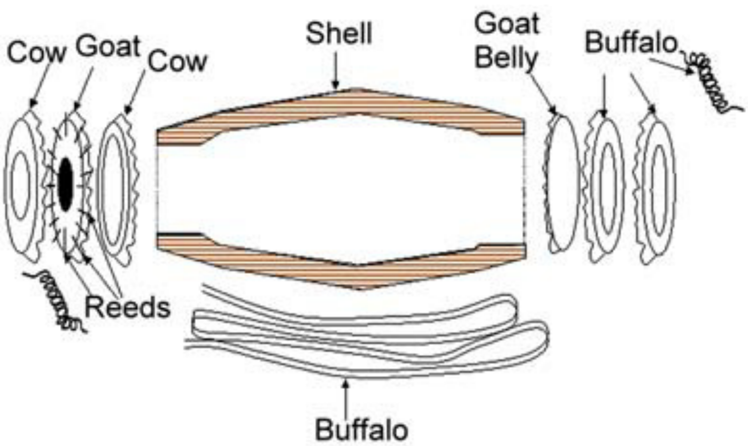


Figure 9.4: Exploded view of a Mridangam (reproduced from [SRN10])

drum heads are some times replaced with nylon straps or ropes, but other parts are predominantly leather. A membrane assembly is called a *mootu*; three layers of hides make up each of the *mootu*.

The *Valanthalai*, the harmonic membrane assembly is on the left side of Figure 9.4. It has two annular flaps of leathers made from cowhide and an in-between layer made from goat hide, which is the chief vibrating membrane. The three pieces of leather are stitched together with multiple strands of buffalo leather to form the braided hoop. As mentioned earlier, the fastening strap/metal hook runs through this hoop and applies tension on the membrane. Artisans also add small reeds (*kutchi*) or stones (*kappi*) in-between the outer flap and the middle layer; this presumably decreases damping and provide better resonance. The central mass loading seen as the black patch is on the primary vibrating membrane, the middle layer. This patch is an adherent composition of iron oxide, charcoal and starch. It has a slightly convex profile and is coated in layers; each coat is dried and polished with a smooth stone before applying the next. Rubbing with stone creates tiny cracks, making it flexible. In general, this finished patch is named karanai (or *saadham* or *soru*) and the whole *mootu* is *valanthalai*. In this paper, however, the usage karanai (followed from here) refers to the drum head.

The *edanthalai* or the thoppi *mootu*, shown on the right side of Figure 9.4, also has two annular flaps. But unlike karanai, these do not sandwich the vibrating membrane. The two thick annular pieces of leather from buffalo hides form the outermost layers, and the vibrating leather from goat hide forms the innermost layer. The buffalo layers are stitched together with leather straps forming a braided hoop, and the fastening leather/mechanism attaches to it. Goat layer, however, is not joined to this braid. It is instead separately stitched on to the outer leathers. Tension applied on the hoop thus does not get transferred to the inner layer directly but through the buffalo layers. The thoppi (used to refer to the drum head from hereon) is designed to produce bass sounds and hence is larger compared to the karanai. Thoppi is not harmonic and does not have any central mass loading like the karanai. Players, though, sometimes place a small lump of dough made from *rava* (semolina) on the thoppi before recitals. This practice is to tune the membrane to the desired pitch raman. One can refer [Soo18] for a more detailed insight into the process of making a Mridangam.

9.1.2 Motivation

There have been significant contributions in the first half of the twentieth century to the physics of Mridangam. Most of these contributions came from Raman raman1920,raman1927,raman. He did a lot of experimental studies on the harmonic drum heads of Mridangam and Tabla. A few others also ventured into the field like [RS54], trying to model the harmonic membrane mathematically. But the latter part of the century has not seen much interest in the subject. One can also see that most of the research related to the Mridangam is predominantly on the primary harmonic membrane, the karanai. While experiments have been conducted on the full instrument, none have studied its parts individually. That was the case until recently when [Soo18] performed a very detailed investigation He disassembled the Mridangam, and experimentally examined them independently. He also made use of numerical methods to build a vibro-acoustic model, which no one did previously. His observations were that the karanai was harmonic on its own, and the shell and enclosed acoustic cavity did not influence the membrane vibrations. Natural questions arising from these findings are: why is then the Mridangam shaped in an asymmetrical double cone shape? Does the shape of the shell matter at all? If it doesn't, what alternate geometries can be considered for it? These questions are the motivation for the present work, and the paper tries to answer some of them. A vibro-acoustic model of the Mridangam is necessary to fulfil this quest. A semi-analytical model is preferred over a numerical one as it would offer quicker insights on the effects of Mridangam cavity shape on the vibro-acoustic behaviour.

9.1.3 Literature Review

Raman's first publication on Mridangam was in the year 1920 with Kumar raman1920. It was probably the first contemporary scientific documentation on Mridangam. In this short writeup, they explained their observations of harmonicity on the drum head of an unspecified Indian percussion instrument. However, it will become apparent in Raman's later publications that this drum head is common to a class of Indian percussion instruments and not just one. Even though he authored two other articles raman1922, raman1927, it was in his 1934 paper raman that he rigorously described his observations on the harmonic drum heads. Focusing on the Mridangam and the Tabla heads, he explained their similarities and a few minor differences. He meticulously explains the different parts and their functions on the head. An interesting note is his remark on using heavy wooden shells for the instruments. He thinks that a massive frame will help sustain the membrane vibrations, owing to less energy transfer, and thus gives a prolonged tonal character. It should be something to ponder for someone designing lightweight alternatives for the shell.

A significant part of the paper, though, is concerning the membrane overtones, their excitation with finger strokes and the harmonic relationships. Raman used Tabla drum head for demonstrations (not explicitly mentioned), but the observations still hold for the karanai membrane of Mridangam. All the tones come from the first nine natural modes, and higher modes get excessively damped by the annular flaps. While the mode shapes are similar to a regular circular membrane, its frequencies are integer multiples of the fundamental. Some of these form degenerate pairs to give five distinct tones for the drum head. The first tone is the fundamental, having no nodal circles or diameters, while the second one is from the mode having a nodal diameter. Higher harmonics come from the superposition of two or more overtones. Mode with one nodal circle and mode with two nodal diameters constitute the third harmonic. At the same time, the fourth harmonic combines mode with one each nodal circle and diameter, and one with three nodal diameters. It is tricky to get the fifth harmonic except in precisely made instruments. This tone is a superposition of three modes: one with four nodal diameters, one with two nodal diameters and a nodal circle, and one with two nodal circles. The paper presents comprehensive figures showing Chladni patterns obtained from the first four harmonics.

When Raman published his articles, a few others engaged in mathematically quantifying his observations. [Gho22] was the first to report such an attempt. He idealised the harmonic membrane in two ways; in the first model, density varies inversely with radius, and in the second it varies inversely with the square of the radius. His assumptions are impractical as the membrane mass becomes infinite. Later [Rao38] also tried to model the membrane considering the density to vary inversely with the power of radial distance. Even though his model created a harmonic sequence, it was only with the modes having nodal diameters and was not in line with Raman's observations. It was [RS54] who finally came up with an adequate model to mimic the actual behaviour. Their idea was to consider karanai as a composite membrane having two concentric density regions, one for the black patch and another for the leather. This model successfully produced the harmonic sequence and degeneracies observed in the experiments. This paper will indeed follow their model for the karanai membrane. [Soo18] put forward for the first time the concept of concentric multi-density membranes for numerically modelling the thoppi drumhead and satisfactorily matched his experimental results. His numerical model is transformed into an analytical one here for analysing the thoppi.

[RS54] provide a model to extract the natural frequencies and mode shapes of the drum heads. But modelling the forced response is crucial for understanding its behaviour while playing. [Vod63] published a paper dealing with the forced vibrations of circular concentric multi-density membranes. His approach was to use the Laplace transform to arrive at the solutions. But the method of superposition of modes seems straightforward here since the mode shapes will be extracted. However, no documented literature on using modal superposition to solve multi-density membrane vibrations is available to the author's knowledge. The approach is followed, probably for the first time. Although the actual nature of the Mridangam strokes is transient, harmonic forces are considered in this paper due to the approximations used for analytically modelling the cavity. Interestingly, [Vod62] has also solved the free vibration problem of circular composite membranes subject to initial conditions.

Some simplifications are necessary to develop an analytical model for the acoustic cavity enclosed by the wooden shell and the drum heads. These involve rendering the complex profile inside the body to a simple double conical shape and considering the acoustics only in the axial direction (reasons detailed later). Propagation of waves in an arbitrarily variable section has been studied from the middle of the eighteenth century Eisner. Daniel Bernoulli, Euler, and Lagrange, all worked on forming the equation and made compelling efforts to arrive at the solution. In the early 1900s, [Lor16] and [Web19] worked independently and published papers on the theory of horns. The governing differential equation later became known as "Webster horn equation". According to Eisner, Rayleigh's article is more engrossing, and the title is not justified. The equation, however, has to be solved for a conical horn to get the axial acoustic pressure distribution in the Mridangam cavity. [Pil17] derived this solution using the Frobenius method as a part of her PhD thesis. Those functions will govern the acoustics for the two conical portions in the simplified cavity.

Modelling the interactions between the drum heads and the enclosed air is essential to meet the objectives of this study. As far as the cavity is concerned, the vibrating drum heads are boundary excitations at its two ends. Cavity response in the form of acoustic pressure produces a load on both of the membranes. [DV63] proposed a model to find the acoustic field inside a cavity for uniform wall motions using its rigid cavity modes. Their method though was unable to satisfy the continuity of velocity at the moving fluid-structure interfaces. [Gin10] gave a better formulation employing the same basis functions (rigid cavity modes) and Green's theorem on a one-dimensional waveguide excited by a single degree of freedom oscillator at one end. This model is an ideal match for the requirements in this paper and will be used to get the pressure field inside the Mridangam cavity due to membrane vibrations. It does not consider structural interaction, though.

Coupling the cavity acoustics with the membrane vibrations is challenging because the vibration variables are not present in the one-dimensional acoustic functions. Again, to the authors' knowledge, nobody has addressed this in the literature.

9.1.4 Objectives and Scope

The motivation for this study stems from Sooraj's conclusion that the membrane vibrations are independent on its own, and the acoustic cavity of Mridangam plays little role in sound generation. Research questions arising from his observations were mentioned previously in Section 9.1.2. Following objectives are outlined to answer these questions:

1. Develop analytical models for the acoustic cavity and the karanai and thoppi drum heads, and validate them with numerical models and experimental results from [Soo18]
2. Formulate a procedure for solving the coupled dynamics equations combining the acoustics and vibration models
3. Conduct shape sensitivity studies by introducing alternate geometries for the air cavity and understand the effects of the shape of Mridangam on membrane vibrations

The scope of the paper, while meeting these objectives, is limited to mathematical and numerical examination of vibro-acoustic interactions between the drum heads and the enclosed cavity. Other elements, like the wooden shell, are disregarded. The dynamics are idealised into linear models, overlooking any nonlinear effects that might be present. Also, all the analyses are harmonic in time due to limitations of the acoustic model.

9.2 Cavity Acoustics Model

The primary objective of this research is to investigate if the air cavity inside the Mridangam shell plays any role in the sound generated. This section describes the mathematical modelling of the cavity for this purpose. The inside surface of the shell along with the membranes defines the shape of the cavity. The original shell has a barrel-like shape, with cylindrical ends for connecting the membranes. As the actual barrel volume is difficult to describe mathematically, the analytical model approximates it as a diverging-converging conical horn. [Soo18] had carried out extensive numerical investigations on the Mridangam using both uncoupled and coupled vibro-acoustic models. He observed that only the axial modes of the cavity couple with the membrane vibrations. This insight from his study indicates that for studying the vibro-acoustics of a Mridangam, the three dimensional numerical model of the acoustic cavity can be replaced by a one-dimensional analytical model. The methodology of modelling, the process of extracting cavity modes and obtaining boundary excitation responses using modal superposition are explained in this section. Further, alternate geometries considered for the cavity for shape sensitivity analysis are introduced along with a comparison of their acoustic characteristics with the baseline geometry.

9.2.1 Cavity Approximation

Figure 9.5 shows the approximated geometry for the cavity, along with the actual shape. As mentioned, the original barrel-like shape is complex to represent mathematically; segments 2 and 3 together form the conical approximations for it. The karanai and thoppi membranes are respectively attached to the cylindrical portions 1 and 4. The largest diameter common to both the cones is equal to the largest diameter inside the Mridangam shell. The diameter of segment 1 is equal to the smallest diameter of segment 2 and equals the karanai membrane diameter. Likewise, segment 3's smallest diameter and segment 4's diameter are both equal to the diameter of the thoppi membrane. Since a one-dimensional model of the cavity is deemed to be adequate, the wave propagation is assumed to be planar and has no radial or azimuthal variations. The differential equations for plane wave propagation in cylindrical and conical segments are different. Hence, to represent the pressure variation along the overall length of the cavity, four piecewise functions are needed.

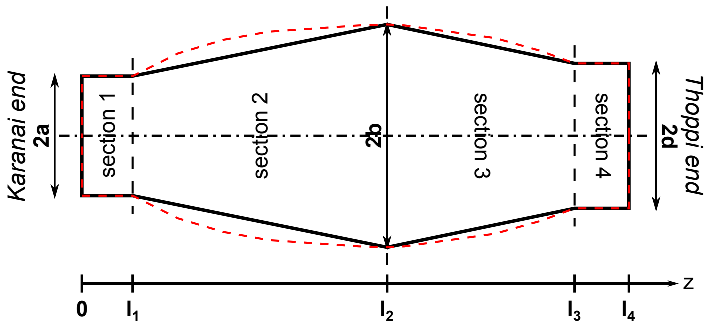


Figure 9.5: Cavity approximation for modelling along with the actual shape (in red)

The problem of plane wave propagation in a cylinder is well known kinsler. The governing 1D differential equation is

$$\frac{\partial^2 p}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}, \tag{9.1}$$

where $p(z,t)$ is the acoustic pressure, z the axial variable and t represents time. The speed of sound in the medium is given by $c = \sqrt{\kappa/\rho}$, with κ being the adiabatic bulk modulus and ρ being the mass density. For harmonic analysis ($p = P(z)\mathrm{e}^{j\omega t}$) with angular frequency ω, the general solution for 9.1 is

$$P(z) = A^{+}\mathrm{e}^{-jkz} + A^{-}\mathrm{e}^{jkz} \tag{9.2}$$

Here $k = \omega/c$ is the wavenumber, and the amplitudes A^+ and A^- depend on the boundary conditions. A^+ is the component for the forward travelling wave (positive z-direction) while A^- is that for the backward travelling wave (negative z-direction). Equation 9.2 is the generic equation for pressure variations in segmentss 1 and 4.

For the conical segments, one has to solve the Webster's horn equation Eisner to get the pressure distribution. It is a one-dimensional approximation for low-frequency wave propagation inside a rigid cavity with a variable cross-sectional area $S(z)$, written as Martin,

$$\frac{1}{S}\frac{\mathrm{d}}{\mathrm{d}z}\left(S\frac{\mathrm{d}P}{\mathrm{d}z}\right) + k^2 P = 0. \tag{9.3}$$

If the length of the cone is l, and inlet and outlet radii are a and b respectively, the area of cross-section varies as $S(z) = \pi(nz+a)^2$, where $n = (b-a)/l$. Substituting this expression into equation 9.3 and using the Frobenius method minu, the pressure distribution equation becomes

$$P(z) = \frac{1}{k(z+a/n)}\left(\tilde{A}^{+}\mathrm{e}^{-jk(z+a/n)} + \tilde{A}^{-}\mathrm{e}^{jk(z+a/n)}\right). \tag{9.4}$$

Amplitudes $\tilde{A}^+ = \frac{k}{n}\frac{a_0}{2} + j\frac{a_1}{2}$ and $\tilde{A}^- = \frac{k}{n}\frac{a_0}{2} - j\frac{a_1}{2}$ depend on the angular frequency ω, and the boundary conditions give the constants a_0 and a_1. Equation 9.4 is the generic equation for pressure variations in segments 2 and 3.

The pressure over the entire cavity is given by

$$P(z) = \begin{cases} P_1(z), & 0 \le z \le l_1 \\ P_2(z), & l_1 \le z \le l_2 \\ P_3(z), & l_2 \le z \le l_3 \\ P_4(z), & l_3 \le z \le l_4 \end{cases}. \tag{9.5}$$

Subscripts 1 and 2 denote respectively functions at cylindrical and conical segments near the karanai end. Subscripts 3 and 4 are respectively those at conical and cylindrical segments near the thoppi end. Lengths l_1, l_2, l_3 and l_4 are measured along the axis from the karanai end and are shown in Figure 9.5. From equations 9.2 and 9.4, these functions are written as

$$
\begin{aligned}
P_1(z) &= A_1^+ \mathrm{e}^{-jkz} + A_1^- \mathrm{e}^{jkz} \\
P_2(z) &= \frac{1}{k(z - l_1 + \frac{a}{n_1})} \left(A_2^+ \mathrm{e}^{-jk(z-l_1+\frac{a}{n_1})} + A_2^- \mathrm{e}^{jk(z-l_1+\frac{a}{n_1})} \right) \\
P_3(z) &= \frac{1}{k(z - l_2 + \frac{b}{n_2})} \left(A_3^+ \mathrm{e}^{-jk(z-l_2+\frac{b}{n_2})} + A_3^- \mathrm{e}^{jk(z-l_2+\frac{b}{n_2})} \right) \\
P_4(z) &= A_4^+ \mathrm{e}^{-jkz} + A_4^- \mathrm{e}^{jkz}.
\end{aligned} \tag{9.6}
$$

Here a is the radius of the karanai cylindrical segment and d is that of the thoppi cylindrical segment. b is the radius of the largest cross-sectional area, where the conical segments meet. In p_2, $A_2^+ = \frac{k}{n_1}\frac{a_0}{2} + j\frac{a_1}{2}$, $A_2^- = \frac{k}{n_1}\frac{a_0}{2} - j\frac{a_1}{2}$ and $n_1 = \frac{b-a}{l_2-l_1}$. Similarly, in p_3, $A_3^+ = \frac{k}{n_2}\frac{b_0}{2} + j\frac{b_1}{2}$, $A_3^- = \frac{k}{n_2}\frac{b_0}{2} - j\frac{b_1}{2}$ and $n_2 = \frac{d-b}{l_3-l_2}$. A_1^+, A_1^-, $a_0, a_1, b_0, b_1, A_4^+$ and A_4^- are constants to be determined. Note that while deriving equation 9.4, the origin for the area function $S(z)$ is at the beginning of the cone. Hence, to have a common coordinate z, the origin in p_2 and p_3 are translated to the left by l_1 and l_2 respectively. Euler's momentum equation $\rho\frac{\partial u}{\partial t} = -\frac{\partial p}{\partial z}$ gives the acoustic particle velocity $u\,(= U(z)\mathrm{e}^{j\omega t})$. Like the pressure function, $U(z)$ is also piecewise in the domain of the cavity. Its subfunctions are as follows.

$$
\begin{aligned}
U_1(z) &= \frac{1}{\rho c}(A_1^+ \mathrm{e}^{-jkz} - A_1^- \mathrm{e}^{jkz}) \\
U_2(z) &= \frac{1}{\rho c k(z - l_1 + \frac{a}{n_1})} \Bigg(A_2^+ \left(1 - j\frac{1}{k(z - l_1 + \frac{a}{n_1})}\right) \mathrm{e}^{-jk(z-l_1+\frac{a}{n_1})} \\
&\qquad - A_2^- \left(1 + j\frac{1}{k(z - l_1 + \frac{a}{n_1})}\right) \mathrm{e}^{jk(z-l_1+\frac{a}{n_1})} \Bigg) \\
U_3(z) &= \frac{1}{\rho c k(z - l_2 + \frac{b}{n_2})} \Bigg(A_3^+ \left(1 - j\frac{1}{k(z - l_2 + \frac{b}{n_2})}\right) \mathrm{e}^{-jk(z-l_2+\frac{b}{n_2})} \\
&\qquad - A_3^- \left(1 + j\frac{1}{k(z - l_2 + \frac{b}{n_2})}\right) \mathrm{e}^{jk(z-l_2+\frac{b}{n_2})} \Bigg) \\
U_4(z) &= \frac{1}{\rho c}(A_4^+ \mathrm{e}^{-jkz} - A_4^- \mathrm{e}^{jkz})
\end{aligned} \tag{9.7}
$$

9.2.2 Natural Frequencies and Modes

Now that the governing equations are in place, it can be used to find the natural frequencies and mode shape functions of the Mridangam cavity. [Soo18] did measurements on an actual Mridangam which form the basis for the dimensions of the cavity approximation; Table 9.1 gives these values. Air is the medium assumed, and its density ρ and bulk modulus κ are taken as 1.225 kg/m^3 and 142 kPa respectively. The motive is to find the natural frequencies and modes for rigid boundaries at the ends. These modes will later help in getting solutions to boundary excitations on the cavity. Moreover, it is an exercise to compare the approximate model with the actual cavity.

Table 9.1: Parameters from measurements of an actual mridangam

Parameter	Length(m)
a	0.066
b	0.123
d	0.08
l_1	0.056
l_2	0.337
l_3	0.544
l_4	0.604

Observe that there are eight unknowns in equation 9.6. Hence, it requires eight equations to form the eigenvalue problem. Rigid walls imply that the acoustic velocity is zero at the boundaries. It means that at the karanai end, $U_1(0) = 0$, and at the thoppi end, $U_4(l_4) = 0$. Ensuring continuity of pressure and its gradient at the three interfaces between the segments give the other six equations. At the karanai end cylinder-cone interface, $P_1(l_1) = P_2(l_1)$ and $U_1(l_1) = U_2(l_1)$. At the interface between the cones, $P_2(l_2) = P_3(l_2)$ and $U_2(l_2) = U_3(l_2)$. Finally, at the thoppi end cone-cylinder interface, $P_3(l_3) = P_4(l_3)$ and $U_3(l_3) = U_4(l_3)$. This set of boundary and compatibility conditions together, when written in matrix form, give

$$[CM]\{V\} = 0, \tag{9.8}$$

where $V = \{A_1^+, A_1^-, a_0, a_1, b_0, b_1, A_4^+, A_4^-\}^T$ is the vector of constants and CM is its coefficient matrix. Equation 9.8 has solutions when the determinant of CM is zero. Values of k for which the determinant is zero give the natural frequencies of the cavity. The nullspace of CM for every such k provides the constants V which in turn gives the corresponding mode shape function P_r.

Symbolic computation software Mathematica Mathematica is used for solving equation 9.8. It obtains the coefficient matrix from the boundary and compatibility equations and then finds the roots for which its determinant is zero. The first six natural frequencies are taken from these roots. The nullspace of the coefficient matrix after plugging in the frequencies gives the corresponding vector of constants which are then used in equation 9.6 to get the piecewise mode shape functions.

A 3D finite element (FE) model is also developed based on the geometry used for analytical approximation. This model will verify the accuracy of the plane wave assumption made. Commercial FE analysis software ABAQUS ABAQUS is used for building and analysing the model. The geometry is made using the same dimensions as given in Table 9.1. Material properties, bulk modulus and density of air, are entered in the property module. While creating the step with procedure type frequency, the default Lanczos eigensolver is selected, with the maximum frequency of interest set to 1700 Hz. ABAQUS by default assumes fixed boundaries for acoustic analysis, and hence no boundary conditions need to be given. The mesh uses hexahedral AC3D20 acoustic elements. It is a 20 node brick element with quadratic interpolation. For a given number of degrees of freedom (DOF), second-order elements are more accurate than first-order elements. Average internodal interval in the axial and radial directions given is 5 mm.

Table 9.2: Natural frequencies (in Hz) of Mridangam air cavity

Axial mode	Analytical	Finite element	Experimental*
1	354.5	351.3	370
2	575.1	569.6	628
3	853.8	847.0	893
4	1123	1113	1152
5	1402	1390	-
6	1690	1677	-

* Values from [Soo18].

Table 9.2 compares the analytical frequencies obtained with the numerical results from the ABAQUS simulation. The table also shows the natural frequencies from experiments carried out by [Soo18]. The slight difference in the experimental data is due to the approximations made for the analytical and ABAQUS models. Figure 9.6 compares the mode shapes obtained from the analytical method with the numerical simulation results. Note that since the FE model is three dimensional, non-axial modes too are present in the output. Both the analytical and FE mode shapes are normalised.

9.2.3 Boundary Excitation

After establishing that the modes of the mathematical model are excellent approximations to the actual cavity, these modes are used to get its boundary excitation responses. The karanai and thoppi membranes are attached to the cavity at its boundaries. Vibrations of these membranes induce acoustic particle accelerations at the interface. These accelerations cause the pressure to build up inside the cavity. The pressure, in turn, acts as loading on the membranes. This coupled behaviour is studied later in Section 9.4. Thus it is important to

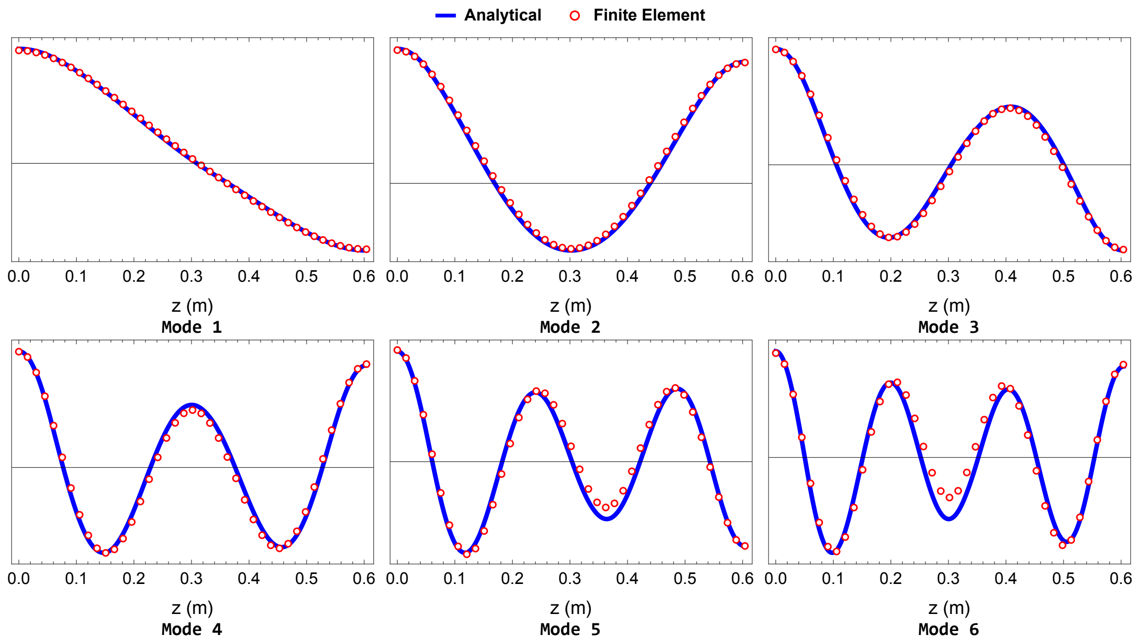


Figure 9.6: Pressure plots for axial modes of Mridangam cavity

model how pressure varies inside the cavity because of these boundary fluctuations. The modelling is done based on a formulation by [Gin10], using Green's theorem and modal summation. The formulation assumes piston-like motion of walls and does not consider structural interaction.

9.2.3.1 Mathematical Formulation

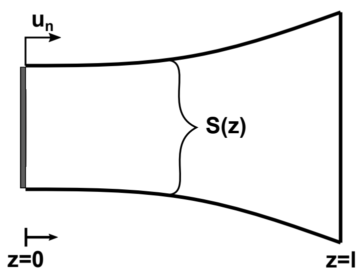


Figure 9.7: Cavity with variable cross-section excited at a boundary

To arrive at the expression for pressure response, consider an air cavity with a variable cross-sectional area $S(z)$ as shown in Figure 9.7. Assume that either/both the walls at the boundaries ($z = 0, l$) execute uniform, simple harmonic motion. The one-dimensional governing equation for acoustic pressure in this cavity, as mentioned earlier, is

$$\frac{1}{S}\frac{\mathrm{d}}{\mathrm{d}z}\left(S\frac{\mathrm{d}P}{\mathrm{d}z}\right) + k^2 P = 0, \qquad \text{(9.3 revisited)}$$

where $k = \Omega/c$, for the driving angular frequency Ω. The rigid wall modes P_r of this cavity satisfy

$$\begin{aligned} &\frac{1}{S}\frac{\mathrm{d}}{\mathrm{d}z}\left(S\frac{\mathrm{d}P_r}{\mathrm{d}z}\right) + k_r^2 P_r = 0, \\ &\left.\frac{\mathrm{d}P_r}{\mathrm{d}z}\right|_0 = \left.\frac{\mathrm{d}P_r}{\mathrm{d}z}\right|_l = 0, \end{aligned} \qquad (9.9)$$

where k_r is the wavenumber for the r^{th} mode. Pressure in equation 9.3 can be written as a linear expansion of the mode shape functions as

$$P(z)=\sum_{r=1}^{R}\alpha_r P_r(z) \tag{9.10}$$

To find the coefficients α_r, consider the following equation similar to Green's second identity.

$$\begin{aligned}\int_0^l\left(P_r\frac{1}{S}\frac{\mathrm{d}}{\mathrm{d}z}S\frac{\mathrm{d}P}{\mathrm{d}z}-P\frac{1}{S}\frac{\mathrm{d}}{\mathrm{d}z}S\frac{\mathrm{d}P_r}{\mathrm{d}z}\right)dz=P_r\frac{\mathrm{d}P}{\mathrm{d}z}\Big|_0^l-P\frac{\mathrm{d}P_r}{\mathrm{d}z}\Big|_0^l-\\ \int_0^l\left((\frac{\mathrm{d}}{\mathrm{d}z}\frac{P_r}{S})S\frac{\mathrm{d}P}{\mathrm{d}z}\right)dz+\int_0^l\left((\frac{\mathrm{d}}{\mathrm{d}z}\frac{P}{S})S\frac{\mathrm{d}P_r}{\mathrm{d}z}\right)dz\end{aligned} \tag{9.11}$$

From Euler's momentum equation, and equations 9.3 and 9.9, this reduces to

$$\begin{aligned}(k_r^2-k^2)\int_0^l PP_r dz=-j\rho\Omega\left(P_r(l)U(l)-P_r(0)U(0)\right)-\\ \int_0^l\left((\frac{\mathrm{d}}{\mathrm{d}z}\frac{P_r}{S})S\frac{\mathrm{d}P}{\mathrm{d}z}-(\frac{\mathrm{d}}{\mathrm{d}z}\frac{P}{S})S\frac{\mathrm{d}P_r}{\mathrm{d}z}\right)dz.\end{aligned} \tag{9.12}$$

However, this equation does not account for any energy dissipation. For that, modal damping, with damping ratio ζ, is introduced. The final equation becomes

$$\begin{aligned}(k_r^2-k^2+j2\zeta k_r k)\int_0^l PP_r dz=-j\rho\Omega\left(P_r(l)U(l)-P_r(0)U(0)\right)-\\ \int_0^l\left((\frac{\mathrm{d}}{\mathrm{d}z}\frac{P_r}{S})S\frac{\mathrm{d}P}{\mathrm{d}z}-(\frac{\mathrm{d}}{\mathrm{d}z}\frac{P}{S})S\frac{\mathrm{d}P_r}{\mathrm{d}z}\right)dz.\end{aligned} \tag{9.13}$$

Based on the number of modes taken (R), the above equation gives a set of R algebraic equations for the unknowns α_r. The solution of these equations gives the coefficients which, when substituted into equation 9.10 gives the pressure function. Ensuring continuity of the velocities at the interface between the walls and the medium gives the inputs $U(0)$ and $U(l)$.

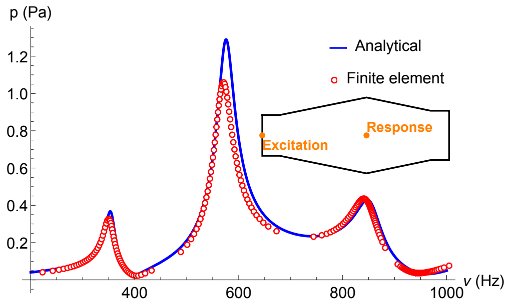
9.2.3.2 Mridangam Cavity Excitation

Equation 9.13 can be expanded to the full Mridangam cavity model presented earlier. The issue is that the functions P, P_r, U and S are piecewise along the length. Therefore, the integration is to be carried out in each of the domains $[0,l_1]$, $[l_1,l_2]$, $[l_2,l_3]$ and $[l_3,l_4]$. The subfunctions in these domains for the modes P_r follow equation 9.6, with the constants evaluated. Pressure P, as in equation 9.10, is written as a summation of these piecewise modes. The velocity inputs for the equation become $U_1(0)$ at the karanai end and $U_4(l_4)$ at the thoppi end. The piecewise area function for the cavity is

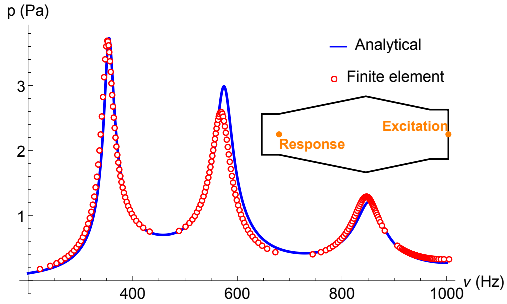
$$S(z)=\begin{cases}S_1(z)=\pi a^2, & 0\le z\le l_1\\ S_2(z)=\pi(a+n_1(z-l_1))^2, & l_1\le z\le l_2\\ S_3(z)=\pi(b+n_2(z-l_2))^2, & l_2\le z\le l_3\\ S_4(z)=\pi d^2, & l_3\le z\le l_4\end{cases} \tag{9.14}$$

Figure 9.8 presents the results after solving equation 9.13 for boundary excitations on the Mridangam cavity. Two cases are shown here. In case (a), karanai end is excited with uniform velocity 1 mm/s, keeping the thoppi end fixed ($U_1(0)=1$ mm/s and $U_4(l_4)=0$). The response shown is at $z=l_2$. For the second case (b), karanai end is fixed, and the thoppi end is excited with the same velocity ($U_1(0)=0$ and $U_4(l_4)=1$ mm/s). The response shown in this case is at $z=l_1$. Using six modes of the cavity extracted in the previous section reduces the truncation error in the driving frequency range 200 Hz to 1000 Hz. The damping ratio for each mode is 3%. The analytical calculations are done in Mathematica to get the coefficients.

The FE model used for validation is the same as used in the previous section. An additional step with procedure type - *Steady-state dynamics, Modal* - is required in this instance. In the step dialogue box, the required frequency range is set along with the direct modal damping data for the modes. Note that non-axial modes are also present in this range for the FE model. The acoustic load for this step is input by choosing



(a) Excitation at Karanai, response at $z = l_2$



(b) Excitation at Thoppi, response at $z = l_1$

Figure 9.8: Mridangam cavity response to boundary excitations

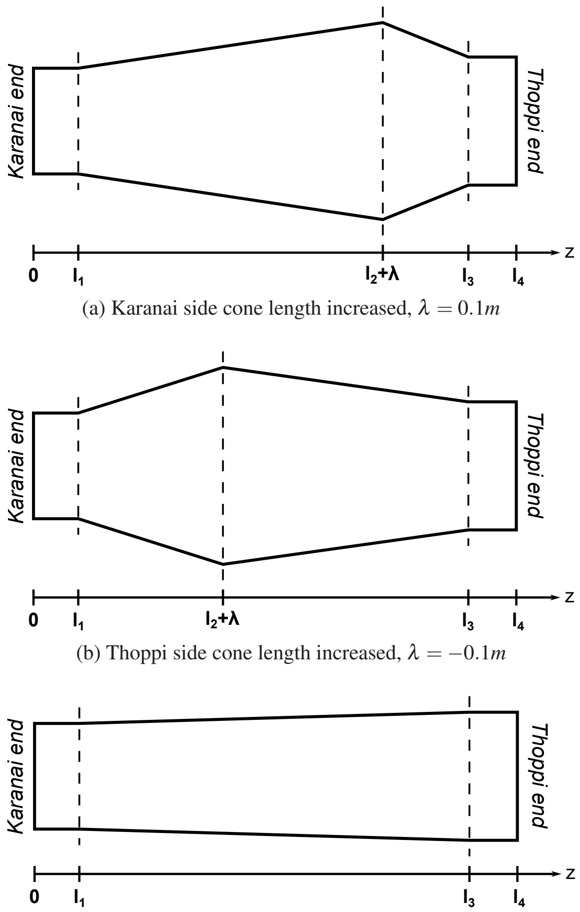
the category '*Acoustic*', and then '*Inward volume acceleration*' from the types. After selecting the nodes of interest, volume acceleration per node (area×velocity / number of nodes) is entered in the magnitude box. A tabular data provided in the amplitude box accounts for the frequency dependence ($2\pi\times$frequency) of magnitude. The comparison shows that the analytical formulation works quite satisfactorily for uniform boundary excitations. There is a slight climb in the mathematical response near the second peak. A possible reason would be the variation in axial pressure between the 1D analytical model and the 3D FE model due to near field effects at the boundary of excitation. More of this will be explained in Section 9.4.

9.2.4 Alternate Geometries for Cavity

To further understand the effect of the cavity on the sound generation by the Mridangam, alternate geometries which are modifications of the baseline model are considered. This section will study the acoustic characteristics of these geometries while later in section 9.4, they will be combined with the membranes to study the coupled behaviour. This cavity shape sensitivity study hopes to realise if alternate, simple geometries can be considered for the Mridangam. See Figure 9.9 for the alternate geometries taken for assessment. In geometry one (9.9a), the largest diameter segment translates nearer to the thoppi side by $\lambda = 0.1$ m (about 20% length of conical segments). The karanai cone length increases by around 35% as a result. In the second geometry (9.9b), thoppi cone length increases by 48% when the big-diameter segment moves towards the karanai side with $\lambda = -0.1$ m. The third one (9.9c) replaces the two cones with a single one connecting the cylindrical segments. The overall length and other parameters in all three cases retain the original cavity values.

The pressure (P), velocity (U) and area (S) functions, for both geometry one and two, follow equations 9.6, 9.7 and 9.14 respectively, except for the fact that the length l_2 changes to $l_2 + \lambda$. For the single cone geometry, only three segments are present: two cylindrical and one conical. Therefore, pressure, velocity and area functions all will have only three subfunctions as given below. Functions with subscript 1 belong to the karanai end cylindrical segment ($0 \leq z \leq l_1$), subscript 2 belong to the conical segment ($l_1 \leq z \leq l_3$) and subscript 3 represents those at the thoppi end cylindrical segment ($l_3 \leq z \leq l_4$).

$$\begin{aligned} P_1(z) &= A_1^+ \, \mathrm{e}^{-jkz} + A_1^- \, \mathrm{e}^{jkz} \\ P_2(z) &= \frac{1}{k(z - l_1 + \frac{a}{n_1})} \left(A_2^+ \, \mathrm{e}^{-jk(z-l_1+\frac{a}{n_1})} + A_2^- \, \mathrm{e}^{jk(z-l_1+\frac{a}{n_1})} \right) \\ P_3(z) &= A_3^+ \, \mathrm{e}^{-jkz} + A_3^- \, \mathrm{e}^{jkz} \end{aligned} \tag{9.15}$$



(a) Karanai side cone length increased, $\lambda = 0.1m$

(b) Thoppi side cone length increased, $\lambda = -0.1m$

(c) Single Cone Cavity

Figure 9.9: Alternate geometries for shape sensitivity studies

$$
\begin{aligned}
U_1(z) &= \frac{1}{\rho c}(A_1^+ \mathrm{e}^{-jkz} - A_1^- \mathrm{e}^{jkz}) \\
U_2(z) &= \frac{1}{\rho c k(z-l_1+\frac{a}{n_1})}\left(A_2^+\left(1-j\frac{1}{k(z-l_1+\frac{a}{n_1})}\right)\mathrm{e}^{-jk(z-l_1+\frac{a}{n_1})}\right. \\
&\qquad \left. -A_2^-\left(1+j\frac{1}{k(z-l_1+\frac{a}{n_1})}\right)\mathrm{e}^{jk(z-l_1+\frac{a}{n_1})}\right) \\
U_3(z) &= \frac{1}{\rho c}(A_3^+ \mathrm{e}^{-jkz} - A_3^- \mathrm{e}^{jkz})
\end{aligned}
\tag{9.16}
$$

$$
\begin{aligned}
S_1(z) &= \pi a^2 \\
S_2(z) &= \pi(a+n_1(z-l_1))^2 \\
S_3(z) &= \pi d^2
\end{aligned}
\tag{9.17}
$$

Here $A_2^+ = \frac{k}{n_1}\frac{a_0}{2} + j\frac{a_1}{2}$, $A_2^- = \frac{k}{n_1}\frac{a_0}{2} - j\frac{a_1}{2}$ and $n_1 = \frac{d-a}{l_3-l_1}$. The constants to be evaluated are A_1^+, A_1^-, a_0, a_1, A_3^+ and A_3^-.

9.2.4.1 Natural Frequencies

The first step is to extract the natural frequencies of these different geometries and compare them with the baseline. The procedure followed in Section 9.2.2 is followed here too for extracting the frequencies and the mode shape functions. Table 9.3 presents the natural frequencies obtained from the analytical models. The numbers in the brackets indicate the percentage deviation from the baseline frequency. As it is already proved that the 1D mathematical model is a good approximation for the 3D FE model, no validation is performed for the new geometries. Note that the single cone fundamental shows 20% deviation while all other modes are within 6% of the original. It has to be seen later, what this shift implies for the coupled dynamics.

Table 9.3: Natural frequencies of alternate geometries

Mode	Frequency (Hz)			
	Baseline	$\lambda = 0.1\ m$	$\lambda = -0.1\ m$	Single cone
1	354.5	336.0 (-5.2)	349.6 (-1.4)	283.4 (-20)
2	575.1	607.0 (+5.5)	591.5 (+2.8)	564.3 (-1.9)
3	853.8	845.1 (-1.0)	846.2 (-0.9)	845.7 (-0.9)
4	1123	1115 (-0.7)	1131 (+0.7)	1127 (+0.4)
5	1402	1407 (+0.4)	1397 (-0.4)	1409 (+0.5)
6	1690	1689 (-0.0)	1690 (-0.0)	1691 (+0.1)

Brackets indicate percentage deviation from original.

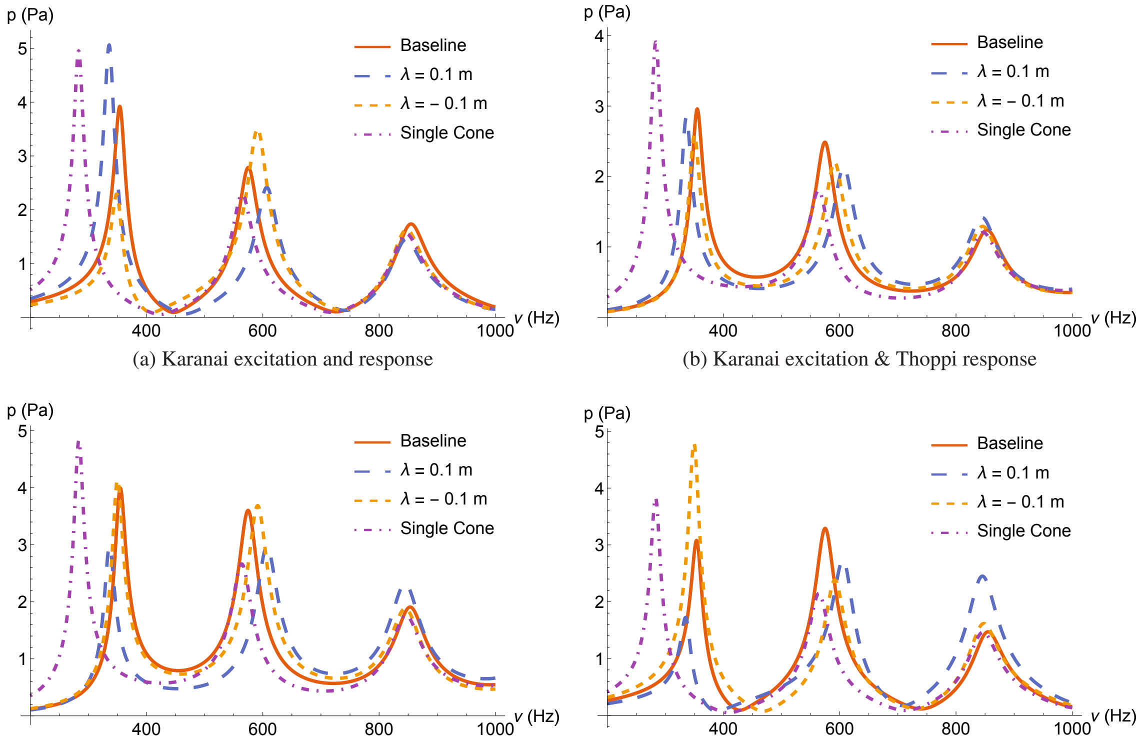
9.2.4.2 Boundary Excitation Responses

Next, the response of the geometries to excitations at the karanai and thoppi boundaries are studied. Equation 9.13 gives the pressure responses for the geometries using the frequencies (k_r) and modes (P_r) extracted above. Damping ratio ζ and properties of air, ρ and κ, are the same as taken before. The frequency range considered is between 200 Hz and 1000 Hz. Writing pressure as a summation of modes, Mathematica solves for the coefficients from the system of integral equations. The integrals are evaluated in four parts for geometries 9.9a and 9.9b, while it is done in three parts for the single cone geometry 9.9c. Input velocity is arbitrarily chosen to be 1 mm/s. Only either one of the ends is excited in an instance, keeping the other fixed. It is because a Mridangam player taps either karanai or thoppi, but not both together, at any point in time. The points of interests for calculating the responses are the boundaries itself since these responses act as loads on the membranes when coupled. Figures 9.10a and 9.10b show the responses of the geometries when the karanai end is excited keeping the thoppi end fixed ($U_1(0) = 1$ mm/s and $U_4(l_4) = 0$). The bottom row (9.10c and 9.10d) shows the results when the thoppi end is excited, keeping the karanai fixed ($U_1(0) = 0$ and $U_4(l_4) = 1$ mm/s). Figures 9.10a and 9.10c give the karanai end responses ($z = 0$) while Figures 9.10b and 9.10d give the thoppi end responses ($z = l_4$).

Near the first natural frequency region, the single cone geometry exhibits higher peak pressure when compared to the baseline geometry as well as the other two modifications. This behaviour could indicate possibly enhanced vibro-acoustic coupling with the single cone geometry. Moving the largest cross-section area towards karanai (geometry 9.9b) reduces the first mode peak pressure amplitude at the karanai boundary when the end is excited (Figure 9.10a, $\lambda = -0.1$ m). At the same time, there is an increase in the pressure at the thoppi side for excitation at the boundary (Figure 9.10d, $\lambda = -0.1$ m). On moving the cross-section towards thoppi (geometry 9.9a), the trends get reversed from the previous case (Figure 9.10a and 9.10d, $\lambda = 0.1$ m). It is generally a weak response near the second mode region except for $\lambda = -0.1$ m curve in Figure 9.10a. $\lambda = 0.1$ m curve gives a spike near the third mode in Figure 9.10d. Since predominant membrane vibration modes are below 500 Hz, the first cavity modes of each geometry are more likely to be coupled with the membranes. It looks like the single cone could affect the membrane vibrations and hence the sound generated. For the other two geometries, there is a distinct increase in pressure, on boundaries opposite to the direction of change in length l_2. The increased pressure could lead to more coupling when compared to the baseline, and this needs further investigation. Also, cross-coupling between membranes might reduce as the largest diameter segment moves to either side of the baseline location.

9.3 Membrane Vibration Models

Membranes are the primary source of sound in a Mridangam. The two circular membranes, karanai and thoppi, are connected to the ends of the wooden shell. Generally, one plays the karanai with the right hand and the thoppi with the left. Karanai has a central mass loading which appears as a black patch, and it makes



(a) Karanai excitation and response

(b) Karanai excitation & Thoppi response

(c) Thoppi excitation & Karanai response

(d) Thoppi excitation and response

Figure 9.10: Effect of cavity shape on the responses to boundary excitations

the membrane harmonic. The thoppi, on the other hand, is not harmonic and is used to play the lower 'bass' frequencies. The introductory section explained the intricate construction of these membranes. However, for modelling, they are simplified as concentric multi-density membranes. In their paper, [RS54], idealised the karanai as a single layer of a membrane with two concentric uniform density regions.Their mathematical model was able to imitate the observations made by [Ram34] convincingly. This paper uses the same model. The thoppi analytical model is an extension of the karanai model from two to three concentric uniform density regions. Using the properties from [Soo18], frequencies and modes are found and compared with experimental and numerical results. Also a mathematical formulation to find the harmonic response of such composite membranes to external forces has been developed by Harikanth (2020). This formulation is used for the coupled vibro-acoustic studies in the next section; one can refer to Sooraj (2018) and Harikanth (2020) for details of the membrane response calculations.

9.4 Coupling of Cavity and Membrane Models

The development of mathematical models for the cavity (in detail) and the leather membranes of Mridangam (in brief) were presented in Sections 9.2 and 9.3, respectively. This section explains how these models are used to study their coupled dynamics. This way, the effect of the cavity on the sound produced by the membranes can be quantified. Note that the membrane models are two-dimensional while the cavity acoustic model is one-dimensional. Thus, the 2D vibration fields have to be mapped to a single-dimensional velocity excitation which will act on the cavity. Also, due to the assumption of plane wave propagation in the cavity, acoustic pressure does not vary across the membranes. Membrane vibrations with and without the cavity loading will be compared first. Later, these results will be presented along with those from the coupled models using alternate cavity geometries introduced in Section 9.2. These analyses rely more on coupled finite element models rather than analytical models for reasons which will be explained later in this section. Although the nature of the actual Mridangam strokes is transient, the present work is limited to harmonic analysis. Another restriction made is that the study is carried only for strokes on the karanai membrane. However, the approach is not different to adapt for thoppi side excitation.

9.4.1 Governing Equations

Equations governing the coupled dynamics of the cavity and the membranes are similar to those derived in previous sections. The presence of the coupling terms is the only difference. Equation 9.13 governing the boundary excitation response of the cavity will now have the additional input terms from the membrane vibration velocities. Similarly, in equations governing the forced response of the membranes (not shown in the paper), the acoustic pressure on the boundaries of the cavity are an additional input apart from the finger stroke based mechanical excitation. The velocity terms couple the membranes to the cavity while the acoustic pressure couples the cavity to the membranes. These equations, when solved together, give the vibro-acoustic response of the Mridangam.

The piecewise equations for the coupled harmonic forced vibrations of the karanai membrane are given by

$$\begin{aligned} \nabla^2\Psi_1 + k_1^2\Psi_1 &= \frac{P_1(r,\theta)+P(0)}{\tau_1} \qquad 0 \le r \le a_1 \\ \nabla^2\Psi_2 + k_2^2\Psi_2 &= \frac{P_1(r,\theta)+P(0)}{\tau_1} \qquad a_1 \le r \le a \end{aligned} \tag{9.18}$$

While for the thoppi membrane, the equations are

$$\begin{aligned} \nabla^2\Psi_3 + k_3^2\Psi_3 &= \frac{P_2(r,\theta)+P(l_4)}{\tau_2} \qquad 0 \le r \le d_1 \\ \nabla^2\Psi_4 + k_4^2\Psi_4 &= \frac{P_2(r,\theta)+P(l_4)}{\tau_2} \qquad d_1 \le r \le d_2 \\ \nabla^2\Psi_5 + k_5^2\Psi_5 &= \frac{P_2(r,\theta)+P(l_4)}{\tau_2} \qquad d_2 \le r \le d \end{aligned} \tag{9.19}$$

Here P_1 and P_2 are the external pressure fields acting on the karanai and thoppi respectively. At the same time, $P(0)$ and $P(l_4)$ are the additional terms to the membrane vibration equations, to account for the pressure loading from the cavity at the karanai and thoppi ends respectively. The cavity load does not vary across the section, unlike the external fields which have distributions. $P(0)$ and $P(l_4)$ arise from the integral equation for the response of the cavity to membrane vibrations. This equation, derived in Section 9.2, gives the pressure distribution due to boundary excitations.

$$\begin{aligned} (k_r^2 - k^2 + j2\zeta k_r k)\int_0^{l_4} PP_r dz = -j\rho\Omega\left(P_r(l_4)U(l_4) - P_r(0)U(0)\right) - \\ \int_0^{l_4}\left((\frac{\mathrm{d}}{\mathrm{d}z}\frac{P_r}{S})S\frac{\mathrm{d}P}{\mathrm{d}z} - (\frac{\mathrm{d}}{\mathrm{d}z}\frac{P}{S})S\frac{\mathrm{d}P_r}{\mathrm{d}z}\right) dz \end{aligned} \tag{9.20}$$

The boundary terms, $U(0)$ and $U(l_4)$, are the coupling terms in this equation arising from the membrane vibrations. They denote the velocities at the karanai and thoppi boundaries. Since the membrane vibrations are functions of both r and θ, transforming them to a single vibration value is necessary.

The transformation used is averaging the 2D velocity distribution over the membrane area. Sooraj's observation forms the rationale for this approach. In his investigation, he found out that only axisymmetric modes of the membranes seemed to couple with the acoustic pressure in the cavity. This observation would suggest that the best approach would be to do an area average of the velocity field. This way, modes having nodal diameters would not contribute, leaving only axisymmetric modes to couple with the cavity. Other options explored include taking root mean square or local maxima. Using the root mean square (RMS) velocity will not eliminate the non-axisymmetric modes as opposed to area averaging and would influence building pressure in the cavity, which otherwise may not be present. Even so, all the options were analysed and compared with the results from the finite element model. It also turned out that area averaging predicts more accurately the pressure inside the cavity. Hence it is chosen as the appropriate way to convert the vibrations into a scalar value.

The area-averaged membrane vibrations at the karanai and the thoppi ends are,

$$U(0) = \frac{j\Omega}{\pi a^2}\left(\int_0^{2\pi}\int_0^{a_1}\Psi_1 r dr d\theta + \int_0^{2\pi}\int_{a_1}^{a}\Psi_2 r dr d\theta\right) \tag{9.21}$$

$$U(l_4) = \frac{j\Omega}{\pi d^2}\left(\int_0^{2\pi}\int_0^{d_1}\Psi_3 r dr d\theta + \int_0^{2\pi}\int_{d_1}^{d_2}\Psi_4 r dr d\theta + \int_0^{2\pi}\int_{d_2}^{d_3}\Psi_5 r dr d\theta\right) \tag{9.22}$$

The functions, Ψ_i, (i=1 to 5), are solutions to equations 9.18 and 9.19. The above equations (9.18 to 9.22) together form the mathematical set for the coupled dynamics. Note that all the variations are assumed to be harmonic in time t, with the driving frequency Ω.

9.4.2 Solving Coupled Equations

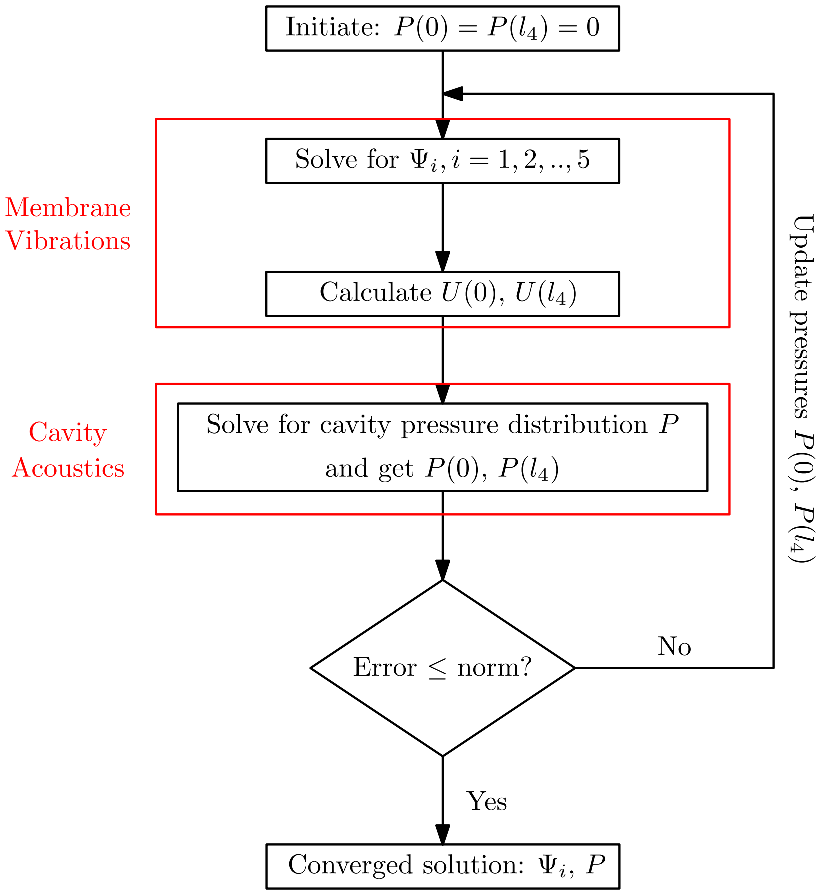


Figure 9.11: Flowchart for iteratively solving coupled equations for a frequency step

To simultaneously solve the set of coupled mathematical equations is quite challenging. A more convenient way is to solve them iteratively, and this approach is followed in this paper. Figure 9.11 shows the flowchart of the iteration procedure. For initialisation, the acoustic pressure loads of the cavity on the membranes is ignored, i.e., $P(0)$ and $P(l_4)$ are initially set to zero. The membrane vibration equations are then the same as the uncoupled equations. Using the solution procedure from [Har20], the responses are obtained. From these, the area-averaged velocities $U(0)$ and $U(l_4)$ can be calculated using equations 9.21 and 9.22. These velocities are then input to equation 9.20, which is solved as described in section 9.2.3.2. The solution gives the pressure distribution to get $P(0)$ and $P(l_4)$, which serve as the acoustic loads on the membranes for the next iteration. The process is repeated with updated right-hand sides in the membrane vibration equations. Membrane vibrations are then solved for and averaged velocities calculated. These velocities are now input into equation 9.20; acoustic pressures $P(0)$ and $P(l_4)$ are updated for the next iteration. The convergence of these cavity pressures is checked before starting a new iteration. The solution is assumed to have converged for a particular frequency if the relative error in pressures between two successive iterations is less than the norm. With converged solutions for one frequency, the process is repeated for the next frequency step.

9.4.2.1 Membranes - Original Cavity Coupled Solutions

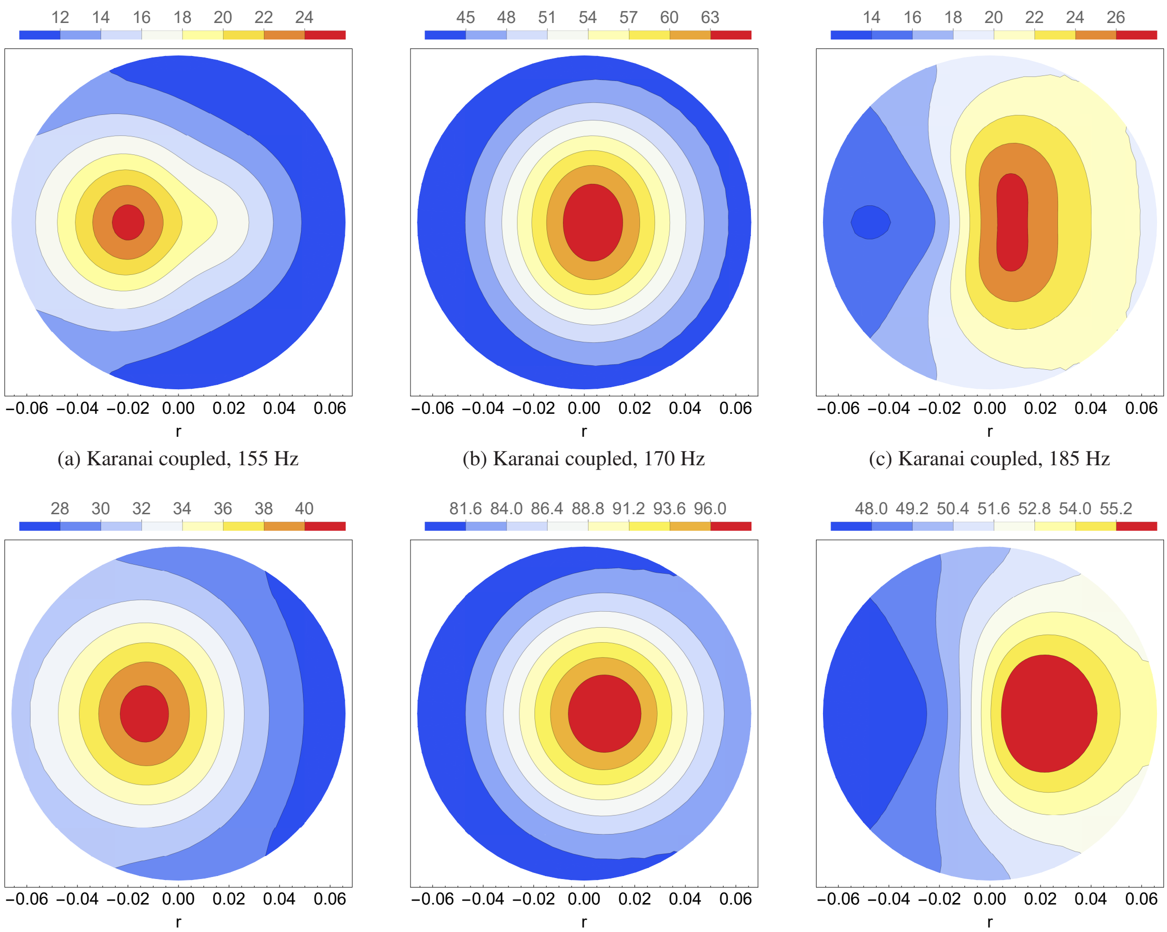
The coupled problem of the membranes and the cavity approximation (section 9.2.1) is approached using the procedure above. Since karanai is the predominant source of sound, the primary motive is to compare its coupled and uncoupled vibrations. In order to be consistent with comparisons, a point load of unit magnitude is applied on the karanai. Note that both P_1 and P_2 are never simultaneously non-zero. The reason for this is that the Mridangam players generally do not hit both the membranes together while playing. Thus P_2 is kept zero. Modes used for summation are the same as used in uncoupled individual cases. Modal damping is incorporated for all modes with a damping ratio of 0.03. Tensions τ_1 and τ_2 are the ones used for getting the modes, given in section 9.3. The first two modes of karanai fall between 100 and 400 Hz. Hence it is the frequency range chosen for comparison.

Symbolic programming for this coupled problem is done in Mathematica Mathematica. First, the uncoupled modes of the membranes and the cavity are individually extracted as described in the previous sections. Then, the integrals not constituting the inputs are evaluated; the left-hand side of the membrane vibration equations and terms in equation 9.13 except those with $U(0)$ and $U(l_4)$. Doing so decreases the computational effort within the iteration loop. To implement the flowchart shown in Figure 9.11 for a particular frequency, *While* loop in Mathematica is employed. The test for exiting this loop is the convergence criterion. The modulus of pressure at the karanai end ($|P(0)|$) should not have more than 1% error relative to the previous iteration to end the evaluation. Since all quantities are interconnected, the convergence of one value can imply that the displacement functions of membranes and the pressure distribution inside the cavity also converges. The *Do* function enables the above iteration to be run in steps of frequencies. The evaluation runs from 100 Hz to 400 Hz in frequency steps of 5 Hz. It is ensured through proper sign convention that the transverse displacement coordinates of the membranes and the axial coordinates of the cavity align together.

However, the above algorithm faces convergence issues near the first mode frequency of the thoppi membrane. The moduli of the complex coupling terms (end pressures $P(0)$ and $P(l_4)$, and the averaged velocities $U(0)$ and $U(l_4)$) keep shooting up with iterations at around 190 Hz, impeding convergence of the solution. Also, the argument of these values oscillates between quadrants for different iterations at around 190 Hz. For identifying the source of this issue, each membrane was individually coupled with the cavity, keeping the opposite end of cavity rigid. The membrane on the rigid side of the cavity is neglected from computations. Coupling karanai alone, keeping the thoppi end of the cavity rigid ($U(l_4) = 0$ in equation 9.20), the procedure gave solutions in the frequency range considered. However, by using thoppi alone ($U(0) = 0$), the very issue with the fully coupled problem reoccured; convergence did not happen near the thoppi fundamental frequency. As a result of this issue, finite element models are used for further study of the coupled problems.

The coupled numerical model is also built based on Sooraj's thesis. In ABAQUS, the parts used earlier for karanai, thoppi and the cavity are imported to a new model. These parts form an assembly, and they are put in their respective locations using position constraints. For the structural-acoustic fully coupled analysis, ABAQUS recommends using a surface-based approach rather than using acoustic-structural interface (ASI) elements ABAQUS. In this approach, tie constraints couple the structural accelerations and the acoustic pressures at the interfaces. One tie constraint joins the inner surface of karanai with the karanai end boundary surface on the cavity and another one couples the thoppi inner surface with the corresponding boundary surface on the cavity. The membrane surfaces serve as the master surfaces while the cavity surfaces serve as the slaves. The membranes are prestressed with edge tensions in step one. In the second, its peripheries are fixed. For the mode-based steady-state dynamic analysis, the coupled modes are extracted in the next step (*Frequency*), before the final *Steady-state dynamics, Modal* step. Frequency range 100 to 400 Hz demands modes up to 600 Hz to reduce errors, and hence it is the maximum frequency of interest input in the *Frequency* step. Modal damping of 3% is input in the final step for all modes. Boundary conditions for the membranes are applied as explained for the uncoupled cases. ABAQUS requires, for accuracy, a finer mesh on the slave surface compared to the master surface ABAQUS. The average internodal distance given in the radial direction at the coupling interfaces of the cavity is 2 mm. For reducing computing expenses, this distance in the axial direction is increased to 20 mm. For the membranes too, the element edge length in the radial direction is increased to 2.5 mm. These increases are warranted as the analysis maximum frequency is much below compared to the uncoupled membrane analyses.

Figure 9.12 shows the surface pressure distribution at the karanai end of the cavity for two cases. The top row shows the distribution when karanai alone is coupled keeping thoppi end fixed. This analysis is done in ABAQUS after suppressing thoppi part and its coupling with the cavity. The bottom one shows the fully coupled case. The frequencies shown are close to the karanai fundamental (about 165 Hz). Observe that in the case where karanai alone is coupled, as frequency increases, contour with the highest pressure (red) moves from the point of application of load $(3a_1/4, \pi)$ to near the centre. There is a slight drift towards the opposite side of the diameter at 185 Hz, but not significant. While for the fully coupled case, this drift is substantial. The maximum pressure contour moves diametrically opposite to the point of application of load. This shift might be due to the influence of the coupled thoppi membrane mode near 190 Hz. Analytical cavity model, being one dimensional, is not being able to capture this phenomenon. It would explain the constant shift in quadrants of the arguments of the pressures and the vibrations in the fully coupled analytical model. The coupling is unstable, causing the magnitudes to build up, and thus solutions do not converge. Figure 9.13 gives another comparison where pressure magnitudes are plotted from the 3D FE and analytical models when karanai alone is coupled. The pressures are taken at the centres of karanai and thoppi end surfaces of the cavity. Near the first resonance, the analytical model predicts higher acoustic pressure at the karanai end while predicting lower at the thoppi end. The near field pressure distribution close to the vibrating membrane in the cavity, which a 1D analytical model cannot capture, might be the reason for this.

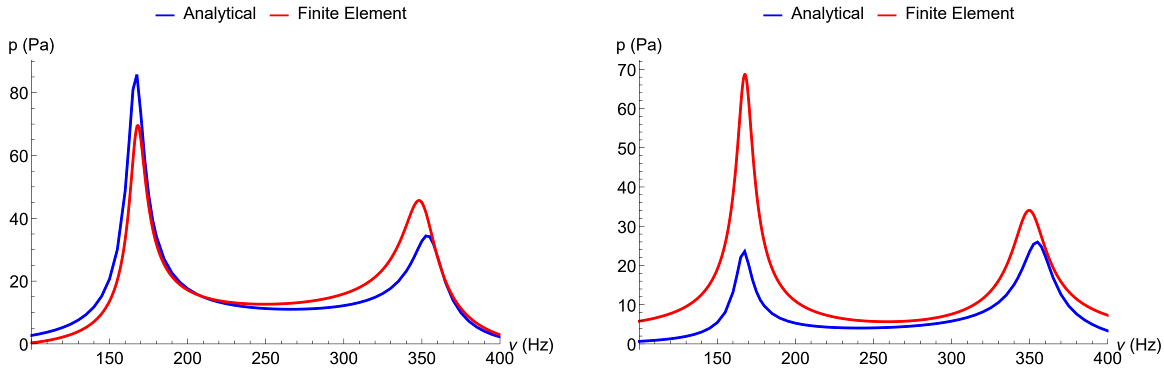


(a) Karanai coupled, 155 Hz (b) Karanai coupled, 170 Hz (c) Karanai coupled, 185 Hz

(d) Fully coupled, 155 Hz (e) Fully coupled, 170 Hz (f) Fully coupled, 185 Hz

Figure 9.12: Surface pressure contour plots (in Pa) at the karanai end of the cavity from FE models

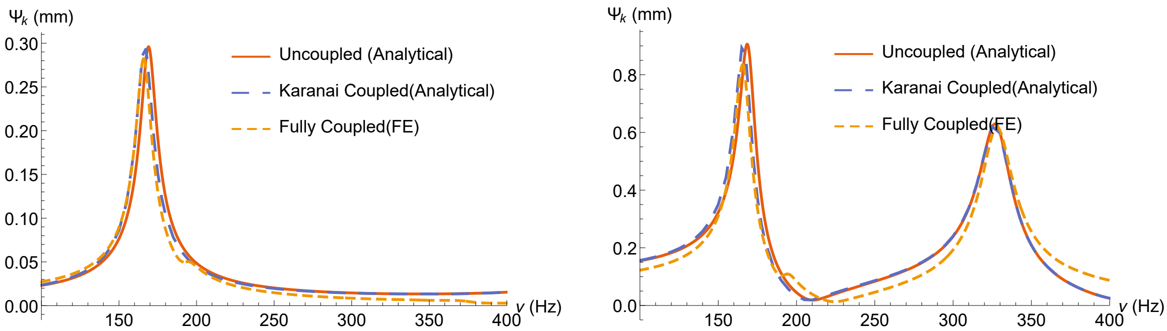
The above discussion suggests that the 1D cavity model has limitations when trying to simulate the fully coupled dynamics of Mridangam. However, the analytical model can give converged solutions in the frequency range from 100 to 400 Hz after removing the thoppi membrane and keeping the cavity end rigid. These results can be compared with the fully coupled FE results. Figure 9.14 shows this comparison, along



(a) Pressure at centre of Karanai surface (b) Pressure at centre of Thoppi surface

Figure 9.13: Comparison of pressure between analytical and FE models when karanai alone is coupled to cavity

with the uncoupled karanai vibrations. Observe that despite the absence of the thoppi membrane, the partially coupled analytical model is a close match to the fully coupled FE model. The first resonant peak is slightly higher in the analytical results, and there is a small spike near 190 Hz due to the presence of thoppi mode in the numerical results. Other than that, there is no considerable difference between the two of them. Both these curves are also very close to the uncoupled vibrations, except for a slight shift in fundamental frequency. These observations prove that the cavity acoustics or thoppi vibrations do not significantly influence the karanai vibrations. A shape sensitivity study using coupled models with alternate geometries for the cavity can further strengthen this statement.



(a) Karanai response at $(\frac{a_1+a}{2}, \pi/2)$ (b) Karanai response at $(3a_1/4, \pi)$

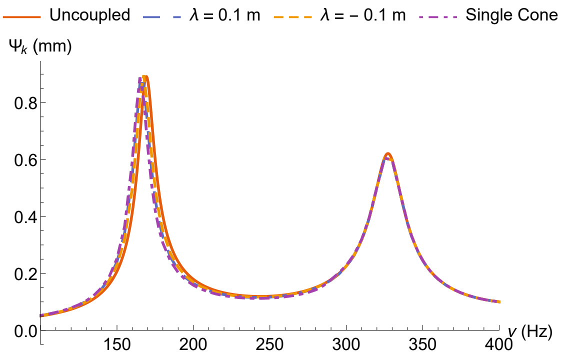
Figure 9.14: Comparison of harmonic response from karanai coupled analytical model and fully coupled FE model

9.4.3 Coupled Cavity-Shape Sensitivity Studies

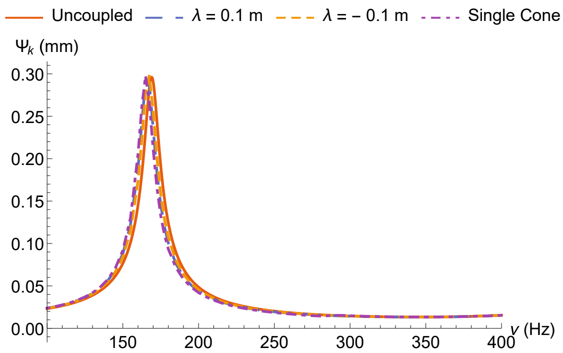
The set of alternate geometries for the Mridangam cavity was introduced in Section 9.2.4. Revisiting, geometry one (Figure 9.9a) had the largest diameter segment towards the thoppi side ($\lambda = 0.1m$) while geometry two (Figure 9.9b) had it towards the karanai side ($\lambda = -0.1m$). The third one was the single cone case (Figure 9.9c). The previous section revealed that the fully coupled analytical model has convergence issues. A partially coupled model after removing the thoppi did not predict the acoustic pressure well but seems to be a good one for predicting the vibration response of the karanai. Hence, design studies are possible using this model to understand the effect of cavity shape on karanai vibrations. However, fully coupled analyses are still necessary to back up these results. Due to the incapability of the analytical model to handle the full dynamics, this investigation is done using numerical methods. The coupled shape sensitivity study is therefore done in two ways: partially coupled analysis using analytical models and fully coupled analysis using FE models.

9.4.3.1 Partially Coupled Analytical Analysis

The previous section described the iterative process for analytically solving the coupling of membrane vibrations and the cavity acoustics. For conducting the shape sensitivity study, the original cavity model is replaced with alternate geometry models. Section 9.2.4 gives these acoustic models, built up using modal superposition. For convergence of solutions, the iterations do not include the thoppi vibrations, and the corresponding cavity end remains rigid. All the analyses are conducted for a harmonic point force at $(3a_1/4, \pi/2)$ on the karanai. Figure 9.15 shows the plots from these partially coupled analyses for the comparison. One can see that all geometries closely follow the uncoupled karanai vibration spectrum; the only noticeable difference being a slight lowering of the fundamental frequency. This shift was visible even with the actual cavity in Figure 9.14. The results confirm that the cavity shape does not alter the karanai vibrations remarkably. A final claim can be made with the fully coupled analyses using numerical models.



(a) Response at $(3a_1/4, 0)$



(b) Response at $(\frac{a_1+a}{2}, \pi/2)$

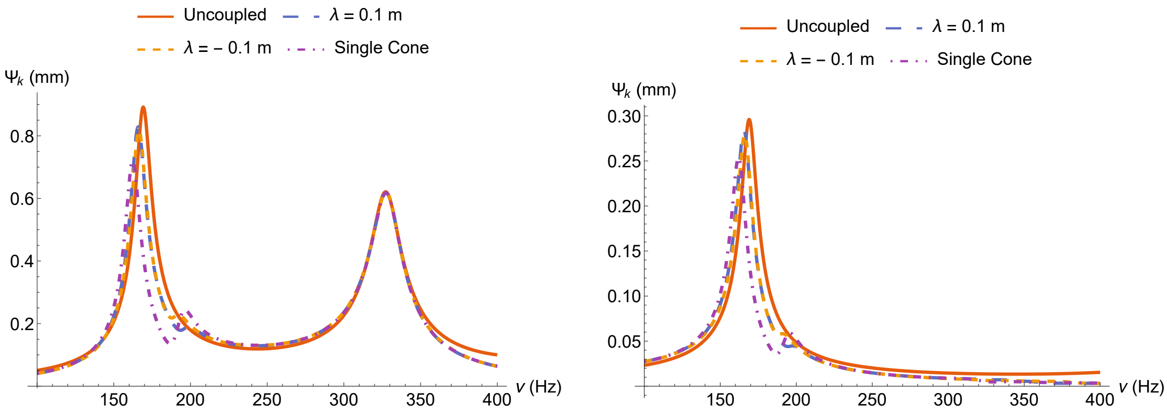
Figure 9.15: Partially coupled shape sensitivity analysis of karanai to a point load at $(3a_1/4, \pi)$

9.4.3.2 Fully Coupled Finite Element Analysis

The alternate geometries are numerically coupled with the membranes to do the shape sensitivity analysis. The FE models are similar in all respects to the fully coupled Mridangam model used in the previous section except for the dimensions of the cavity. They are sketched in ABAQUS to their lengths and meshed with elements similar in size as the original. Forcing is also consistent with previous analyses, given on karanai black patch region at point $(3a_1/4, \pi/2)$.

Figure 9.16 gives the karanai membrane responses captured from the geometries in comparison to the uncoupled analytical response. In alignment with the coupled analytical results, all the cavities show a slight shift in the fundamental frequency of the membrane. This shift is insignificant as it can be adjusted by changing the membrane tension, and players usually tune the instrument this way before recitals. There is also a small dip in the peak amplitude around the karanai first natural frequency, with the single cone cavity showing the most significant change. This reduction is intuitive since the cavity pressure is expected to oppose the external load on the membrane and dampen the vibration. When one looks at the second mode response, there is no change due to the cavity geometry, confirming that there is no coupling between the non-axisymmetric membrane modes and the cavity. Note also that the single cone case gives a more pronounced spike near the thoppi mode at about 190 Hz. Boundary excitation response from the single cone cavity (subsection 9.2.4.2) predicted this behaviour. Single cone gave higher responses at the opposite ends to the excitation (figure 9.10b and 9.10c). The peak displacement at the karanai first mode is not significantly altered by the presence of the cavity and is reasonably insensitive to its shape. Hence, it is not expected to bring notable changes to the vibration pattern of karanai. Experiments using a single cone shell-cavity, a possible extension to this work, can verify this for certainty.

All the coupled curves, both analytical and numerical, more or less follow the uncoupled vibration curve of the karanai disregarding small deviations. This finding is a clear indication that the shape of the Mridangam cavity does not modify the karanai vibrations significantly. The wooden shell can take almost any shape and



(a) Response at $(3a_1/4, 0)$

(b) Response at $(\frac{a_1+a}{2}, \pi/2)$

Figure 9.16: Fully coupled shape sensitivity analysis of karanai to a point load at $(3a_1/4, \pi)$

still do not alter the nature of the sound produced. Future research can verify these observations by doing experiments on Mridangam manufactured with differently shaped shell/cavity. Positive outcomes from these studies can significantly reduce the complexity of making and using the instrument.

9.5 CONCLUSIONS

The investigation started with questions about the importance of the shape of the Mridangam in the sound produced. Previous research had already suggested that the cavity enclosed plays little role in it. This paper, with a central focus on the vibro-acoustics of the instrument, further substantiates this inference. The coupled study showed that the air inside did not influence the karanai vibrations. This conclusion should naturally mean that any cavity shape could replace the original and the karanai should sound mostly the same. The coupled cavity-shape sensitivity studies confirmed that it is precisely the case. The three alternate geometries made only minimal differences to the uncoupled vibration spectrum of the karanai. Thus, one can infer that the Mridangam shell could be of any shape, and the karanai vibrations will remain unaffected.

However, not enough study was conducted on the thoppi membrane due to the difficulty in convergence when coupled. Further research has to be conducted to understand the effect of cavity coupling on thoppi. Experimental investigations using alternate geometries for the Mridangam shell will substantiate the findings here. Effects on transients due to different geometries is also a gap to be filled in the future.

Any scientific study on this subject is incomplete without contributions from the practising vidwans. Many of them are very particular about the tonal and transient characteristics of the instrument. Thus, practical investigations with artistic feedback on sound quality and ergonomics is a vital step forward towards developing alternately shaped Mridangams.

References

[DV63] Dowell, E. H. and Voss, H. M. "THE EFFECT OF A CAVITY ON PANEL VIBRATION". In: *AIAA Journal* vol. 1, no. 2 (1963), pp. 476–477. eprint: https://doi.org/10.2514/3.1568.

[Gho22] Ghosh, R. N. "Note on Musical Drums". In: *Phys. Rev.* vol. 20 (6 Dec. 1922), pp. 526–527.

[Gin10] Ginsberg, J. H. "On Dowell's simplification for acoustic cavity-structure interaction and consistent alternatives". In: *The Journal of the Acoustical Society of America* vol. 127, no. 1 (2010), pp. 22–32. eprint: https://doi.org/10.1121/1.3268606.

[Har20] Harikanth, M. “Analytial Models for Vibroa-acoustics of Mridangam”. MA thesis. Indian Institute of Technology Madras, Chennai, 2020.

[Lor16] Lord Rayleigh, O. F. “XII. On the propagation of sound in narrow tubes of variable section”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* vol. 31, no. 182 (1916), pp. 89–96. eprint: https://doi.org/10.1080/14786440208635477.

[Pil17] Pillai, M. A. “Analysis and design of acoustic energy harvesters using duct resonators and piezoelectric unimorphs”. PhD thesis. National Institute of Technology, Tiruchirappalli, 2017.

[Rag53] Raghavan, V. “Why is the Mridanga so-called?” In: *The Journal of the Music Academy* vol. 24 (1953), pp. 135–136.

[Ram22] Raman, C. V. “The acoustical knowledge of the ancient Hindus”. In: *Sir Asutosh Mookerjee Silver Jubilee Volumes*. Vol. 2. University Press, Calcutta, 1922, pp. 473–478.

[Ram34] Raman, C. V. “The Indian musical drums”. In: *Proceedings of the Indian Academy of Sciences-Section A*. Vol. 1. Springer. 1934, pp. 179–188.

[Rao38] Rao, K. N. “Theory of the Indian musical drums—Part I”. In: *Proceedings of the Indian Academy of Sciences - Section A* vol. 7, no. 2 (Feb. 1938), pp. 75–84.

[RS54] Ramakrishna, B. S. and Sondhi, M. M. “Vibrations of Indian Musical Drums Regarded as Composite Membranes”. In: *The Journal of the Acoustical Society of America* vol. 26, no. 4 (1954), pp. 523–529. eprint: https://doi.org/10.1121/1.1907369.

[Soo18] Sooraj, G. “Vibro-acoustics of a Mridangam”. MA thesis. Indian Institute of Technology Madras, Chennai, 2018.

[SRN10] Sivaraman, V. U., Ramasami, T., and Naresh, M. *Science for Musical Excellence (Sixth Raja Ramanna memorial Lecture)*. Lecture. Bangalore: NIAS, 2010.

[Vod62] Vodička, V. “Free Vibrations of a Composite Circular Membrane”. In: *Journal of the Physical Society of Japan* vol. 17, no. 4 (1962), pp. 698–702. eprint: https://doi.org/10.1143/JPSJ.17.698.

[Vod63] Vodička, V. “On forced vibrations of composite circular membranes”. In: *Czechoslovak Journal of Physics* vol. 13, no. 7 (July 1963), pp. 493–498.

[Web19] Webster, A. G. “Acoustical Impedance and the Theory of Horns and of the Phonograph”. In: *Proceedings of the National Academy of Sciences* vol. 5, no. 7 (1919), pp. 275–282. eprint: https://www.pnas.org/content/5/7/275.full.pdf.

Chapter 10

Design & Development of Electronic Indian Musical Instruments

G. Raj Narayan

Founder & MD, Radel Electronics Pvt. Ltd., Bengaluru

This article captures five decades of efforts of an engineer-musician hailing from a family of performing classical musicians to find innovative solutions to problems of musicians performing or learning Indian classical music.

10.1 Introduction

While electronics was introduced into music almost a century ago through various experiments, it was formally launched as a usable product in the form of the Moog Synthesizer in the 1960s when this author was in his teens. With a deep grounding in classical music as well as a degree in electrical engineering, this innovation captured his attention and, a few years later, drove him to dive deep into it for his own needs as a performing Indian classical musician.

The primary need of Indian classical musicians, especially instrument performers, is for Sruthi and Taala support during their practice sessions. This necessitated another individual playing the tambura and maintaining taala as support. In order to overcome this handicap, the feasibility of using an electronic circuit that could simulate these functions was explored. This led to the successful development of a reed-sound generator circuit in 1971, followed by a string-sound generator circuit in 1979. This served the needs of a constant Sruthi (Drone) accompaniment without the need for another person to perform that role. The next step was the invention of the 'Talometer', a digital taala machine that was developed in 1978. All these instruments were demonstrated at the Annual Music Conference of the Music Academy, Chennai, in December 1979.

The products were launched into formal commercial production in 1980 with periodic upgradation in their technology. Radel Electronics has been manufacturing a complete range of digital electronic sruthi-boxes, tamburas, veena, tabla, Talometer, Swarmandal, etc., over the last four decades.

10.2 Historical Background of Indian musical instruments

Indian musical instruments (excluding the Violin, Clarionet and Harmonium) include a large variety of string, wind and percussion types used both in Hindustani (North Indian) and Karnatic (South Indian) music. Many of these trace back to a few centuries when they used natural wood and bamboo, animal skin of various types, vegetable and fruit gourds as well as animal guts for strings. Apart from the change to metallic strings in a few of these instruments, very little else has undergone any change. Frets of a veena are still embedded into a bed of wax, Sitar and often Veena gourds too, are still made from dried vegetables, membranes of mridangam, tabla, khanjira, dholak, etc., are still made out of animal and reptile hides, bamboo flutes are still manufactured in a 'trial and error' process, etc. Very little scientific research, if any at all, is conducted into the design, operation, manufacture and testing of such traditional Indian instruments. Most of these instruments are manufactured using manual labour as single pieces or in small batches using hand tools instead of modern power-operated machines ('Numerous strings in my lute', The Hindu Magazine section, page 2, 29th Jan 2021). This is in sharp contrast to the western instruments that have had significant engineering inputs in their design and manufacturing processes in the last two centuries. There are many reasons for this state of affairs in this country. It is relevant to capture these for a complete appreciation of the constraints that both creative artists, skilled artisans and even qualified engineers have had to endure over many decades.

Firstly, the traditional Indian instruments are mythologically connected to Hindu deities and are considered to be the physical embodiment of these deities (Refer links provided at the end of this article). Hence, any drastic change in their appearance or construction is considered by many purists as bad as blasphemy.

Secondly, Indian musicians are far more conservative than western musicians when it comes to tradition, physical appearance, tonal qualities, etc. They do not seem to accept a scientific solution to either overcome inherent deficiencies of the instruments or an improvement in their features to permit the enhancement of their creative skills.

Thirdly, most of the Indian classical musicians are averse to the use of new technologies and scientific solutions that are perceived as complex and undesirable, if not altogether unwanted. Scientific explanations were looked at with suspicion and waved off with disdain.

Hence, it required a considerable amount of patience, perseverance and persistence to convince Indian classical musicians to appreciate the real benefits of the revolutionary inventions in the last five decades.

10.3 Needs of Indian Classical Musicians

Indian classical musicians need basic sruthi (Tonic drone) and Taala (pattern of rhythmic count) support both during their practice and concert performances. The long, unwieldy, fragile and delicate Tambura/ Tanpura used for sruthi accompaniment were difficult to carry especially for long distance travel. Further, for any artist performing on instruments, another person was required for playing the Tambura accompaniment. These issues therefore required a solution.

Taala accompaniment was specifically required for artists performing on instruments. This again required another person sitting with the artist and indicating the taala patterns on his/her hands in the case of Karnatic music or playing on the tabla in the case of Hindustani music. In the case of professional musicians who practice for 8 to 12 hours per day, it was almost impossible to find such support over such large periods of time.

The reed type Sruthi-box or Sur-peti, which was widely used for practice sessions, had only one or two sruthis (sur) built into it as per the requirements of the user. It was not tunable at all apart from selection of Madhyama or Panchama notes. While the stringed tambura was capable of being micro-tuned, there were limitations in tuning them over more than a note or two, beyond which the strings had to be changed to different gauges. This was laborious and inconvenient. The same is true with the traditional tabla used for Hindustani music. In addition to these limitations, the bellows of a reed sruthi-box would start leaking over time and the reeds would also corrode and deteriorate in their performance. The strings of a Tambura would rust or deteriorate due to metal fatigue and ultimately break. The skins of a tabla or similar instrument deteriorate rapidly with use and require replacement quite frequently for restoration of good resonance. All these constituted additional maintenance and usage issues for a musician.

10.4 Design & Development of electronic Indian instruments

Due to the increasing use of electronics in music, it was decided to find modern solutions for the provision of a steady sruthi drone support as well as taala indication without the need for another person. In other words, it had to be a gadget adjustable to any pitch, taala and tempo. The first attempt was made in 1971 to simulate the sound of a reed type sruthi-box. This was a battery operated compact (Approx. 200x125x125mm) unit that could generate three independently tunable notes corresponding to two shadjas (an octave part) and the intermediate panchama or madhyama as needed (Ref. Figure 10.1). The continuously adjustable tuning range of these notes was a little over one octave and hence could be used by anyone. The circuit was based on astable multivibrators implemented using discrete germanium transistors that were available in the commercial market. The square wave generated by the multivibrators were passed through differentiating filters to generate an approximate sawtooth waveform that is characteristic of a reed.

Figure 10.1: Picture of Dhruva Sruthi box

The critical issue here was to achieve a reasonably stable frequency to ensure that the musician did not need to retune it frequently at least over an hour or two. This issue was addressed by the use of a zener regulated power supply for the oscillators and use of low temperature coefficient metallised polyester capacitors in the timing circuits. In spite of a frequency stability close to 1% that was not perceptible by itself to the human ear, the slow beats resulting from the harmonic related frequencies of the other two notes could be heard. This therefore required some periodic re-tuning at intervals of about half an hour. However, this instrument was good enough for use during extended practice sessions and received acceptance from many musicians that ranged from students to professional concert performers, primarily due to the advantages it provided over the traditional reed instruments. Years later, in 1979, it was commercially launched and sold in the Indian market, with later improved models having ultra stability with the use of microcontrollers.

Having solved the need for an automated sruthi accompaniment, the focus then shifted to a Taala machine. Having been exposed to digital ICs and LEDs that had started becoming available in the commercial market around 1977, the first ever electronic machine for the audio-visual indication of Karnatic classical taalas was designed. Taalas are maintained or indicated in Karnatic music on the hands of a person. This is a complex, but standardised system that uses a beat of the palm (facing downward), wave of the palm (facing upwards) and count of fingers extending upto 9. Karnatic music primarily uses 7 Taalas with each of them comprising basic elements called 'Laghu', 'Dhrutham' and 'Anudhrutham'. A Laghu comprises a beat of the palm followed by count of fingers. The total count of a Laghu can be 3, 4, 5, 7 or 9 and these are referred to as 'Jaathi'. The 7 Taalas are made up of different combinations of these basic elements. The requirements of an automated electronic Taala instrument was therefore to reproduce these basic elements in a form that was as close to a visual representation on a human hand. In the absence of affordable technologies in those days for simulating pictorial hand movements, the best alternative was to use a sequence of LED lights in association with an audible sound for each of the counts of a rhythmic cycle. Figure 10.2 shows how this was implemented in the first prototype which was demonstrated at a few music conferences in Bengaluru and Chennai in 1978 and 1979. (Refer Figure 10.3). The instrument consisted of two rotary selector switches for selection of the Taala and Jathi, a tempo knob for adjustment of the speed of progression, a pitch knob for adjustment of the pitch of the sound generated with each count. The instrument could be set to all 35

Figure 10.2: Picture of first 'Talometer'

combinations of Taala and Jaathi and any desired tempo to enable a musician to independently practice. The instrument used standard 74 series TTL logic ICs for the implementation of programmable counters and LED decoder drivers. This instrument proved very successful and useful for this author and his wife Radhika, both being proficient in performing Karnatic music on the flute and veena respectively. This instrument was also commercially launched in 1980 with later versions using a microprocessor (1988) and a microcontroller (1999).

Electronic Tambura 'Saarang' - A repetitive desire expressed by many musicians who witnessed or used the electronic sruthi-box with the reed sound described above, was for an instrument that simulated the sound of a traditional tambura. It was easy to appreciate that the sound of a plucked string possesses a decaying amplitude asymptotically reducing to zero over time. However, what was really challenging was the complexity of the vibrational characteristics of the tambura string. In the absence of a storage oscilloscope that was unaffordable as well as inaccessible to a hobby electronics enthusiast in 1979, the only option was to explore experimentally. Once again, discrete transistor based multivibrators were used to generate square waves of the desired frequency. This was amplitude-modulated to create the required decaying envelope of a string by the use of a capacitor charge/discharge sequence serving as the collector supply of a transistor. The resultant signal was then differentiated and filtered to obtain a harmonically rich tone. While this was a crude method by current standards, it produced a reasonably acceptable tone to enable professional musicians to use it even for their concerts. The first ever such electronic tambura, named 'Saarang', was demonstrated along with the Sruthi-box and the 'Talometer' during the lecture demonstration sessions at the annual music conference of the Music Academy, Chennai in Dec. 1979. (Figure 10.4)

Vidwan Dr. M Balamuralikrishna, who was present as one of experts in the audience, evinced keen interest in using the product and studied the instrument in detail. All credit for the successful acceptance of this invention by the music fraternity should go to him since he started using it in all his concerts without even the traditional tambura by his side. Many other senior musicians followed suit due to the convenience of portability, volume adjustments and sruthi stability.

Launch of the electronic tambura brought forth considerable interest from not only musicians but also researchers and academicians. Some papers published by researchers at the ITC Sangeeth Research Academy,

3-12-78
Times of India Sunday Magazine

Electronic Aids For Classical Music

IS there a place for twentieth century electronics in tradition-bound Indian classical music?

Purists would normally say no, but two novel instruments — the talometer and the electronic shruthi box — invented by a young engineer-flutist from Bangalore have been hailed without reservations by traditionalists and avant-gardists alike.

The talometer is a battery operated, audio-visual aid that marks time and keeps track of tala cycles. It is designed for use in South Indian music and can be set to perform any of the 35 major talas of Carnatic music. It can also be adapted to reckon North Indian tala patterns.

In Indian music, a performer has not only to maintain a steady speed, he has to keep track of the tala cycles, too. This is done with the hand if he is a vocalist. An instrumentalist is handicapped to the extent that his hands are already occupied in playing his instrument. He therefore has to use the services of another musician to mark tala. An out-station performer has to choose between paying for an additional member in his party and relying on some local. unfamiliar individual — with results that are not always happy. The talometer solves this problem by virtually adding an extra pair of hands to an instrumentalist and taking on the work of the 'time keeper.'

This it does with the help of a row of miniature bulbs that light up sequentially, one at a time, to indicate the particular stage of the tala cycle at that instant. Each bulb represents a finger count or beat. A bell is also provided to sound out the divisions of each cycle. By manipulating the switches for tala, jati and speed as required, a musician can set the talometer on a precise, unvarying and continuous course of cycles in any of the 35 talas. An expert musician can even adjust it to an additional 112 talas.

If this devise is useful to violinists and veena and flute players, it is indispensable for percussionists, like mridangam players, who cannot practise or perform without visual tala and normally require someone to mark the cycles with the hands.

When the talometer was demonstrated in Bangalore recently, it received enthusiastic notices from musicians as well as critics and scholars. The noted violinist, Professor T. N. Krishnan, who is now the principal of the College of Carnatic Music, Madras, was among those present at the demonstration; he pronounced it to be an invaluable accessory for musicians and congratulated the inventor.

There is nothing in the talometer to offend purists, nothing gimmicky to detract from or 'dilute' the classical merits of the music itself. On the other hand, a musician can improve his sense of rhythm by pitting himself against the machine and checking if the cycles coincide.

The inventor of the talometer, 29-year-old G. Raj Narayan, is an engineer working with Hindustan Aeronautics, in Bangalore. He comes from a family of musicians and is himself a promising flutist broadcasting regularly over AIR, Bangalore. He also plays the veena, the guitar and the tabla (which he learnt as part of his training in Hindustani music). Narayan's wife Radhika is a veena player in her own right and the couple often give concerts in duet.

Besides the talometer, Narayan has also designed an electronic shruthi box which was demonstrated at the annual conference of the Bangalore Gana Kala Parishad, in 1971.

A shruthi box is a drone instrument often used along with, or as a substitute for, the tanpura. Because of the arrangement of bellows on which it works, the sound produced is sometimes jarring and discontinuous. Narayan's electronic shruthi box gives not only a continuous and uniform drone, it can also be adjusted for any pitch required (the conventional shruthi box cannot be tuned, except by unscrewing and replacing the reeds; in the case of the tanpura, the strings will have to be changed for large variations in pitch). Since there are no moving parts, it can hold its note stable for any length of time.

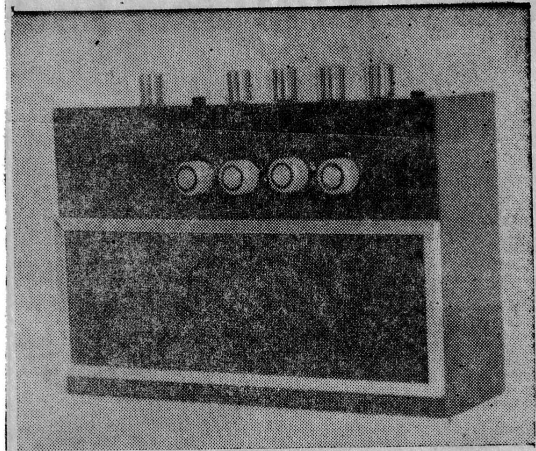
"One has to strike a compromise between stability and cost," Narayan explained. "The more sophisticated I make it, the better its notes — but the higher the cost too."

A volume control knob is also provided in this gadget to adjust loudness to cater to different needs. It is a compact box whereas the tanpura is both unwieldy and fragile. The greatest attraction of the electronic version, particularly to instrumentalists, is that, like the talometer, it does not require the help of another person for providing the shruthi.

Sakuntala Narasimhan

Figure 10.3: Press report on Talometer

DECCAN HERALD
JAN 7 1980



The electronic thambura design ed by G. Raj Narayan, a City engineer and flutist.

City engineer makes electronic Thamboora

By S. N. CHANDRASEKHAR

Thamboora, the four-stringed, fretless instrument is an integral part of a performing musician, vocalist and instrumentalist alike. Producing as it does the all important 'sruti'. it is really difficult to imagine a concert without the key-filled top of it jutting out over the heads of the performers.

But electronics are out to bring about a change in this familiar setting. G. Raj Narayan, a City engineer-cum-amateur flutist, has designed a device which may soon prove a boon to the artiste, both because of its handy size and manoeuvrability. The device produces the plucked sound of the Thamboora with the same intensity, precision and accuracy without the risk of having to trim the tuning every now and again, if the artiste is sensitive to the 'sruti'.

This musical sound which is so essential for accurate tonality is generated electronically, neither the strings nor any other mechanical parts accounting for it. Nor has one to strain his fingers by constantly plucking the strings to produce the sound. Much less is the need to replace the worn-out strings every now and then.

Yet, the device has the facility to fix the "Sruti" of more than one octave range, which means it can be interchanged from the male or female range with the least confusion. It must be noted that the Thamboora cannot be adjusted for any variation beyond a quarter octave.

For Raj Narayan, this new device is not the first achievement. His 'sruti-box' and 'talameter' are already popular with musicians all over the country. Like those inventions which were introduced less than three seasons ago, the electronic Thamboora is also operated on battery, though it could be as easily converted to mains. More importantly, the device is rugged in its make-up and does not call for delicate handling.

The box-like device is compact and measures 250 x 175 x 80 m.m. The adjustments depend on four major knobs. the other four knobs facilitating finer tuning. The volume of the 'sruti' can also be much louder than the conventional Thamboora. That it is automatic also eliminates the obiquitous Thamboora-player.

As Raj Narayan explained, the device was demonstrated before the Experts Committee of the Madras Music Academy Conference recently.

It is not surprising that it was received with great enthusiasm and appreciation, as was the case with the other two instruments by all the leading artistes.

With all this, the cost of the Electronic Thamboora does not exceed Rs. 850-00, a price which is almost the same as for the conventional instrument.

Figure 10.4: Press reports about Tambura

Calcutta (Kolkata) in the 1980's had reported their findings of the rich harmonic content. In experiments conducted by this author in 1995, the following interesting observations were made.

1. The spectral content of a vibrating tambura string varied continuously during its ADSR (Attack Decay Sustain Release) envelope.
2. The maximum harmonic content, going upto the 19th harmonic, was present during the peak resonance phase.
3. The attack and decay phases had significantly lower content of the rich harmonics.
4. Certain phases of the vibration possessed a periodic amplitude modulated waveform. with the higher harmonics modulated by the fundamental and lower harmonics.

A very important point to note in the study of tambura string vibrations is the fact that a characteristic 'soft suspension' is used on the brass or ivory bridge of a Tambura by the placement of a cotton or wool string. This is placed such that a point along the string at a few millimeters from the point of suspension strikes the bridge repeatedly as it vibrates. This in association with the soft suspension provides a positive stimulus to increase the sustenance as well as produce the characteristically rich harmonics. The placement and thickness of the cotton string is dependent on the gauge and tension of the tambura string and hence its position would need to be altered if the string is tuned to a different pitch. There is an optimal placement of this soft suspension to produce the best combination of harmonics as well as sustenance. The desirable harmonic content of the string is a highly subjective matter among musicians. While some musicians like a very rich tone of the Tambura, others prefer a tone with reduced or minimal harmonics.

Provision of all these features of a traditional tambura in a digital electronic tambura poses quite a few complexities and hence affects its cost and feasibility. The feasibility of modeling this into an electronic tambura was explored with the use of a Texas Instruments TMS320F241 DSP chip in the year 1998. However, this was not pursued further since PCM based solutions were found to be simpler and more cost effective due to drop in memory prices. This boiled down to provision of a stored digital sample of a string within the instrument and replaying it at different sampling frequencies to generate the required notes. Hence the quality of the reproduced sound would basically depend on the quality and type of string sample stored in memory. With the processing power available in today's smartphones and other personal devices, it is now feasible to provide multiple samples of strings to enable a mix and match facility to realise a variety of tambura tones as desired by our musicians.

Electronic Tabla 'Taalmala' - With the basic needs of a Karnatic musician having been met with the electronic instruments described above, the focus shifted to the needs of the Hindustani musicians. Hindustani musicians, such as singers and instrument performers, need tabla accompaniment even during their practice. The characteristic 'theka' produced on the tabla for each taal, provides a reference to the rhythmic count (matras). This is therefore an auditory reference to the main performer who does not mark taals on his/her hands as done in Karnatic music. This reliance on the tabla constituted a severe handicap for Hindustani musicians for their practice sessions that would often commence at 9 in the night and extend well past midnight. This was therefore seen as an opportunity for innovation and the handicap was overcome through the invention of an electronic tabla in 1987. This was a far more challenging task than the tambura and 'Talometer' since it involved synthesis of different types of strokes on a pair of tablas (Dagga and Tabla, played on the left and right hands) and these strokes had to be generated in different sequences corresponding to various taals at different tempos and pitch as needed by the musician. Use of pre-recorded tapes (in 1980's) was out of question since tape speeds could not be easily altered. Even if this was made possible, both tempo and pitch would change. Hence, the only solution was to synthesize it such that tempo and pitch could be independently controlled. Further, the patterns and sequences of strokes had to be stored in memory in some codified form so as to be capable of being reproduced at the desired pitch and time intervals depending on the tempo selected by the musician.

Microprocessors had just begun to be commercially available in the Indian market and it was the opportune time to use it in the design of the electronic Tabla. Z80, an 8 bit microprocessor, along with its family of peripheral chips was chosen for the application. A basic development kit called 'Micro-Professor' was used. This enabled keying in programs in hexadecimal machine code into its RAM and then executing it. A Cassette tape interface on the kit enabled the loaded programs to be saved into an ordinary cassette tape-recorder as well as reloaded into RAM as required. A very basic multi-digit 7-segment LED alphanumeric display along with a keypad constituted the user interface for the development kit. Programs had to be written down on paper in Assembly language and then manually converted into machine codes and then keyed into the kit. Though this was laborious, there was no other alternative in the absence of the ubiquitous PCs that became available 10 years later.

The strategy adopted in the design of the electronic Tabla was to use the microprocessor as only a logical control and timing device with its own memory for storing coded patterns of taal thekas and to trigger and control the separate sound generation circuits. Sound generation circuits employed both digital and analog techniques to produce the decaying amplitude of percussion sounds of the pair of tablas.

Players of Indian percussion instruments such as the Tabla and Mridangam, produce a very large variety of strokes with infinitely variable intensities and tones, depending on the skill level of the performer. This is how unique styles associated with each maestro are created.

Another characteristic of the Tabla is that both the instruments (Tabla and Dagga) use loaded membranes. The membranes are supported around the circular periphery with another narrow strip of leather all round that adds to the tonal quality of the vibration. Apart from its own vibration characteristics, the membranes of the ‘Daayan’ are struck with fingers at different points along the radial and each of these strokes produces a distinctive tone with a distinct spectral content. On the other hand, the Dagga is struck predominantly at its center, but is also modulated with pressure applied by the heel of the palm. This results in a frequency modulated tone. A skilled artist can create almost infinite varieties of frequency modulated patterns.

Different parts of the vibrating membrane produce distinctly different sounds depending on its position on the membrane and the manner of striking it with the hands or fingers. Studies into the various modes of vibration of the loaded membranes of a Tabla or mridangam have been carried out by many researchers including Sir CV Raman and these have been published in various papers. These studies have highlighted the multi-modal nature of vibrations that produce multi-timbral tones. The ‘Dagga’ (normally played with the left hand and hence also called as the ‘Baayaan’) has a lower resonant frequency in the range of 40 to 100 Hz whereas the Tabla (normally played with the right hand and hence also called as the ‘Daayaan’) has a much higher resonant frequency in the region of 200 to 400 Hz. It needs to be tuned to align with the sruthi of the main performer. Since the membrane of the Tabla cannot be tuned beyond a range of two semitones, different tablas need to be used for accompanying say, gents or lady vocalists, or other instruments. This itself constitutes another handicap for tabla artists since they need to possess a number of tablas tuned to various sruthis of an octave.

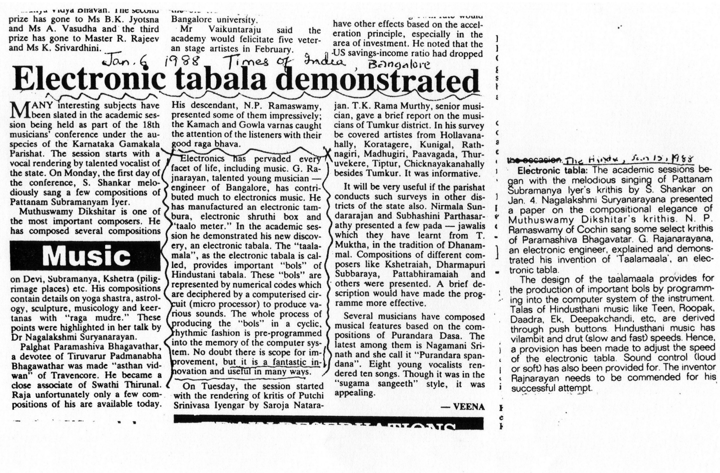
The foregoing explanations describe the challenges that were faced in simulating and synthesizing the complete taal thekas for Hindustani classical music. In order to arrive at a basic instrument suitable just to enable musicians to practice independently, the various tabla strokes forming part of the taal thekas were analysed to identify a minimal set of acceptable strokes that could uniquely distinguish the required theka patterns. This resulted in the following basic ‘bol-s’ (strokes on each of the instrument) required to be generated.

- Tabla - ‘Na’, ‘Tha’, ‘Thun’, ‘Thi’, ‘Ru’, ‘Ta’
- Dagga - ‘Ki’, ‘Ge’ frequency modulated in 7 different but fixed patterns.

It was quite interesting to find that the ‘Na’ played at the edge of the ‘Dayan’ comprises an emphasised 3rd harmonic (Panchamam) in addition to the second harmonic. The fundamental is quite subdued almost to the point of being absent.

The ‘Tha’ played in the area between the edge and the loaded disc, is rich in the second and 4th harmonics. ‘Thi’ ‘Ru’ and ‘Ta’ are muted sounds produced on the ‘Daayaan’. So is the ‘Ki’ produced on the ‘Baayan’.

‘Ge’ is generated on the Dagga by gently resting the heel of the palm in the area beside the loading disc and striking with the tips of fingers either on the loaded disc or a little beyond it. Frequency modulation of the resultant tone is performed by either pressing with the heel or moving it towards the centre with pressure. This frequency modulation was also analysed and reduced to 7 basic types of patterns as per common practice of playing the theka-s.



prize has gone to Ms B.K. Jyotsna and Ms A. Vasudha and the third prize has gone to Master R. Rajeev and Ms K. Srivardhini.

Bangalore university.

Mr Vaikuntaraju said the academy would felicitate five veteran stage artistes in February.

have other effects based on the acceleration principle, especially in the area of investment. He noted that the US savings-income ratio had dropped

Jan. 6 1988 Times of India, Bangalore

Electronic tabala demonstrated

MANY interesting subjects have been slated in the academic session being held as part of the 18th musicians' conference under the auspecies of the Karnataka Gamakala Parishat. The session starts with a vocal rendering by talented vocalist of the state. On Monday, the first day of the conference, S. Shankar melodiously sang a few compositions of Pattanam Subramanyam Iyer.

Muthuswamy Dikshitar is one of the most important composers. He has composed several compositions

Music

on Devi, Subramanya, Kshetra (piligrimage places) etc. His compositions contain details on yoga shastra, astrology, sculpture, musicology and keertanas with "raga mudre." These points were highlighted in her talk by Dr Nagalakshmi Suryanarayan.

Palghat Paramashiva Bhagavathar, a devotee of Tiruvarur Padmanabha Bhagawathar was made "asthan vidwan" of Travencore. He became a close associate of Swathi Thirunal. Raja unfortunately only a few compositions of his are available today.

His descendant, N.P. Ramaswamy, presented some of them impressively; the Kamach and Gowla varnas caught the attention of the listeners with their good raga bhava.

Electronics has pervaded every facet of life, including music. G. Rajnarayan, talented young musician — engineer of Bangalore, has contributed much to electronics music. He has manufactured an electronic tambura, electronic shruthi box and "taalo meter." In the academic session he demonstrated his new discovery, an electronic tabala. The "taalamala", as the electronic tabala is called, provides important "bols" of Hindustani tabala. These "bols" are represented by numerical codes which are deciphered by a computerised circuit (micro processor) to produce various sounds. The whole process of producing the "bols" in a cyclic, rhythmic fashion is pre-programmed into the memory of the computer system. No doubt there is scope for improvement, but it is a fantastic innovation and useful in many ways.

On Tuesday, the session started with the rendering of kritis of Putchi Srinivasa Iyengar by Saroja Natarajan. T.K. Rama Murthy, senior musician, gave a brief report on the musicians of Tumkur district. In his survey be covered artistes from Hollavanahally, Koratagere, Kunigal, Rathnagiri, Madhugiri, Paavagada, Thuruvekere, Tiptur, Chicknayakanahally besides Tumkur. It was informative.

It will be very useful if the parishat conducts such surveys in other districts of the state also. Nirmala Sundararajan and Subhashini Parthasarathy presented a few pada — jawalis which they have learnt from T. Muktha, in the tradition of Dhanammal. Compositions of different composers like Kshetraiah, Dharmapuri Subbaraya, Pattabhiramaiah and others were presented. A brief description would have made the programme more effective.

Several musicians have composed musical features based on the compositions of Purandara Dasa. The latest among them is Nagamani Srinath and she call it "Purandara spandana". Eight young vocalists rendered ten songs. Though it was in the "sugama sangeeth" style, it was appealing.

— VEENA

~~the occasion.~~ The Hindu, Jan 13, 1988

Electronic tabla: The academic sessions began with the melodious singing of Pattanam Subramanya Iyer's krithis by S. Shankar on Jan. 4. Nagalakshmi Suryanarayana presented a paper on the compositional elegance of Muthuswamy Dikshitar's krithis. N. P. Ramaswamy of Cochin sang some select krithis of Paramashiva Bhagavatar. G. Rajanarayana, an electronic engineer, explained and demonstrated his invention of 'Taalamaala', an electronic tabla.

The design of the taalamaala provides for the production of important bols by programming into the computer system of the instrument. Talas of Hindusthani music like Teen, Roopak, Daadra, Ek, Deepakchandi, etc, are derived through push buttons. Hindusthani music has vilambit and drut (slow and fast) speeds. Hence, a provision has been made to adjust the speed of the electronic tabla. Sound control (loud or soft) has also been provided for. The inventor Rajnarayan needs to be commended for his successful attempt.

Taalmala: a landmark in electronics rhythm

By Our Music Critic

FIRST, it was the "surpeti", in 1970. Then it was the "talometer", in 1978. And now, it is the "Taalmala". These three electronic gadgets add up to the achievements of G. Raj Narayan (39), of Bangalore, who recently won the IMM-Binatone electronics engineers' award in the "entertainment electronics" category, recently in New Delhi.

The music department of Bombay University has sponsored an hour-long demonstration of Mr Narayan's latest innovation in the presence of a large audience on Thursday evening. Simply put, the taalmala is an electronics tabla, designed to produce the actual sounds of the age-old percussion instrument, as played by an artiste in a live performance. All one has to do is simply to switch the instrument on, select the *taal* and *laya* desired, and press the 'start' button. Hey presto! you have a maestro at your command!

The instrument can produce, with uncanny fidelity, the *theka* (rhythmic sounds of the tabla) for several commonly used Hindustani *taals*, such as *ektaal, teentaal, jhaptaal, roopak, dadra, deepchandi, keharwa, ada-chautaal, jhumra* and the like.

Compact and portable, the box-type taalmala has controls, which are easy to use, to facilitate right adjustment of volume (loudness), pitch (basic *shadja*) of the tabla speed (tempo) and balance control for relative volumes of the left hand ("bayan") and the right hand ("dayan") sounds.

The selection of various *taals* is secured through a calculator-type keyboard, which is also equipped with the "start" and "stop" keys. Various ranges of *laya*, such as *drut vilambit* and *ati-vilambit* can also be set to obtain the desired speed-range. There is a knob designed to adjust the speed continuously within any of these ranges.

The instrument automatically inserts additional *bols* (sounds) in the *theka* for *vilambit* and *ati-vilambit* modes to simulate a true-to-life rhythmic accompaniment.

Mr Narayan is a perceptive flutist as well. He punctuated his enlightening narration with brief renditions of *Jhinjboti* and *Bageshree*, both set to *teentaal*, while explaining the basic idea or the object of his instrument. And these he played to the accompaniment of the simulated *taal* that emanated from the taalmala.

Mr Narayan demonstrated how the basic elements of the tabla *bols* were simulated. He also showed how the frequency modulations of the "bayan" could be secalled after putting them into the memory of the machine.

The taalmala

"The Times of India" - Bombay Wednesday March 2nd 1988

Figure 10.5: Press report about Taalmala

The above requirements were then encoded into data bytes and a table of such data bytes constituted a complete theka starting with the first matra and ending with the last, after which the whole sequence was repeated endlessly.

One more issue faced during interactions with stalwarts of Hindustani music was the lack of standardisation among musicians from different regions. This is characterised by different 'gharana-s' (schools). It was found that differences existed in the naming as well as the theka patterns of the taal-s. While a musician from one gharana may want the bol 'Na' for a specific matra of a taal, another from another gharana would want a 'Dha' or 'Tha' for the same matra. This meant that the theka created for one gharana may not be acceptable to another. This therefore created the need for making the taal thekas either editable by the user or alternate versions to be provided. This was indeed done in later versions of the electronic tabla.

The embedded software of the instrument decoded the data patterns in real-time and sent out triggers to the sound generation circuits at different points of time as defined by the tempo selected as well as defined by the data patterns themselves. Timing was provided by a Counter-Timer peripheral of the Z80 family and was connected to the crystal clock used for the microprocessor circuit.The first ever electronic tabla, named 'Taalmala', was demonstrated at a lecture demonstration held at Bengaluru in Dec. 1987 and subsequently at the Bombay University in Feb. 1988. (Ref. Figure 10.5)

While the sound simulation of the first model was quite crude by current standards, the instrument was received well by many performing musicians of repute primarily due to the fact that it eliminated the need for a tabla artist for their practice sessions. This was a huge benefit since it allowed them to practice everyday at any time of the day or night and for as long as they wanted. Starting with 12 basic thekas of most commonly used taals, another model was provided with 24 thekas. As part of the continuous development, the technology was subsequently upgraded to incorporate PCM encoded sampled tones in 1998. The present generation of digital tablas provide 108 thekas besides allowing a user to create his/her own theka, provide 'Fill-in' and

'Random play' features to create a live accompaniment feel for the user. 'Fill-in' is a feature to introduce a catchy phrase of rhythm to synchronously end at the first count of the next cycle. This phrase would be picked out in realtime from a table of pre-defined phrases so as to match the total number of counts remaining in the present cycle of the theka. 'Short Fill-in' and 'Long Fill-in' options have also been provided as a choice for the user. User programmability, where the user can create his/her own theka by keying in a sequence of bol-s and store it in memory, was also incorporated. All these efforts were aimed at making the instrument as close to live accompaniment as possible. A MIDI interface was also incorporated to enable the instrument to be capable of communicating with other digital musical instruments or with a PC. Such modern features have enabled these instruments to be used in concerts across the world including fusion music with western artists.

'Sunadamala' Lehera instrument - Once the needs of a lead singer or instrumental musician was met, it was the turn of the accompanists themselves to benefit from the introduction of electronics. Just as the lead musician requires a tabla accompaniment even during practice sessions, the tabla artist also requires a repetitive tune to be played as a reference for his taal reference. This repetitive line of music is referred to as a 'Lehera' in Hindustani music. This requirement used to be met by another musician playing the lehera on either a harmonium or a Saarangi. This therefore led to the development of the electronic lehera machine named 'Sunadamala'. This was again a microprocessor based instrument that provided 200 pre-programmed tunes set to different taals of Hindustani music in 20 groups. Each group had 10 tunes in different ragas since Hindustani music utilises different ragas for different times of the day or night. The lehera tunes, simulating the sound of a harmonium, could be set to any sruthi or tempo as desired. The first such instrument was launched in the presence of the renowned sitar maestro, Pt. Ravi Shankar, in an event organised in Bengaluru in 1993, in the presence of an august gathering of many musicians. (Ref. Figure 10.6). Pt. Ravi Shankar spoke briefly appreciating the efforts of the inventor in his quest for modern solutions to problems of Indian musicians.

Electronic & Digital Veena - The veena is a beautiful Indian stringed instrument that is mythologically associated with Goddess Saraswati. It has a sweet twangy sound and it was used as an accompaniment for Karnatic musicians many decades ago. This author has a very intimate association with the instrument since his mother, Smt. Sugandha Raman performed both Karnatic and Hindustani music on it over the Indian broadcast system, All India Radio (later renamed Akashvani).

The Veena has many inherent deficiencies in its original acoustic form. The main deficiency is its low volume of sound consequently leading to the perception of poor sound sustenance. The sound is so weak that very often the playing artist is unable to hear its sound in the presence of much louder accompanying percussion instruments. The second deficiency lies in the fact that its frets are mounted in a bed of specially formulated compound of wax, which is very susceptible to distortions caused by variations of temperature and humidity, besides being very delicate. This requires periodic re-setting by a trained artisan. The third drawback is the fact that it is a bulky and fragile instrument, prone to breakage during transportation. Two of these three primary deficiencies were addressed first in the development of an electric veena named 'Sunadavinodini', in 1971 (Refer Figure 10.7). This was an adaptation of the electric guitar by the incorporation of a magnetic pickup directly under the strings and elimination of the bulky acoustic resonator (Kudam). This innovation enabled the support gourds to be capable of being dismantled easily so as to make the instrument very portable. Brass frets were fitted permanently on the wooden fretboard to eliminate the use of wax. However this was not found practical since the position of the frets also depended on the gauge of the strings and their tension corresponding to their pitch settings. Hence, the frets needed to be adjustable to provide for all environmental variations as well as string properties. However, the above mentioned experimental veena demonstrated the improvements in loudness as well as portability. This instrument was also used by Smt. Radhika Rajnarayan in a live concert at the Sur Singar Samsad, Bombay in 1986 (Figure 10.8). The next improvement was made in 1998 by the incorporation of an electronic tambura as well as an ampli-speaker into the gourds of the veena. This resulted in a self-contained instrument that was far more portable than the traditional instrument and loud enough for the performer even in the midst of a loud percussion accompaniment. A fully engineered version was launched for the commercial market in the year 2001 with adjustable frets mounted on two rails using clamps with screw and nuts. This seemed to solve almost all the deficiencies of the traditional instrument.

The only deficiency still remaining was the need for periodic tuning of the strings as one performed on it, especially since the Karnatic style involved frequent transverse pull on the strings to produce 'gamaka-s or

MUSIC

Sound waves

Once he designed a cockpit simulator. Now he invents hi-tech musical instruments

THE mechanical engineer wanted to play his harmonium with all his limbs. He designed a stand to manipulate the bellows with his knees. His son, an electronics engineer who found it difficult to operate, came up with a better one: a *shruti* box that would play on its own. That was just the first of G. Raj Narayan's electronic musical aids, which are blending the hi-tech with Indian classical music. His latest is called Sunadamala, an electronic *lehera* instrument.

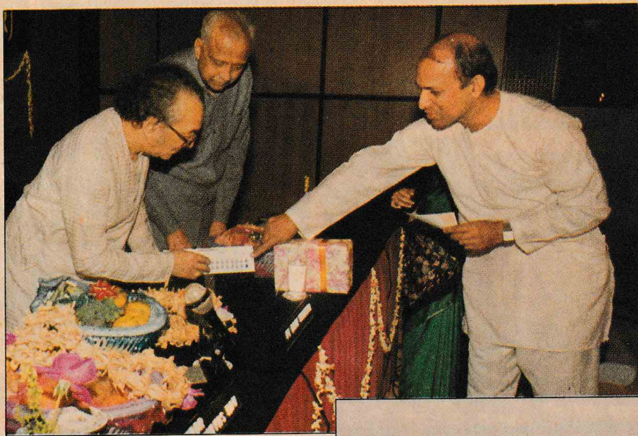
"I am worried," sitar virtuoso Ravi Shankar said mock-seriously while releasing the instrument recently in Bangalore. "My worry is that classical artistes like us will not even be needed in the concerts of the future. Perhaps there will just be photographs of us on stage with electronic instruments playing. I do hope that electronic music does not progress to that level."

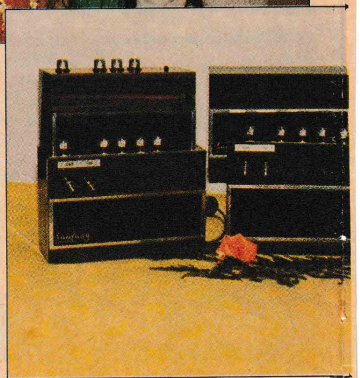
Synthesisers were becoming popular in the west when Narayan, a flautist like his father, designed the *shruti* box in 1971. "Electronics had just come into music in the west," he recalls. "I just wanted to reproduce the sound of the *shruti* electronically, so that I would not be dependent on an accompanist. I had no intention of commercialising it."

He didn't, for the next decade. He was then working as a systems design engineer at the Hindustan Aeronautics Ltd (HAL), where he designed a cockpit simulator for the air force. Simulation has been his forte in music, too.

Over the next few years, he developed a simple audio-visual aid which he calls Talometer. This small, box-like gadget can be set to keep track of tala (rhythm) cycles and to display them to the musician with tiny colour lights. It spares the Carnatic music instrumentalist of the mental gymnastics he normally does since his hands are engaged. The Talometer displays 28 tala cycles at different tempos.

"If you keep it in front of you on stage, you do not have to ask a friend to sit in the front row of the audience, or beside you on stage, and mark the cycle for you, so that you do not slip up," says P. Ramamurthy, a student instrumentalist.





Touch of class. *Pandit Ravi Shankar operates the Sunadamala helped by Narayan (above); the five instruments Narayan has designed (right); a demonstration of the Sunadamala and the Taalamala (far right)*

But many musicians were not entirely satisfied with the *shruti* box, which they found less resonant and melodious than a tanpura. So Narayan decided to simulate electronically the sequential plucked sound of this stringed accompaniment instrument. The outcome of three months of dedicated effort was another little box (24x17x8 cm) that he named the Saarang.

The conventional tanpura is a long and delicate wooden instrument, with either a gourd (as in the north) or a wooden bowl as resonator. It is unwieldy and highly influenced by temperature and humidity, unlike Narayan's instrument which can be set to a steady pitch.

Appreciation of the Saarang came from maestros. Famous violinist brothers L. Sankar and L. Subramaniam and eminent vocalist M. Balamuralikrishna, all of whom are known for their love for innovation, asked him to make them electronic tanpuras.

But everything was not sweet music in his life and he felt forced to leave HAL. "I did not resign just because I wanted to make these instruments," he says. "I quit simply because a govern-

52 THE WEEK ■ JAN.31, 1993

Figure 10.6: Press report about Sunadamala

ನೂತನ ವಾದ್ಯ-ಸುನಾದ ವಿನೋದಿನಿ

Prajavani 1971

Translated Version

Translation of article published in Prajavani, Bangalore 1971

“ SUNADA VINODINI” A NEW MUSICAL INSTRUMENT

On the occasion of the second annual music experts' conference of the Karnataka Gana Kala Parishat, many unique innovations were exhibited

Sowma gaana was a special presentation made on the first day's evening session.

All the grammatical texts of music pronounce that music was born out of Saama Veda. Very little efforts are made to understand their inter relationship. From this point of view, the Sowma Gaana presentation by Shri Mysore V Subramanya was significant. To begin with, he explained the background of Vaidika and Loukika singing. He further delineated on the relationship between the Rig Veda and the Saama Veda. He chose certain portions of Saama Veda and the three notations used in reciting of Rig Veda and sang them in the Saama Veda style. While listening to his singing in Saama Veda style, one could experience the seven notes akin to the scale of Kharaharapriya. This was a very significant point.

ELECTRONIC MUSICAL INSTRUMENTS

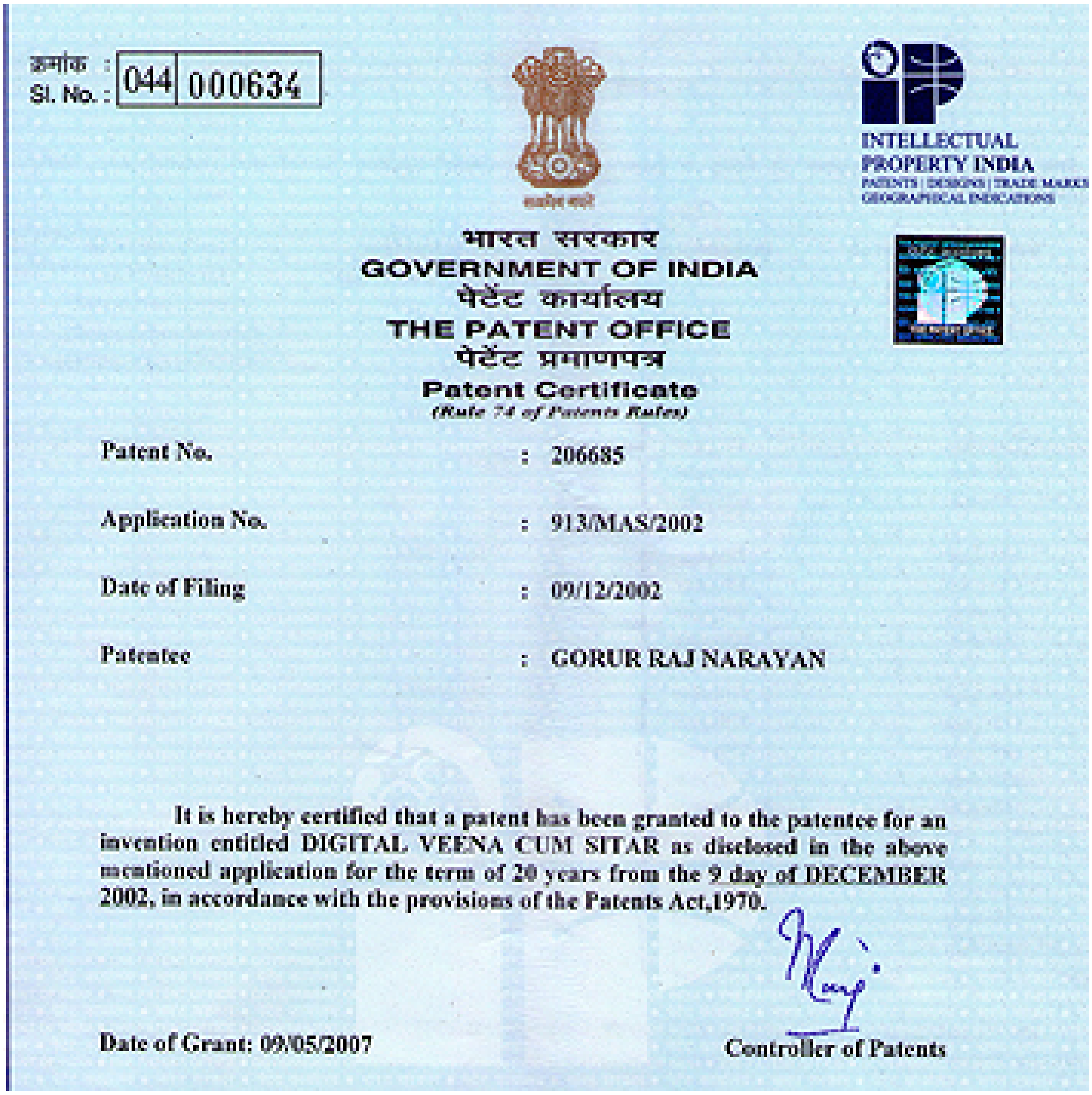
On Tuesday morning, one could feel the impact and the influence of the electronic age on Indian music. The first instrument demonstrated was “Solovox” by Shri T V Karigiri Achar. This unique instrument could produce the sounds of many wind instruments.

The second instrument presented was “Sunaada Vinodini”, a stringed musical instrument constructed on a wooden plank with frets. An electronic gadget is provided to amplify the sound of the instrument. The gourds provided as support to the wooden plank do not serve any other purpose. However, the electronic gadgets help in highlighting with clarity the delicate musical phrases played on the instrument. Apart from this, it was also demonstrated that the instrument could be played with only the left hand without any pluck with the right hand.

Shri G. Rajanarayan, the designer of this instrument is an engineer by profession, but has taken music and new experiments in music as a hobby. This instrument has been displayed in the exhibition organised as part of the conference. - **Murali**

Figure 10.7: Press report about Sunadavinodini

Figure 10.8: Sursingar concert with e-veena



क्रमांक :
Sl. No. : 044 000634

INTELLECTUAL PROPERTY INDIA
PATENTS | DESIGNS | TRADE MARKS
GEOGRAPHICAL INDICATIONS

भारत सरकार
GOVERNMENT OF INDIA
पेटेंट कार्यालय
THE PATENT OFFICE
पेटेंट प्रमाणपत्र
Patent Certificate
(Rule 74 of Patents Rules)

Patent No. : 206685

Application No. : 913/MAS/2002

Date of Filing : 09/12/2002

Patentee : GORUR RAJ NARAYAN

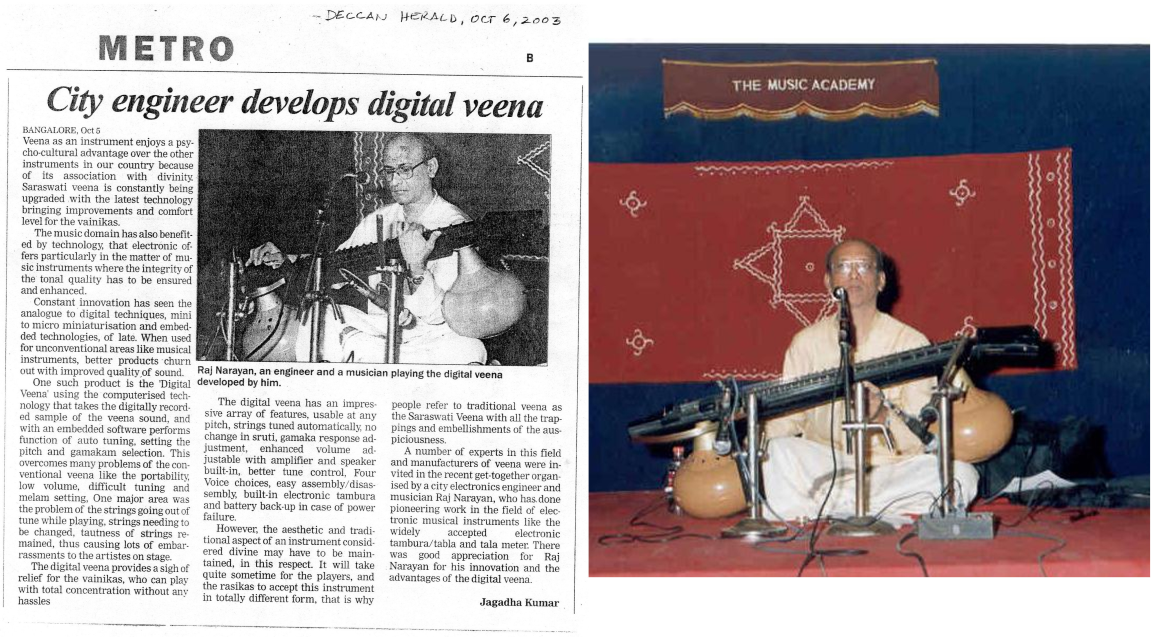
It is hereby certified that a patent has been granted to the patentee for an invention entitled DIGITAL VEENA CUM SITAR as disclosed in the above mentioned application for the term of 20 years from the 9 day of DECEMBER 2002, in accordance with the provisions of the Patents Act,1970.

Date of Grant: 09/05/2007

Controller of Patents

Figure 10.9: Digital Veena Patent

meend' (Pitch bend). This resulted in the tension of the open string being lowered, thereby affecting all the notes. This issue along with the need for string changes corresponding to changes in sruthi led to the invention of the 'Digital Veena' (Indian Patent No. 206685, Figure 10.9). This is a full fledged electronic synthesizer using the physical form of the Veena as the user interface. In other words, the playing strings and frets were only sensors which in conjunction with a pluck sensor and string tension sensors generated a sampled tone sound of the veena as per the exact frequency ratios of a Just Intonation scale. The tension sensors enabled the frequency modulation of the basic note corresponding to each fret and string in proportion to the transverse pull. The instrument therefore completely emulated the features of a traditional veena without its inherent deficiencies. Pitch (Sruthi) of the instrument can be selected through buttons on a keypad without any change in the tension of the strings. All seven strings automatically tune to their relevant notes. The frets never need to be adjusted since the gauge of the strings would never need to be changed. A striking feature of this instrument is that the notes produced on all 24 frets of any string is perfectly aligned with any selected sruthi (tonic), which is impossible to achieve in a traditional instrument. This is due to the fact that the embedded software of the instrument generates the note frequency corresponding to the number (not physical position!) of the fret and the plucked string. This therefore was a revolutionary invention of a classical Indian musical instrument using completely digital techniques and yet retaining the playing interface so that any trained



- DECCAN HERALD, OCT 6, 2003

METRO

B

City engineer develops digital veena

BANGALORE, Oct 5

Veena as an instrument enjoys a psycho-cultural advantage over the other instruments in our country because of its association with divinity. Saraswati veena is constantly being upgraded with the latest technology bringing improvements and comfort level for the vainikas.

The music domain has also benefited by technology, that electronic offers particularly in the matter of music instruments where the integrity of the tonal quality has to be ensured and enhanced.

Constant innovation has seen the analogue to digital techniques, mini to micro miniaturisation and embedded technologies, of late. When used for unconventional areas like musical instruments, better products churn out with improved quality of sound.

One such product is the 'Digital Veena' using the computerised technology that takes the digitally recorded sample of the veena sound, and with an embedded software performs function of auto tuning, setting the pitch and gamakam selection. This overcomes many problems of the conventional veena like the portability, low volume, difficult tuning and melam setting, One major area was the problem of the strings going out of tune while playing, strings needing to be changed, tautness of strings remained, thus causing lots of embarrassments to the artistes on stage.

The digital veena provides a sigh of relief for the vainikas, who can play with total concentration without any hassles

Raj Narayan, an engineer and a musician playing the digital veena developed by him.

The digital veena has an impressive array of features, usable at any pitch, strings tuned automatically, no change in sruti, gamaka response adjustment, enhanced volume adjustable with amplifier and speaker built-in, better tune control, Four Voice choices, easy assembly/disassembly, built-in electronic tambura and battery back-up in case of power failure.

However, the aesthetic and traditional aspect of an instrument considered divine may have to be maintained, in this respect. It will take quite sometime for the players, and the rasikas to accept this instrument in totally different form, that is why people refer to traditional veena as the Saraswati Veena with all the trappings and embellishments of the auspiciousness.

A number of experts in this field and manufacturers of veena were invited in the recent get-together organised by a city electronics engineer and musician Raj Narayan, who has done pioneering work in the field of electronic musical instruments like the widely accepted electronic tambura/tabla and tala meter. There was good appreciation for Raj Narayan for his innovation and the advantages of the digital veena.

Jagadha Kumar

Figure 10.10: Digi-veena demo

veena performer could switch to it without any discomfort. A prototype of this instrument was demonstrated in the experts committee session during the Annual conference of the Music Academy, Chennai in 2002. (Figure 10.10)

Since this was a synthesizer based on digital sound samples, it was possible to incorporate additional voices such as double stringed veena, violin, flute, mandolin, nadaswaram, saxophone, etc., though this was only for demonstrating feasibility.

Digital Harmonium: The successful development of the Digital Veena logically shifted attention to the traditional Harmonium used mostly by Hindustani classical musicians as an accompaniment for vocal music. One of the primary reasons for this attention was the fact that the traditional bellows operated reed instrument was routinely manufactured as per the chromatic (equal tempered) scale whereas Indian music is based on the Just Intonation scale. This subtle difference between the two scales was either not appreciated by most Indian musicians or just ignored by them for the sake of convenience. Some, who were extremely fastidious, insisted on their Harmonium artistes possessing a specially made Harmonium as per the Just Intonation scale with their sruthi as the tonic. The other issue faced by many Harmonium players is that they are comfortable playing only on one specific key (Eg; white 2 or black 2, etc) as the 'Sa' and hence cannot play on other sruthi-s. The so-called Scale Changer Harmoniums addressed this issue, but would work only with an equal tempered scale.

The traditional Harmonium is primarily a wind instrument using banks of reeds and hence they are referred to as 'single reed', double reed' and 'triple reed' types depending on the type of tones desired. Without going into further details, the following main deficiencies of a traditional Harmonium can be listed.

1. It is based on Chromatic scale and hence not accurately suitable for Indian music. If it is specifically manufactured to Just Intonation scale, it is applicable to only one tonic key (Sa) and not playable on any other key as tonic.
2. It is not tunable, if one needs to align it with an unstandard sruthi.
3. It cannot produce 'gamakam-s', which is the essence of Indian music.
4. It is bulky and difficult to carry even for concerts within a city.
5. Being a mechanical instrument with bellows, it needs periodic maintenance.

6. Manufacturing process is extremely labour intensive and involves traditional wood working skills as well as a musical ear for grinding (tuning) of the reeds. Traditional artisans are dying out.

Many western keyboards have the Harmonium voice provided in them, but the facility of a Just Intonation scale is rare. It was therefore decided to develop a digital Harmonium eliminating all the above deficiencies. Further, it was also decided to make all the parts of the instrument standardized so that manufacturing and maintenance costs could be reduced. A prototype of this instrument, based on PCM sampled tones, was demonstrated at the 'World Harmonium Summit' of the Bijapure Harmonium Foundation, Bengaluru, in Jan. 2018 (Figure 10.11). This had the following prominent features:

1. 8 different types of voices simulating single, double and triple reed instruments.
2. Gamak (Pitch Bend) Lever giving a maximum of 8 notes deviation on any key.
3. An LCD user interface to select the pitch as well as designate the 'Sa' key (tonic) which could be any white or black key of an octave. All other notes would be automatically tuned as per the Just Intonation scale.
4. Each key of an octave could be micro-tuned either manually or selected as per the 22 sruthi system of Indian music, thus catering to individual tastes or a characteric raga. Further, this tuned scale can be stored in internal memory for future recall.
5. An Emphasis lever was provided to enable creative intonation of passages by the User.
6. Very light and portable to enable any User to carry it in a back-pack bag even on a 2-wheeler.
7. Volume and Tone controls to set as per convenience.

The reaction of prospective users was very positive and plans are afoot to launch this instrument into the commercial market.

Other instruments - In addition to the instruments described above, an electronic 'Swarmandal', a multi-stringed instrument that simulated the traditional swarmandal, but eliminated the laborious physical tuning of each string, was developed in 2004. This instrument generated the 12 notes of an octave using digital techniques with reference to the sruthi selected by the user and any of these notes could be mapped onto a set of 24 or 36 strings through a simple selection process using a keypad User interface. Different voices with tonal qualities were provided as a choice for the user. A memory for storing any raga composed on the strings was also provided so that it could be instantaneously recalled on the performing stage. This instrument is far more convenient to use than the conventional one.

The 'Swaravali' teaching aid is another unique instrument developed for beginners of Indian music. This plays all the standard practice exercises of Karnatic music (Sarali, Jantai varisais and Alankarams) in three speeds after being set to any tempo and pitch so that it provides a reference to a beginner during practice at home.

10.5 Manufacture of traditional Indian instruments

As already stated in the historical background, most traditional Indian musical instruments are manufactured by individuals or groups of artisans dispersed across India. These artisans have acquired knowledge about their trade over generations in an unorganised manner with hardly any scientific inputs. With the absence of any scientific method of either manufacturing or testing, there is a wide variation in the tonal quality and cost. It has been the practice among renowned musicians to get their instruments manufactured and specially tailored to suit their ears. Eg; position of the frets or the finish of a bridge of a veena would be customised in their presence. Even to this day, various types of harmoniums are manufactured in a 'suit and match' process where there is no standardisation or interchangeability of parts and sub-assemblies. This leads to a labour intensive process that yields non-uniform parts. The resulting small volumes of production and sales makes the trade unattractive for younger generations of the artisans. This is leading to a shortage of skilled workers and instruments of good quality. There is a dire need for the infusion of engineering and technology in the following areas:

Figure 10.11: Demo of digital Harmonium 2018

1. Use of affordable electrical machines, such as circular saw, jig saw, router, planer, sander, spray polishers, etc. This would increase production and lower costs.
2. Standardisation of parts, sub-assemblies and designs so that these could be manufactured by out-sourced vendors in volumes at lower cost. Eg; Tuning knobs and bridges of sitars, veenas, sarod, etc.,
3. Gourds which serve only as supports and do not form the resonant chambers, can be made of alternative materials such as plastic and moulded in numbers so as to reduce labour and cost. So is the case with tuning knobs.
4. Improve quality of sound and manufacturing and reduce wastage by using standardized processes and using measuring instruments such as ultrasonic thickness gauges, verniers, audio level meters, etc. Govts of each State can facilitate the absorption of technology through a common facilities center in Clusters where such equipment can be made available to manufacturers.
5. Establishment of specialised training centers for infusion of science and technology into the products as well as processes. Better understanding of acoustics would also lead to improved designs and better tonal quality.

10.6 Reactions & Experiences of users

As already stated, purists of Indian classical music have in general been averse to the introduction of modern technology in the evolution of Indian musical instruments. Hence, considerable resistance existed for the acceptance of the instruments developed. It took at least 15 years before many musicians started using it freely and hence the growth of this industry was extremely slow. It was primarily the passion of this author and his wife, Smt. Radhika, that helped sustain the manufacture as well as promote further development of these electronic Indian instruments. However, it must be acknowledged that many reputed maestros were very supportive as well as appreciative of these efforts. Some of them are Dr. Balamuralikrishna, Veena maestro Dr. S. Balachander, Pt. Ravi Shankar, Ustad Vilayat Khan, Ustad Amjad Ali Khan, Sri. KV Narayanaswamy, Sri. Lalgudi G. Jayaraman and many others, who even went to the extent of making valuable suggestions on improving the features built into the products.

At the launch of the electronic lehera instrument in Bengaluru in 1993, sitar maestro Pt. Ravi Shankar very affectionately appreciated the utility of electronic instruments and humorously added, “I hope Raj Narayan does not come up with a machine that can play the sitar and replace me”! At a workshop held in Bombay in 1992, Ustad Vilayat Khan commented with reference to the electronic tabla, “Hindustani musicians no longer have any excuse for not practicing enough”. Ustad Amjad Ali Khan wrote in his letter, “These instruments are proving to be very useful”. During a concert in Bengaluru, Vidwan TV Sankarnarayanan, with a smile, kept switching quickly between panchamam, madhyaman and nishadam on his electronic tambura as he sang a ragamalika, something impossible on a traditional instrument.

Yet, there are many even today who hesitate to use some of the electronic instruments on stage due to the fear of being perceived by audiences as ‘too modern’ or ‘tradition breakers’. Even manufacturing processes of traditional instruments have not received the engineering inputs that they truly can benefit from. These issues deserve more attention from groups comprising scientists, engineers as well as progressive minded musicians. Indian music can then scale new heights of creativity without affecting purity and bhakthi ingrained in it. The ultimate objective should be to increase the listeners’ experience as well as the performer’s convenience using the best of modern science and technology.

Articles on the Divine Veena: [1]

[1]
https://www.academia.edu/37850149/Divine_Musical_Instruments
https://www.newindianexpress.com/cities/chennai/2012/aug/13/veena---the-divine-instrument-396488.html
https://pavanimehra.wordpress.com/2017/10/12/indian-mythology-and-music/
https://www.jayanthikumaresh.com/about-the-veena/
https://mme.iitm.ac.in/vsarma/personalweb/veena2.html

Chapter 11

Current Practices and Perspectives in Karnatik Music

Vignesh Ishwar, Rithvik Raja

11

Abstract

The general notion about anything traditional, classical or artistic in the Indian ethos, is that it is over two thousand years old. This is also the case with Indian art music traditions, namely Karnātik music and Hindustani music. Karnātik music as we know it today can be roughly traced back to the 17th century thus making it not more than 400 years old. It is then important to ask the question as to what can be defined as recent practices and perspectives for such an art form. The performance of Karnātik music has evolved to what it is today over the last century. It is the culmination of efforts and changes brought about by many artists and institutions. In the 21st century, with the advancements in technology, art is being consumed through various mediums. The use of high quality microphones and other forms of technology coupled with the current COVID-19 pandemic situation, there is a surge in online consumption of art and music which has led to an evolution in the way Karnātik music is being presented and received. This article will cover the brief history of Karnātik music and its evolution over the past 400 years, along with the practice of the artform over the last century, which has led to the understanding of Karnātik music as it is presented today.

11.1 A Brief History of Performance

Karṇātik music is the art music tradition from the southern region of the Indian subcontinent. It has a history dating back to the —-th century. This article deals with the current perspectives and practises in Karnātik music performance. Before we delve into current practises, it is important to look at the history of performance in Karnātik music.

Art traditions in India are generally termed as ancient, but, when we look at the timeline of events occurring leading to the Karnātik music we listen to today, it is very important to trace the practises back in time and determine as to what we can call ancient and what we can call recent or current.

In order to look at the history of performance in Karnātik music, we must look at music before and after the trinity of Karnātik music, Tyāgarāja, Muddusvami Diksitar and Syāmā Sāstri. Looking back before the 18th century (the time period of the trinity), the Tanjāvur belt, and more importantly the Nāyak period (1532-1673) [Southern Music] is an important period when Karnātik music flourished. Many treatises exploring and formalizing various theories in the art form were written during this period. Performances were also patronized majorly in Tanjāvur and other princely states such as Travancore, Mysore, Rāmanāthapuram, Pudukottai, Ettayapuram, Bobbili, Karvetinagaram, Vijayanagaram and Pithapuram. Both dance and music were patronized in these states [Southern Music].

11.1.1 Performance Pre-Trinity

According to the Chaturdandi Prakasika (a 1620 treatise on Karnatik music by Venkatamakhin) [TMKPaper, Chaturdandi] Karnatik music presentation had 4 parts, alapa and thaya, which were believed to be spontaneous improvisations and prabandha and gita which were composed music. It is also clear from literary pieces written on the Chaturdandi Prakasika that the music described in it was presented mainly on the instrument vina, though in ancient times, the term vina referred to any generic stringed instrument [Southern Music].

There are detailed accounts of how vainikas(vina artists) prepared their instruments before a performance and interestingly, artists today also prepare their instruments in a similar manner. There are also appreciation terms such as 'bale tayam' and 'aha rakti pada' recorded. 'Bale' and 'Aha' are used even today for appreciation in a performance. [Southern Music]

From the various poems and stories written by court poets and courtesans and sometimes kings, valuable information of performance practices of the Nayaka period can be deciphered. Music, dance performances and discourses and discussions were held in an auditorium in Tanjavur during these times. Dance was extensively patronized by the Nayakas. It is a fact that performances were mainly by women artists. It is in fact accounted for that women artists have played almost all musical instruments including percussive instruments. The distinction between dancers and musicians was not so apparent at the time. Both music and dance were a part of the artform as a whole performed mainly by the *Dēvadāsi*[1] community [Southern Music]. The late Nayaka and early Maratha paintings (1660-1700) give an insight into how a musical orchestra was. The frescoes in the Tiruvarur temple describe the 'chinna melam' a small orchestra with the Devasis, nattuvanars, mridanga, flautist, tutti and a shankha[Southern Music][Tanjavur as a seat of music]. These performances were generally a part of a procession or some event.

With the advent of the Maratha rule in Tanjavur, the music also seems to have changed towards what we know as Karnatik music today. Karnatik music was not presented as we know it today in the first half of the 18th century [Southern Music]. The Maratha period brought about development of various ragas from various regions. The travels of Kshetrayya (a composer known for his compositions in compositional form called pada) to Tanjāvur brought about the compositional structure of pallavi, anupallavi, charana that we know today in keerthanas. It is during this time that the compositional forms used in the dance(Sadir) would have changed. It is to be noted that the compositional forms of pada, varna, savarajathi, thillana etc. were predominantly connected to Sadir(dance) and not Karnatik music as an art music practice. The advent of the composition form tana-varna in the late 18th century indicates the focus shifting towards purely musically intended compositions branching away from Sadir(dance)[Southern Music][Tanjavur as a seat of music].

[1] The Dēvadāsi community....

Chapter 12

Applause and Aesthetic Experience

Mundur V. N. Murthy

The Institute of Mathematical Sciences, Chennai 600 113 (Retired)

Long ago, when I was a student in Mysore, my friend, Manu Shetty (an artist and a philosopher), said that the aesthetic experience in the Indian art, especially in its architecture, is in detail where as in the European art the real experience is in the perspective. The experience is dictated by where you stand with respect to the object of admiration.[1]

As a lay person who is interested in art and music, I would like to know if there are some parallels in the way we appreciate art and music in general. In a limited exercise, I want to look at architecture and music in particular and the inherent cultural influences in the way we appreciate finer points. Some generalisations, from the limited number of examples, are necessary but I am aware that there are always exceptions[2].

12.1 General background

Some thirty years ago, my wife Hema and I were in the middle of St. Peter's square (Piazza San Peitro, actually it is almost a circle) in Rome standing in front of the Basilica facing the large circular edifice with hundreds of columns all along the perimeter.

At the center of the square is a four-thousand-year-old Egyptian obelisk, erected at the current site in 1568. Gian Lorenzo Bernini designed the square/circle almost 100 years later, including the massive Tuscan colonnades, four columns deep, which embrace visitors in "the maternal arms of Mother Church." A granite fountain constructed by Bernini in 1675 matches another fountain designed by Carlo Maderno dating to 1613.

To appreciate the grandeur you have to situate yourself in a position from where the symmetry of the structure as a whole is visible and obvious. While taking a walk all around, the columns themselves appeared rather dumb and monotonous without any character. It is the whole and not the individual parts that provide the aesthetic experience of the beauty and grandeur of St Peter's square.

This was made even more apparent recently, in the year 2013 to be specific, when we were at the Acropolis in Athens. Even in the middle of all the ruins accumulated over centuries of building, destruction and rebuilding, the Parthenon stands tall on the small hill in the middle of Athens.

The Parthenon is a temple on the Athenian Acropolis, Greece, dedicated to the maiden goddess Athena, whom the people of Athens considered their patron (or matron) deity. Its construction began in 447 BC when the Athenian Empire was at the height of its power. It was completed in 438 BC, although decoration of the building continued until 432 BC. It is the most important surviving building of Classical Greece. Its decorative sculptures are considered some of the high points of Greek art. The Parthenon is regarded as an enduring symbol of Ancient Greece, Athenian democracy and so called cradle of western civilization. It has gone through many phases of building, destruction and rebuilding. Almost every proportion in these buildings seem to be in Golden ratio–considered aesthetically the most desirable proportion[3].

Its massive pillars, even without the supported super-structure (destroyed long ago), ignite the imagination enough to reconstruct what it would have looked like from a distance. Each pillar, though impressive in size, again is rather simple looking cylindrical object of stone with long glitches, minimalist at best. But the

[1]This chapter is an expanded version of an article that first appeared on the compmusic website: https://compmusic.upf.edu/ar/node/151

[2]In here applause is a generic reference to various forms of appreciation communicated by the audience to the performer

[3]The golden ratio is solution of the equation $\phi = 1 + 1/\phi$ and has the value $\phi = 1.618033...$ When two length dimensions are in this ratio it is supposed to be aesthetic.

Figure 12.1: Saint Peter's Square in Rome-Photo by DAVID ILIFF. License: CC BY-SA 3.0; Reproduced from https://en.wikipedia.org/wiki/St._Peter_Square

whole, in its shear size and conception, is mind boggling. There is only one way to appreciate- from a suitable distance from where the whole is framed by the eye.

Manu's thesis becomes apparent when the Greeco-Roman architecture is compared with the South Indian temple architecture from the Pallava monuments to the Hoysala temples, an evolution over a period of thousand years taking in many influences, becoming more and more intricate over time. The duality of scales, between Greeko-Roman and Indian, plays out in fullness- grandeur versus intricate, distant versus proximate, macroscopic versus microscopic.

This is especially so with later period Hoysala temples at Belur and Halebid– often referred to as the poetry in stone. It is a culmination of the development of the stone art in South India combining various influences. I have shown in the picture, the later, more complete example, of the Hoysala architecture located at Somanathapura–a town located 35 km from the historic Mysore city. Somanathapura is famous for the Chennakesava Temple built by Soma, a commander in 1268 CE under the Hoysala Empire King Narasimha III, when the Hoysalas were a major power in South India. The Keshava temple is one of the finest and the most complete examples of Hoysala architecture and is in a very well preserved condition.

From a distance, on a scale that the perspective of St.Peters square or Parthenon is enjoyed, the temples may appear small and fairly non-descript, like any other temple in the region. However, as one goes closer each panel from top to bottom and left to right is littered with exquisite carvings frozen in stone revealing layers upon layers of figures and panels each of which tell a story. Unless the temple is incomplete, there is no empty canvass in stone.

As one moves around these temples, the details lead one gasping with exclamations. As one completes a slow revolution around these marvels in stone, it is time to pause in front to admire and applaud the creators taking in the whole, but in a sense which is entirely different from the experience of standing little far away in front of the Parthenon in Acropolis or in the center of St. Peter's square in Rome.

Figure 12.2: Parthenon at Acropolis in Athens. Photo by author.

Figure 12.3: Stone reliefs at the later Hoysala temple at Somanathpur near Mysore. Photos by author.

Rarely the two have merged in temple architecture, perhaps the Brihadeeshwara temple in Tanjavur is an exception to the rule. Completed in 1010 AD, it is known as the big temple. The Kumbam or apex of the temple is carved out of a single stone weighing about 80 tons.

More prominently, in the later Islamic period monuments in India, Central Asia we have some examples of a systematic synthesis of both traditions, differential versus integrated, but not necessarily influenced by them. It is said that Timurlane of Timur dynasty (14 Century), for all his bloody conquests, valued craftsmen and artisans in each defeated territory so much that he spared them from death and employed them to build grand, imposing architectural structures leading to the fusion of many different styles in his empire (around modern-day Uzbekistan).

The Taj Mahal is a classic example of the fusion of the scale, symmetry and the intricate as also the Humayun tomb in Delhi.

Figure 12.4: Humayun's tomb in Delhi. Photo by Muhammad Mahdi Karim -GNU Free Documentation License, V1.2, Reproduced from https://en.wikipedia.org/wiki/Humayuns_Tomb

12.2 Appreciation of Music

This spatial appreciation of architecture seems to find a mirror in music appreciation, but in time. Some time back, we were listening to an young and established Carnatic musician, T M Krishna at a concert in Music Academy. We are ardent listeners of Krishna's music, not only because it is very good but also because he challenges the orthodoxy of the so called established traditions. The main piece was in raga *Kambodhi* which included the raga elaboration followed by a *padam*. The padam part was controversial since it is usually sung at the end of the concert without much flourish. The content is usually romantic and not necessarily devotional as traditional compositions like *Keerthanas* of the Trinity of karnatic music.

But it is not what he chose to sing that is contextual, but the way he sang. The piece included some exquisite phrases rendered in a very melodious style with soft gamakas- the audience audibly gasping at the beauty of the phrases followed some times by some energetic phrases demanding instant applause. The audience appreciation, obvious throughout the rendering, was appropriate- it was subtle but clear when the rendering was soft full of gamakas, boisterous and loud when the phrases included fast brighas. Krishna's rendering demanded appreciation here and now much like the temples such as Belur and Halebid or Somanathapura. Of course the conclusion of the piece again deserved an applause which was the sum total of the whole experience. This is not unusual in a Carnatic music concert. Any random selection of renderings by very good performers, there are many now, are indicative of the above trend more or less. The good musical passages are appreciated then and there in varying modes. I know for a fact that the performers also enjoy this instant gratification of their offerings and it encourages them to perform even better.

This is so unlike a performance of European classical music. The audience quietly and expectantly sits through a performance. There is a quiet anticipation and even a nod at some finer points of the music. But the music rendition is never disturbed. It will have to wait for the edifice to be built by an ensemble of musicians and completed. It is only at the end, when the whole process of construction of the musical piece is completed, the audience stands and applauds probably after a gap in time, some times demands even more. It is in-built into the nature of compositions themselves. The compositions like the Jupiter by Mozart (symphony No. 41) or the Emperor by Beethoven (Piano Concerto No. 5) demand that you listen in absolute silence, look at the musical edifice as a whole and applaud the integrated experience in the end. Same is the case with an opera like Carmen rendered by Maria Callas- the movie “Philadelphia" has many snatches of this opera buttressing some very poignant moments in the film. One of my favourites in this category is the string quartet “American” by Antonín Dvořák. One can think of many more examples.

Of course, there are always exceptions to the rule as in any creative field. It does not mean that one can not appreciate, here and now, compositions in European classical music in detail- some obviously are extremely finely crafted. For example the Mozart symphony in G-minor (25th) is exquisite in detail and has been copied in parts in so many places without sacrificing the pleasure of the whole. Nevertheless, the applause is always reserved for the end.

Indeed, there are some compositions also in Carnatic music that coax the listener into silence till the very end. This was brought home to me in a concert by Prasanna Venkataraman in Music Academy - he was rendering a master piece by Muthuswami Dikshitar in the raga Brindavana Saranga in praise of Soundararaja, the presiding deity of the Nagapattinam temple. It was a delicate and beautiful rendition shorn of acrobatics with both violin and percussion following in a subdued manner. The conclusion was followed by a few seconds of silence before the applause broke out. The silence was also part of the musical experience much as it is at the end of a European classical music experience.

Applause can become a nuisance when it interferes with aesthetic experience, as it often happens when listeners tuned to Carnatic music in a European classical music concert leaving the performers, not familiar with the local custom, baffled when the concert is interrupted with applause by locals. Conversely one finds many foreigners in December concerts in Chennai looking bewildered when the “here and now applause" breaks out at frequent intervals. Thinking along similar lines, the appreciation shown at the end of a performance of European music some times reduces to a ritual so much so that it continues on until the performers come back and play a short piece even when it is not needed!

Many Indian musicians in these times, following the European tradition, urge the audience not to show their appreciation in the middle as it disturbs the flow of music. For some it may indeed be the case. However, experience shows that most musicians would like some sort of instant feed back which fires their imagination further. Ustad Vilayat Khan, playing Rag Marwa in Royal Albert Hall, London, laments that he can not see the faces of his audience in the darkness of the auditorium. He says it would be nice to see the expression on the faces of his admirers.

How did this mode of appreciation come about? Perhaps the answer lies in the origins of Indian music, more specifically Carnatic music, lies at the heart of the oral tradition. In any oral tradition knowledge is often communicated and passed on through music of some form and vintage. This is one of the main reasons that musical traditions are stronger and complex in societies where oral tradition dominates.

On its own the music system may further develop its own complicated structure and grammar. This helps not only in communication but also in preserving knowledge in the absence of elaborate and dominant writing systems. Communication, to be effective, also depends on repartee or at least an acknowledgement of understanding. Any teacher would agree that it is unnerving to teach a class of students who are absolutely silent and hold their thoughts till the end. Translated in to a musical performance this amounts to repartee among musicians and an acknowledgement of understanding among connoisseurs of music through a nod or even an applause.

As in the case of Islamic architecture, the twine have indeed met- perhaps in the Hindustani music where one again sees a strong influence of Turcic and Persian cultural traditions. Here an approximate median is achieved since there is a gradual build up from a slow paced alap to the crescendo created by a Tarana or a Jhala. Appreciation of the parts and the whole go together without sacrificing the aesthetic experience–a combination of grandeur and the delicate.

It is therefore fair to say that there is a method in the “here and now" appreciation rooted in the larger art and architectural context in the sub-continent developed over centuries.

Chapter 13

Hindustani rhythm and computational analysis: A musicological perspective

Martin Clayton

Department of Music, Durham University, UK

13

13.1 Introduction

The focus of this chapter is on tala and laya – metre and rhythm, in terms of Western musicology – in Hindustani music. Below I introduce the topic and then outline some areas of current research interest: this presentation is illustrated with examples from publicly shared research corpora ([CLT21]). Reflecting on this work, I conclude by making some suggestions for future computational rhythm analysis in Hindustani music. Firstly however, I offer some introductory thoughts on the nature of music theory, and the broader contribution of computational approaches, from a musicological perspective.[1]

13.1.1 On computational musicology and music theory

Music theory – from the Greek theoria, 'action of viewing, contemplation, sight, spectacle' (Oxford English Dictionary) – is a multifaceted and complex phenomenon.[2] When applied to music the term is broader than what is normally understood as scientific theory: that is, an explanation of the physical world that can be empirically tested. Music theory rarely fits this modern scientific paradigm: it does not seek to explain music in testable ways. Rather, it encompasses a wide range of contemplation and speculation about music. In some societies the articulation of music theory has a relatively minor role: ideas about the right way to play or sing can be conveyed by example, without much recourse to language or graphical devices. In other places and times, overt theorisation becomes more prominent. In these cases people develop names, classification systems, notations and other means, initially perhaps as a practical tool to help them to remember music accurately, but these theories take on a greater significance. The role of theorist emerges in its own right, and since theory can never encompass practice in all its richness, it develops its own internal logic and may diverge significantly from practice. In societies with long traditions of written music theory – India is one of the stand-out examples globally – we can use these texts to trace music's history. In doing so we need to be mindful of the fact that theory is autonomous: sometimes it lags behind, or drifts away from practice. Nonetheless, as musical practice changes, as it has done dramatically over the two millennia of documented music theory in India, so too does theory, and these texts provide more information about that process of change than any other surviving evidence.

Another important aspect of music theory is that it can take on an important role of validating practice at both an individual and societal level. Thus, a musician may be praised because he or she performed in accordance with theory; a society may even congratulate itself on the excellence of its music, citing theory as evidence. At this point theorists may come to see their role in terms of such validation: music theory is the body of knowledge that explains to a society's members why their music should be valued, and perhaps, why the qualities of their music demonstrate the superiority of their culture.

[1] I would like to thank Prof. Preeti Rao and Nithya Shikarpur for their invaluable help with the visualisations in this chapter.

[2] Indian theoretical traditions, and the connotations of terms such as shastra, do not align exactly with concepts derived from Greek in western traditions. This lies beyond the scope of this chapter, but I suggest that the argument outline here holds for both contexts.

This is a process that has been critiqued extensively in the case of Western (Eurogenic) music. Notation, which developed as a practical tool to aid memory and ensemble coordination, came to be seen in its own right as a mark of advanced civilisation. Musical elements seen to be characteristic of European elite music were defined by its theorists as markers of advanced, highly evolved music, and offered (in a circular argument) as proof of the more highly developed status of their own tradition. As academic critics are now beginning to argue, music theory played a role in the perpetuation of myths of racial superiority that justified European colonial domination and racial segregation ([Ewe21], [Wal20]). Indian music theory has been instrumentalised in similar ways. Particularly from the second half of the nineteenth century, music theory and notation became associated with nationalism and resistance to (or accommodation with) colonialism, but also with claims of specific communities to be the true 'owners' of the tradition, and with the exclusion of groups such as courtesans and many hereditary (often Muslim) performers. The supposed lack of theoretical knowledge of these performers was used to justify this marginalisation ([Bak05], [Kat12], [Neu12]).

Theory is thus a hugely important part of the Hindustani music tradition. It is diverse: it includes written textbooks and collections of bandishes (compositions), academic discourse, critical commentary and the knowledge that expert musicians transmit orally to their students. It has practical uses as well as giving us a language to talk about an extremely rich body of shared knowledge. Music theory is also a body of discourse that is shaped by its own history, a history which could potentially bias it in various ways. Theory is a necessary but treacherous tool: we cannot understand music without referring to it, and yet it also has the power to distract and mislead us, even to doubt the evidence of our own ears. As academic researchers our aim should be to build a critical music theory – one which aims to understand music both as performed and as thought and discussed, and to reflect on the relationship between the two.

How could computational methods assist us in doing this? Computational tools offer the musicologist empirical evidence of the ways in which music has been performed, evidence that can be contemplated in more detail than is possible in the moment of performance. But how should we use this opportunity, and in what relationship should computational analysis stand with respect to existing bodies of music theory? A key point here is that we should not limit ourselves to using computational methods to locate, in recorded music, those things which existing music theory tells us are present. In other words, while algorithms to identify ragas and talas, or to match performances against theoretical scales or chalans, are extremely valuable, we should not feel that this is the only possible approach. We might plot the relationship of new studies to existing theory schematically in three categories, as follow:

	Approach	**Example tasks**	**Uses**
Top-down	Search audio recordings for music-theoretical concepts.	Raga and tala identification	Help organise large corpora and save time-consuming work of identifying items
Bottom-up	Ask what patterns and invariants could be discovered in the music in a data-driven fashion	Pattern discovery, in melody or rhythm, from first principles	Reveal important patterns not otherwise recognised by theory, potentially revealing (for example) cognitive processes.
Critical	Explore the gaps between theoretical concepts and practice.	How adequately do raga definitions describe musical practice?	Improve music theory by testing it against practice

My third approach, labelled 'Critical', seems to me to combine the power of computational data analysis with a critical approach more typical of humanities research. It seeks to understand theory critically by making the relationship between theory and practice its concern. It may work in a bottom-up fashion, exploring regularities in performance data; but it then asks, and how does this pattern relate to previous knowledge? Is this something known but ignored because it does not fit dominant paradigms, or because it is simply too hard to put into words? Does it potentially reveal unconscious musical thought? It may work in a more top-down fashion, exploring theoretical categories but also seeking out gaps between theory and practice. Again, it can ask, why is this gap between theory and practice not recognised: what kinds of understanding might written theory be suppressing, and why? This would allow us to see music theory in a deeper perspective still: as an abstraction that captures some parameters of practice well, others poorly, but that will always be a simplification. Later in the chapter I will discuss some topics in the domain of time and rhythm rather than pitch, but the latter domain offers some familiar examples. An obvious point of interest in Hindustani music is the exploration of intonation, and indeed much earlier computational research occupied itself with testing theoretical shruti systems against modern practice (see e.g. the early pitch tracks and histograms computed by [Bel88]; see also [Mee00]). This work will surely continue, but work on scale and intonation can also work on a macro level. Nazir Jairazbhoy once argued that Hindustani raga scales evolve over time as a result of the patterns of consonance and dissonance of different intervals – a rare example of a music theory that makes testable predictions (1971)[Jai71]. What could we learn about this evolutionary process if we had the tools to study large corpora of recordings efficiently – to see how the patterns and motifs, or the intonation of specific notes used in a particular raga vary between performers or over time?

Computational musicologists need to be able to see a bigger picture: our job is not simply to operationalise existing theory, but to take a critical view of existing music theory and use that knowledge in order to dig deeper into practice. We can enrich theory in ways that allow us to see variability and historical change in perspective, to relate our insights to other scientific areas such as cognitive science and psychoacoustics, as well as to compare different forms of music.

13.1.2 On corpus research

In order to explore the richness and complexity of practice, we need the ability to scale up computational analysis from extracts or single performances to corpus scale. We have seen similar processes in medical, astronomical and many other kinds of data: some patterns only emerge when we have the technology to process data far faster than a human expert can manage. There is no reason the same should not be true of music performance data. Realising this in practice requires both the existence of suitable corpora, and development of analytical processes that can be automated and scaled up.

The former may seem a trivial problem – there is a huge amount of recorded music in existence, after all. In order to develop computational approaches and gradually expand the scale of analysis, however, we need dedicated research corpora. Different research groups have taken complementary approaches to this issue. Xavier Serra's CompMusic project is an important example: CompMusic's Duniya database concentrates on enabling access to existing audio recordings of Hindustani music; additional data including markup of the tala structures is also shared. This has the advantage that different projects can study the same recordings, and the database is set up in such a way as to facilitate computational analysis (https://dunya.compmusic.upf.edu/).

The complementary approach I have taken with colleagues is to make new audiovisual recordings, obtaining informed consent from the musicians for their research use, and to share these freely with the research community via Open Science Framework ([CLT21]). At the time of writing our Hindustani music collections comprise 17 complete raga performances, vocal and instrumental, offered as multitrack audio, video (in most cases including more than one static view) and accompanying annotations. Annotations include the tala (i.e. the location of each sam). For those used in the Interpersonal Entrainment in Music Performance (IEMP) project, we also include computationally extracted event onset times for each instrument of the instrumental duos (e.g. sitar and tabla), and provide composite annotation tables including onset times for each instrument labelled with beat and cycle (matra/avartan) numbers alongside onset peak levels and estimations of event density. For some recordings we include extracted movement data, and/or manual annotations of observed gestures or interactions between artists. For the rupak tal khyals sub-collection, we have shared also the R code used for their analysis ([Cla20], [CLT21]).

The advantages of this approach are that video data can be consulted (and movement data can be extracted directly, using pose estimation algorithms such as OpenPose, see [Cao+21]); that multitrack audio allows detailed analysis of each instrument (or voice), and the testing of source separation algorithms; and that annotations are open-ended and can be expanded and enhanced by future projects. By using this material, users agree to a licence that commits them to use "only for the purposes of not-for-profit research or teaching or personal educational development".

Recording ID	Soloist	Instrument	Genre	Co-performers	Contents	Place and date	Duration	Annotations
IEMP North Indian Raga								
ABh_Puriya	Anupama Bhagwat	Sitar	Gat	Gurdain Rayatt (tabla)	Rag Puriya, alap, vilambit teental, drut teental	Gateshead, 29/9/2013	56’ 35”	Metre, Onsets, Structure, Optical flow
DBh_Malhar	Debashish Bhattacharya	Slide guitar	Gat	Gurdain Rayatt (tabla)	Rag Miyan Malhar, alap; dhamar tal, teental, ektal, drut teental	Durham, 31/5/2016	44’	Metre, Onsets, Structure, Optical flow
NGh_Tabla	Nayan Ghosh	Tabla	Tabla solo	Hiranmoy Mitra (harmonium)	Vilambit teental, Drut teental	Kolkata, 3/2/2007	64’09”	Metre, Structure, Optical flow
PrB_Jhinjhoti	Prattyush Banerjee	Sarod	Gat	Gauri Shankar Karmarkar (tabla)	Rag Jhinjhoti, alap; rupak tal, drut teental	Kolkata, 28/1/2007	55’ 24”	Metre, Onsets, Structure, Optical flow
SCh_Malhar	Sudokshina Chatterjee	Vocal	Khyal	Kaviraj Singh (harmonium), Gurdain Rayatt (tabla)	Rag Miyan Malhar, vilambit ektal and drut teental	Durham, 2/3/2016	33’01”	Metre, Structure, Optical flow
ShKh_Jhinjhoti	Shujaat Khan	Sitar	Gat	Shahbaz Hussein (tabla)	Rag Jhinjhoti, alap, vilambit teental, drut teental	Gateshead, 28/9/2012	56’ 52”	Metre, Onsets, Structure, Optical flow
VK_Multani	Vijay Koparkar	Vocal	Khyal	Viswanath Shirodkar (tabla), Seema Shirodkar (harmonium)	Rag Multani, vilambit ektal and drut ektal	Mumbai, 20/05/2005	49’28”	Metre, Structure, Optical flow
VS_Bhoop	Veena Sahasrabuddhe	Vocal	Khyal	Viswanath Shirodkar (tabla), Seema Shirodkar (harmonium)	Rag Bhoop, vilambit ektal and drut teental (Khyal)	Mumbai, 20/05/2005	57’25”	Metre, Structure, Optical flow
NIRP1 Rupak Tal Khyals								
ArunBh_Kedar	Arun Bhaduri	Vocal	Khyal	Tarak Saha (tabla), Gourab Chatterjee (harmonium)	Rag Kedar, alap, rupak tal, drut ektal	Raniganj, 14/2/2007	26’12”	Metre, Tabla onsets
MAK_Jaun	Manjiri Asanare-Kelkar	Vocal	Khyal	Milind Pote (tabla), Chinmay Kolhatka (harmonium)	Rag Jaunpuri, alap, rupak tal, drut teental	Pune, 9/12/2006	30’30”	Metre, Tabla onsets
MAK_Jhin	Manjiri Asanare-Kelkar	Vocal	Khyal	Milind Hingne (tabla), Shubhash Dasakkar (harmonium)	Rag Jhinjhoti, alap, rupak tal, drut teental	Nashik, 13/2/2010	27’46”	Metre, Tabla onsets
MAK_TilakK	Manjiri Asanare-Kelkar	Vocal	Khyal	Vishwanath Shirodkar (tabla), Kaviraj Singh (harmonium)	Rag Tilak Kamod, alap, rupak tal, drut ektal (tarana)	Durham, 30/4/2014	20’41”	Metre, Tabla onsets
MurA_DhaniK	Murad Ali (sarangi)	Vocal	Khyal inst.	Gurdain Rayatt (tabla)	Rag Dhani Kauns, alap, rupak tal, drut teental	Gateshead, 24/9/2012	10’29”	Metre, Tabla onsets
RamD_RagBah	Ram Deshpande	Vocal	Khyal	Viswanath Shirodkar (tabla), Anant Lakhe (harmonium)	Rag Rageshree-Bahar, alap, rupak tal, drut teental	Nashik, 13/2/2010	28’02”	Metre, Tabla onsets
VK_HindBah	Vijay Koparkar	Vocal	Khyal	Ajinkya Joshi (tabla), Rahul Gole (harmonium)	Rag Hindol-Bahar, alap, rupak tal, drut teental	Pune, 25/2/2010	16’15”	Metre, Tabla onsets
VK_RagK	Vijay Koparkar	Vocal	Khyal	Ajinkya Joshi (tabla), Rahul Gole (harmonium)	Rag Rageshree-Kauns, alap, rupak tal, drut ektal	Pune, 25/2/2010	48’28”	Metre, Tabla onsets
VS_Hams	Veena Sahasrabuddhe	Vocal	Khyal	Viswanath Shirodkar (tabla), Seema Shirodkar (harmonium)	Rag Hamsadhwani, alap, rupak tal, madhya lay teental, drut ektal (tarana)	Pune, 15/12/2006	30’36”	Metre, Tabla onsets

Table 13.1

13.1.2.1 OSF Collections of Hindustani music

The remainder of this chapter makes use of extracts from these recordings, with annotations from the published collections. The static illustrations here can also be explored interactively in the form of Jupyter Notebooks that directly use the OSF media and annotation files. See Table 13.1.

13.2 Time and rhythm in Hindustani music

In the remainder of this chapter I will first summarise some basic rhythmic features of the music, illustrating these points with clips from our published OSF materials. This is not a step by step introduction to tala theory, rather an illustration of features highlighted in [Cla00b] in order to demonstrate some of the potential of data visualisation. I then summarise some more recent research on tala and rhythm and discuss prospects and challenges for future research. Although rhythm without tala (i.e. alap) is a fascinating topic in its own right (as noted in [Cla00b]: 94ff), the following material focuses on rhythm in music with tala.

Talas, which we might render in English as 'metric cycles', take diverse forms in Hindustani music. Although the vast majority of the music can be described as using no more than a dozen talas, these structures differ in the number and organisation of matras (basic time units, roughly translated as 'beats'), and particularly in tempo. Calculated as matras per minute, tempo ranges from about 10 to more than 700, an extraordinary ratio. What this means in practice is that the matras are either subdivided (at slow tempi) or grouped together (at fast rates), to produce the perceived tactus pulse (tactus is a Western musicological term meaning the rate of pulsation the listener is most likely to tap or clap along with). The fact that tala cycles range from less than 2 seconds to more than a minute in duration is another significant factor, and together they mean that talas are very diverse in practice. London points out that metres are distinguished not only according to the number of beats per cycle, with (in Western music) many Tempo-Metrical Types ([Lon14]: 153–155): in Hindustani music similarly, a named pattern (such as teental) exists in very different forms at different speeds.

Talas are recognisable in practice in two complementary ways – by the basic patterns of drum strokes called theka, and by distinctive patterns of hand gestures. The usage of these two elements is also tempo-dependent: slow-tempo pieces tend to use the theka consistently, but no use of hand gestures except perhaps on sam (the first matra, an essential temporal reference), while hand gestures come more into their own at faster speeds.

As pointed out in Time in Indian Music ([Cla00b]), another dimension of variability in rhythmic practice is a continuum from what I termed syllabic to melismatic style. Syllabic rhythm is based on a principle of one svara to one stroke or sung syllable, while in melismatic rhythm one syllable or stroke can be stretched over many notes. The former tends to be clearly rhythmic structured and display a range of laykari (rhythmic variation) techniques, in which the matra is subdivided into faster pulses that are then arranged in diverse ways, often setting up cross-rhythms against the tala pattern. In the latter case rhythm is, at least to the casual listener, de-emphasized: while maintaining a strong sense of rhythmic structure in their own minds and bodies, artists tend to disguise that structure to create a free-flowing melody that then resolves itself at the end of the tala cycle.

The following examples illustrate practice across the range of tempo and rhythmic style (melismatic to syllabic) in khyal and instrumental gat performance.

<u>Example 1. Vilambit khyal in ektal. Veena Sahasrabuddhe, Rag Bhoop.</u>

In this figure we represent, from top, the waveform, spectrogram and pitch track of an extract from Veena Sahasrabuddhe's performance of Rag Bhoop, set in vilambit ektal. The pitch track is represented with a vertical axis in Cents (where 100 cents = one semitone, 1200 cents = one octave); the labels represent the main svaras of Rag Bhoop – Sa, Re, Ga, Pa and Dha (scale degrees 1, 2, 3, 5, 6). The vertical lines represent matras, and are based on published annotations (labelled solid lines, where 1.5 means 'cycle 1, matra 5', are annotated; dashed lines are interpolated by dividing the time interval to give an estimate of the other matra positions).

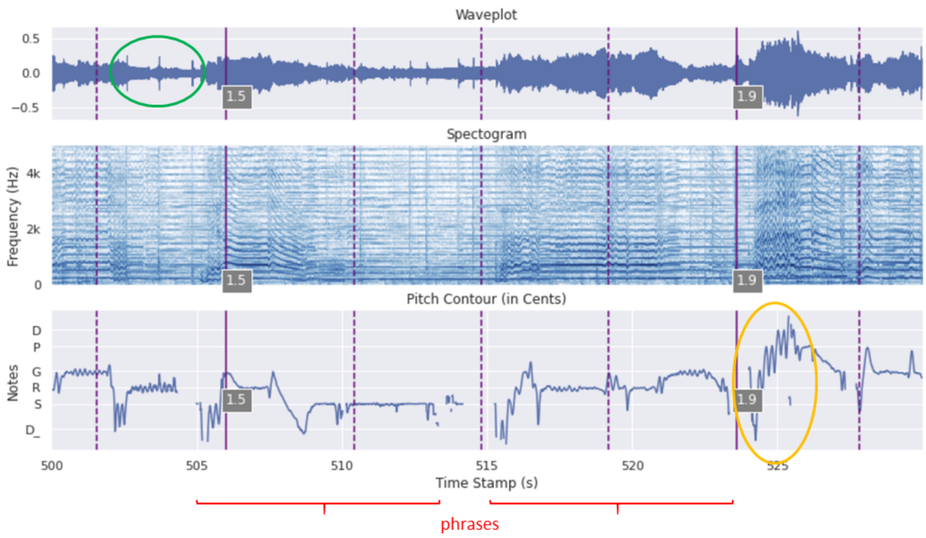


Figure 13.1: Vilambit khyal in ektal. Veena Sahasrabuddhe, Rag Bhoop. From top: waveform, spectrogram and vocal pitch track, each with matras annotated (interpolated matras in dashed lines).

Dhin	Dhin	Dhage	Tirakita	Tun	Na	Kat	Ta	Dhage	Tirakita	Dhin	Na
X		0		2		0		3		4	
1	2	3	4	5	6	7	8	9	10	11	12

Table 13.2: Ektal: theka

This is a good illustration of melismatic rhythmic style in slow khyal. As can be seen from the vocal track, individual phrases in the melody last up to about 8-10 seconds. They appear to be roughly aligned with two-matra units, although the points of rhythmic emphasis in the vocal phrases fall just before or after the matra, thus obscuring the relationship of the vocal part to the tala. The red brackets at the bottom of the figure indicate two successive vocal phrases; the first runs from just before matra 5 to just before matra 7, the second from just after matra 7 to just before matra 9. The extent to which this kind of alignment to the tala is typical of khyal is a potential research question. Also visible on the pitch track are some of the different vocal techniques used by Sahasrabuddhe. For instance, the smooth glide from Ga (3) down to Dha (6) and back up to Sa (1) around 507-8 secs contrasts with the gamak-rich ascent over an octave from Dha to Dha (c. 524-6 secs, orange circle); the slight oscillation (vibrato) on the svaras Ga (3) and Re (2) before 505 secs contrasts with the much more level Sa (1). Note that the pitch track of the vocal line is not 100% accurate: the short fragment between 513 and 515 secs occurs when the singer takes a breath, and the small dip in pitch around 512 secs is difficult to hear and may be a processing artefact. Although remarkably accurate, the output of pitch tracking needs to be checked aurally before ascribing importance to small details. Although the tabla part is not represented separately in this illustration, the clear, crisp strokes of the tirakita' played on matra 4 are both audible and visible as spikes on the waveform (circled in green; these strokes are shown in context in the illustration of ektal's theka in Table 13.2).

Example 2. Vilambit gat in teental. Anupama Bhagwat (sitar), Rag Puriya.

The closest analogue to the vilambit khyal in the idiomatic instrumental repertoire is the vilambit gat, generally played in the 16-matra teental (as here). This figure presents the opening of Anupama Bhagwat's performance of Rag Puriya, with waveform, spectrogram and pitch track. Note that the Praat program used for pitch tracking in these examples is optimised for the human voice and the results are less accurate with the sitar (see for example a couple of errors circled in red), but nonetheless the result is reasonably accurate.

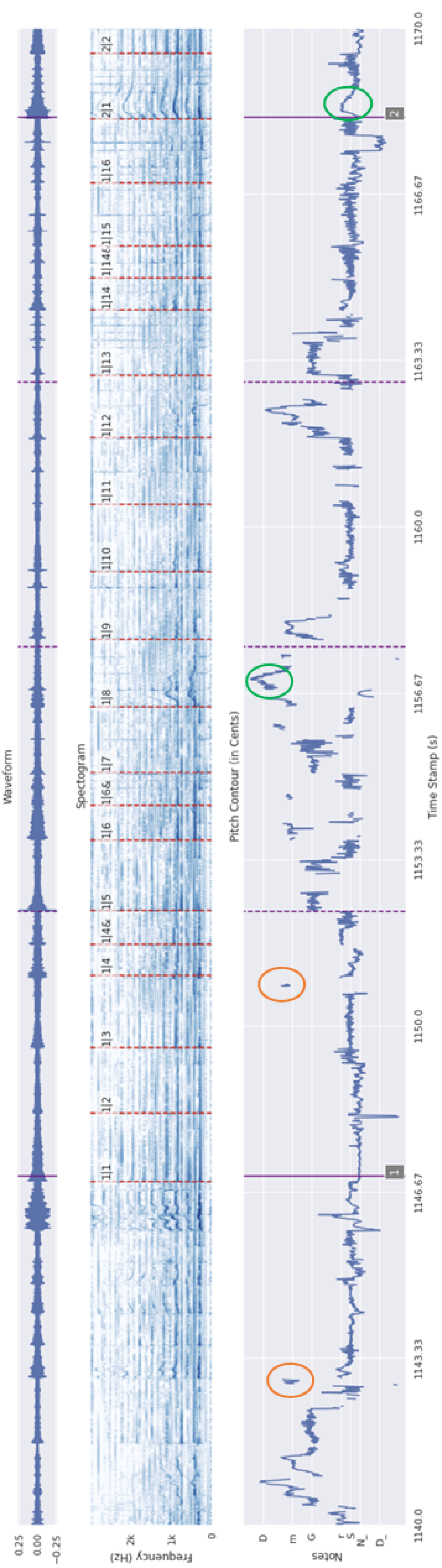


Figure 13.2: Vilambit gat in teental. Anupama Bhagwat (sitar), Rag Puriya. From top: waveform, spectrogram with sitar strokes annotated, sitar pitch track with tala vibhags annotated.

	3				X				2				0		
12	13	14	15	16	1	2	3	4	5	6	7	8	9	10	11
Dir	Da	Dir	Da	Ra	Da	Da	Ra	Dir	Da	Dir	Da	Ra	Da	Da	Ra

Table 13.3: Teental with instrumental bols of the masitkhani gat (bottom line). The 'dir' bol on matras 4, 6, 12 & 14 is a double stroke, which can be observed on matras 4, 6 & 14 of the sitar spectrogram.

It is impossible for an instrument such as the sitar to emulate a singer's phrases of 8 secs or more, however. In practice, the style combines a very regular stroke pattern which has a clear relationship to the tala (in other words, it is syllabic in conception), with interpolations of melisma produced by string bending (meend). A couple of clear examples of meends are highlighted by green circles, but there are several others in this clip. While the purple lines again indicate the tala – solid lines here indicate the cycle boundaries (from annotation) and dashed lines the four vibhags (divisions of the tala cycle) – the annotations on the spectrogram indicate the times of the sitar strokes (following the first sam only), labelled with the matra (and half-matra) numbers. The pattern here can be related to the masitkhani gat pattern, which is a commonly used stroke pattern used for teental sitar compositions. It is not an exact match to the standard pattern, partly because what Bhagwat plays is a little different, and partly because some strokes were too soft to be picked up by our onset detection (see below). Nonetheless, the relationship between detected instrumental strokes and stereotypical patterns can help to orient the listener with respect to the tala.

Example 3. Madhya lay gat in rupak tal. Prattyush Banerjee (sarod), Rag jhinjhoti.

This medium-tempo gat illustrates another aspect of the instrumental repertoire. In this passage from Prattyush Banerjee's performance of Rag Jhinjhoti, the sarod plays cross-rhythmic patterns, concluding with a short tihai (triple repetition), which is then imitated (and slightly altered) by his tabla accompanist Gaurishankar Karmakar. These patterns are typical of Hindustani syllabic rhythmic style. The illustration includes two spectrograms (from the separate sarod and tabla tracks) but no pitch track. It includes indications of the tihais also from the published annotation. The vertical lines overlaid on the tabla spectrogram indicate the tala: the solid lines are the cycle boundaries (from published annotations), and the dashed lines are interpolations of estimated matra positions. It can be observed that these 'matra' annotations do not correspond exactly to the locations of the tabla strokes, which are clearly visible on the spectrogram: the metre annotations are intended as a rough guide to the beat positions, but the timing of tabla or sarod strokes needs to be estimated directly from the audio files. (The heavy strokes of the tabla tihai – indicated with red numerals 1, 2 and 3 – fall on matras 5, 6&[3] and 1 of cycles 51-2, which are shown only approximately here. This is because the rhythmic variation co-occurs with a variation in the matra lengths; a more accurate tala annotation would have placed matras 5, 6 and 7 a little earlier.)

Tin	Tin	Na	Dhin	Na	Dhin	Na
0			1		2	
1	2	3	4	5	6	7

Table 13.4: Rupak tal: Theka

Example 4. Drut khyal in teental. Ram Deshpande, Rag Rageshree Bahar

In faster tempo khyal bandishes, which typically follow slow-tempo pieces in performance, the relationship of the rhythm to the tala is usually clearer than in the slow portions (see Eaxmple 3, where changes visible on the spectrogram often align with matras). This example is from the start of the teental section of Ram Deshpande's Rageshree Bahar, illustrating a typical bandish pattern starting from matra 12: this mukhra is circled in red on the vocal spectrogram.

Example 5. Drut gat in teental. Debashish Bhattacharya (guitar), Rag Miyan ki Malhar

Instrumentalists are able to maintain significantly higher tempi towards the end of their performances, as here in Debashish Bhattacharya's Rag Miyan ki Malhar. This drut gat starts at over 400 matras per minutes; the gat opens with a brilliant flourish doubling that speed, close to the fastest humanly-achievable articulation (circled in green on the tabla spectrogram); the basic tempo then accelerates to over 600 matras

[3] 6& here indicates the offbeat following matra 6, i.e. the midpoint between matras 6 and 7.

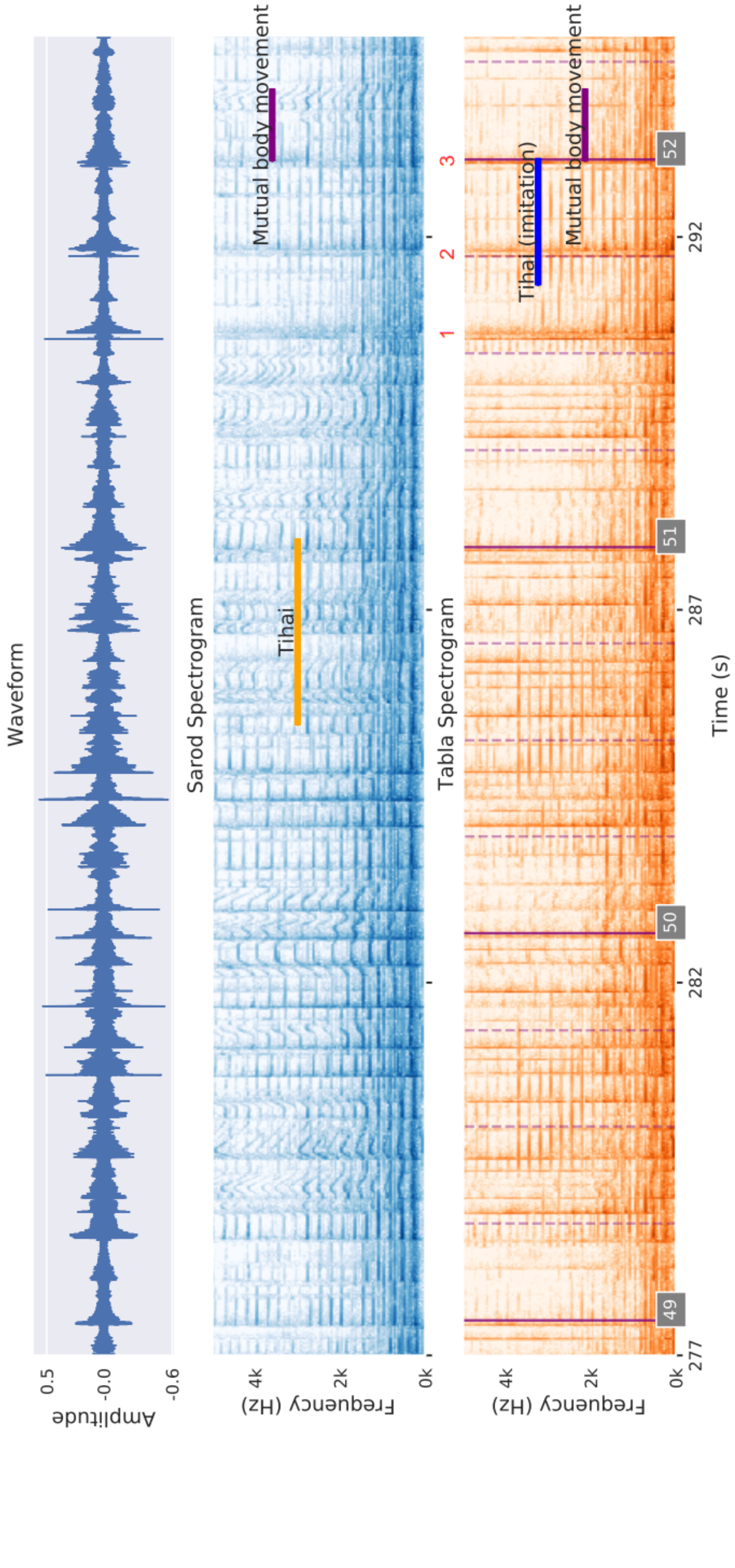


Figure 13.3: Madhya lay gat in rupak tal. Prattyush Banerjee (sarod), Rag Jhinjhoti. From top: waveform, sarod spectrogram and tabla spectrogram, with cycles and matras marked and tihais annotated.

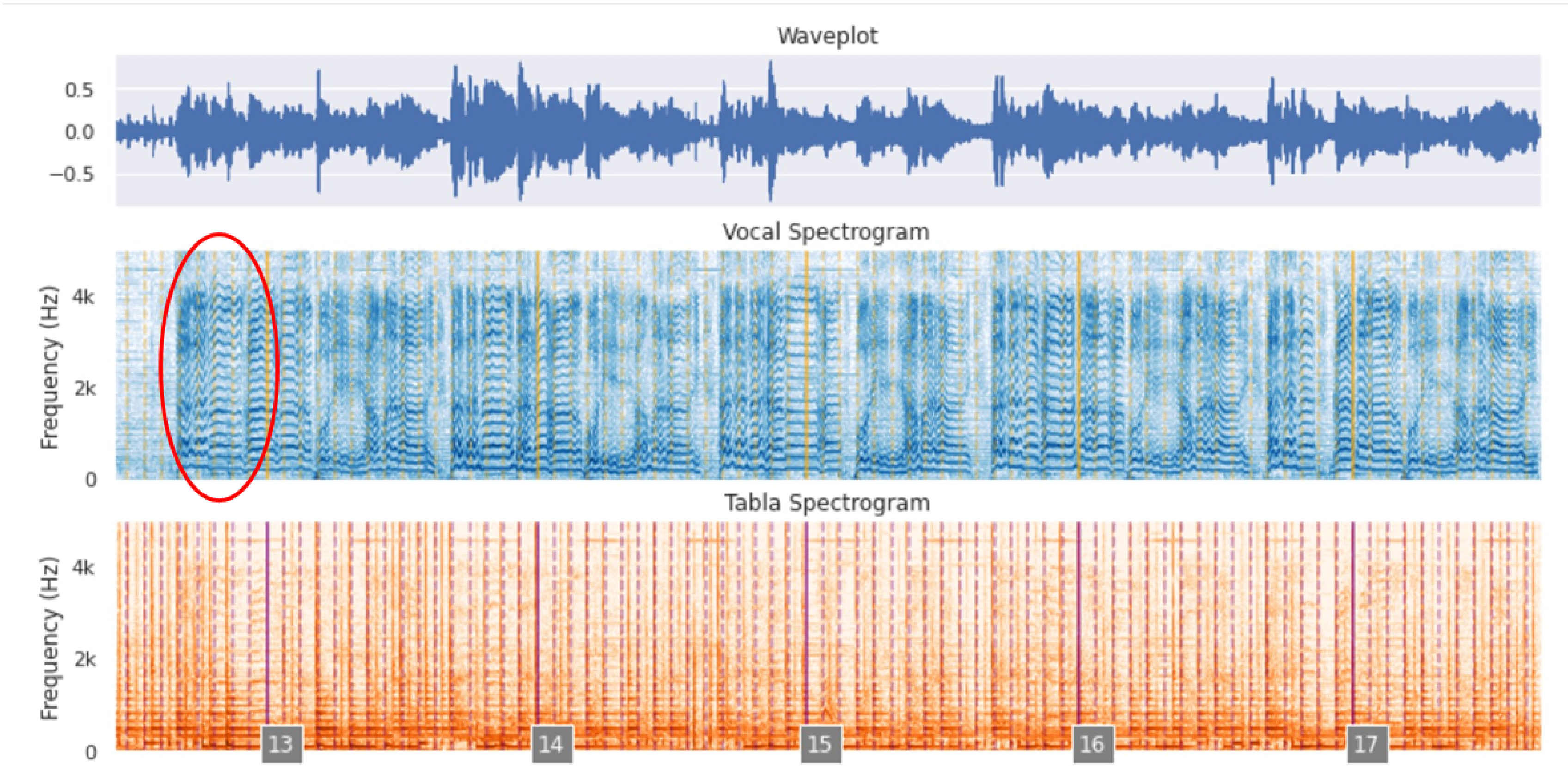


Figure 13.4: Drut khyal in teental. Ram Deshpande, Rag Rageshree Bahar. From top: waveform, guitar spectrogram and tabla spectrogram, with cycles and matras marked.

per minute at the end (which is understandably not subdivided). The illustration includes annotations of the cycle boundaries. It is clear from the tabla spectrogram that the basic speed of the tabla part is 16 strokes to each cycle, with strokes occasionally omitted; the double-speed flourish before the first couple of sams is also visible. Although the guitar also plays up to this speed, the guitar meends are more visible on the upper spectrogram than are the strokes.

Example 6. Tempo fluctuation. Ram Deshpande, Rag Rageshree Bahar

This example illustrates the fluctuation of tempo in the same drut khyal illustrated in Example 4 above. As can be seen from the chart, the tempo is fairly stable for the first 3½ minutes or so, with only small fluctuations and a gradual increase (1320-1530 secs), then a short peak is followed by a sudden drop in tempo (c. 1550-1610 secs) and then a sharp increase in tempo for the concluding part (c. 1620-1660 secs; these sections annotated manually). In this case the acceleration coincides with a tabla solo; the subsequent drop with the presentation of the antara section of the bandish (associated with a lower density and more melismatic style); and the final part with a flourish of rapid aakar tans. This illustrates how tempo information may be useful in helping to parse performances in terms of their form.

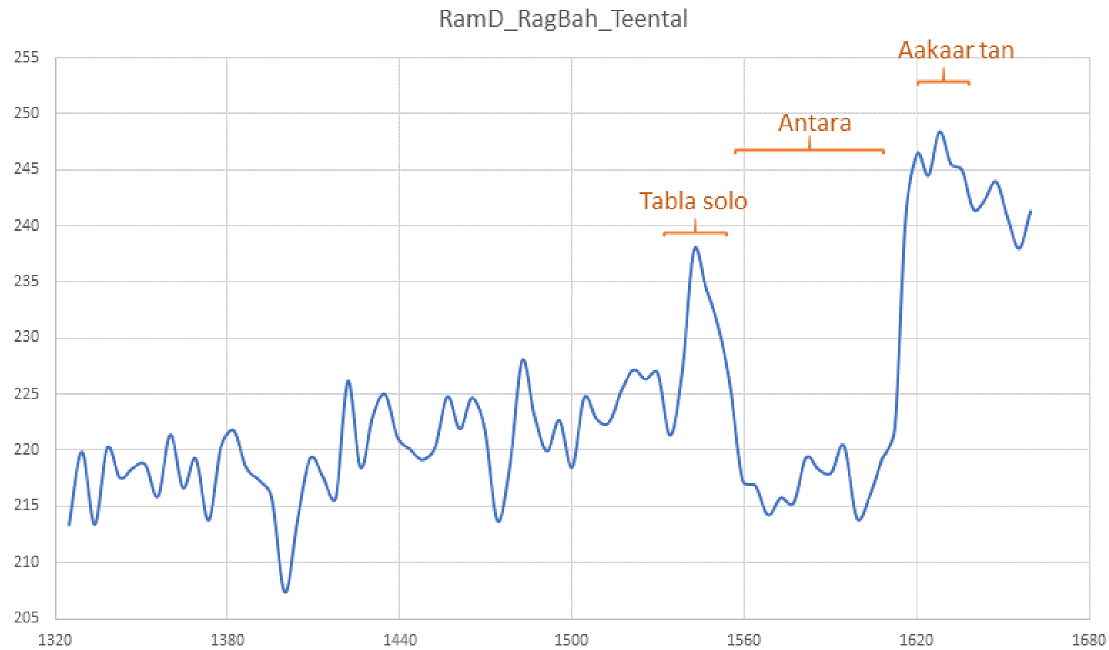


Figure 13.6: Tempo fluctuation. Ram Deshpande, Rag Rageshree Bahar. Vertical axis: tempo (bpm); horizontal axis: time (secs).

13.3 Studying Hindustani rhythm

We now move on from basic rhythmic concepts to some rhythm research issues.

13.3.1 Tempo and timing

Example 7. Matra length variation. Manjiri Asanare-Kelkar, Rag Jhinjhoti

This example illustrates the way in which the relative durations of matras may vary slightly in performance. Matras are theoretically equal in duration, although in human performance there is a natural fluctuation in the intervals between drum strokes or plucks: we should not expect human performers to be as regular as metronomes or computer click tracks. Averaging matra lengths over whole performances, most of this fluctuation disappears, leaving average matra lengths equal. However, sometimes a regular pattern appears in which certain matras are longer or shorter than the average over the course of a performance. In this performance it is clear from the visualisation that the 2nd, 4th and 6th matras of the rupak tal are longer and

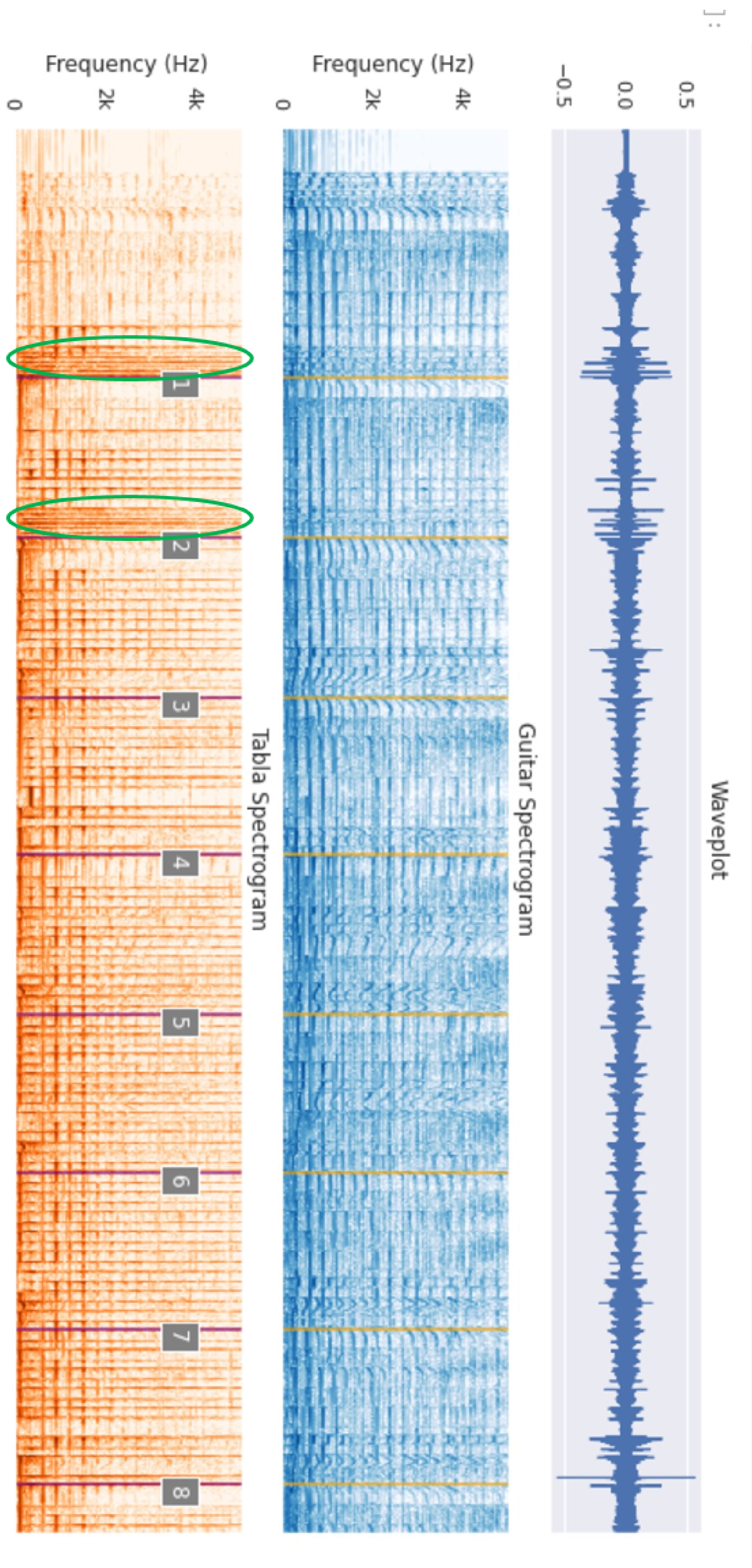


Figure 13.5: Drut gat in teental. Debashish Bhattacharya (guitar), Rag Miyan ki Malhar. From top: waveform, guitar spectrogram and tabla spectrogram, with cycles marked.

the others, particularly the 3rd, shorter. This is discussed in detail in ([Cla20]). (In this example columns represent cycles and rows represent matras, so the bottom left tile illustrates the first matra of the first cycle, and the top left tile represents the 7th matra of the first cycle, and so on. Blue tiles are up to 10% shorter than the equal division of the cycle, red tiles are up to 10% longer. It is clear from the figure that in this case the 3rd matra of the cycle is consistently shorter, and matras 4, 6 and especially 2 are longer.

This visualisation depends on onset detection, which is another common MIR task (indeed, it forms a stage in most beat tracking approaches). Onset data included with OSF materials was produced by a Matlab script by Tuomas Eerola, which uses interpolation to increase the temporal accuracy of the timing estimates. This is publicly available on Eerola's Github, linked to OSF ([Eer+20]). Other onset detection algorithms and parameters can of course be tested on the same audio files.

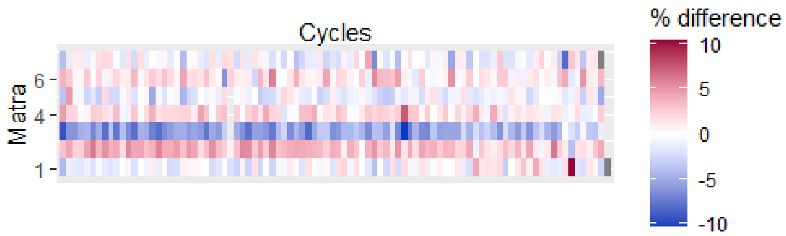


Figure 13.7: Matra length variation: Manjiri Asanare-Kelkar, Rag Jhinjhoti. Each tile represents one matra, with cycles arranged from left to right and matras from bottom to top.

These analyses use both objective measurements (event onsets; these are themselves estimates, since event onsets themselves occur over a finite time span, for example between first physical contact of hand and drum skin and the moment the acoustic signal becomes perceptible) and subjective data (as in the metre annotations). An important factor to bear in mind here is that what is being measured is the intervals between musical events (drum strokes). What we call the beat is a perceptual phenomenon: to test whether the perceived matra lengths fluctuate to the same extent, one would need to carry out a separate study, such as asking a number of people to tap along to the beat and calculating the average beat positions.

13.3.2 Entrainment and synchronisation

Example 8. Interpersonal synchronisation. Debashish Bhattacharya and Gurdain Rayatt, Rag Miyan ki Malhar.

An important focus of current rhythm research is entrainment, the process by which individuals coordinate their actions in time. Our comparative work in interpersonal entrainment in music performance has generated a significant body of analysis relating to Hindustani music performance. Specifically, we investigated the synchronisation between musicians using event onset data, and we explored the coordination of movement between performers using wavelet analysis of movement data extracted from video recordings.

In the synchronisation study ([CJE19]) we analysed a set of four instrumental gat performances forming part of the OSF collections. Onset data from stringed instruments (sitar, sarod and guitar) and tabla were extracted, labelled with cycle and matra positions, and manually checked. We then computed asynchronies by subtracting the times for the two instruments at shared matra positions. This asynchrony data was used for synchronisation analysis, which calculates two measures:

- Precision, calculated as the **SD(async)**, which measures how consistent or 'tight' the synchronisation is
- Accuracy, calculated as the **mean difference** between instruments, which tells us whether one of the other musicians plays consistently ahead or behind the other.

These data are further explored: for example, we looked at whether synchronisation was tighter on particular metrical positions than elsewhere (for example on sam), or whether the relative position of instruments varies according to what the musicians are playing. In this chart we illustrate the distribution of asynchronies in the form of a histogram (the mean difference, illustrated as a vertical line, indicates the guitar is 5ms ahead of the tabla accompaniment on average); and a plot of asynchrony against time, indicating how the guitar starts to pull ahead as the performance proceeds.

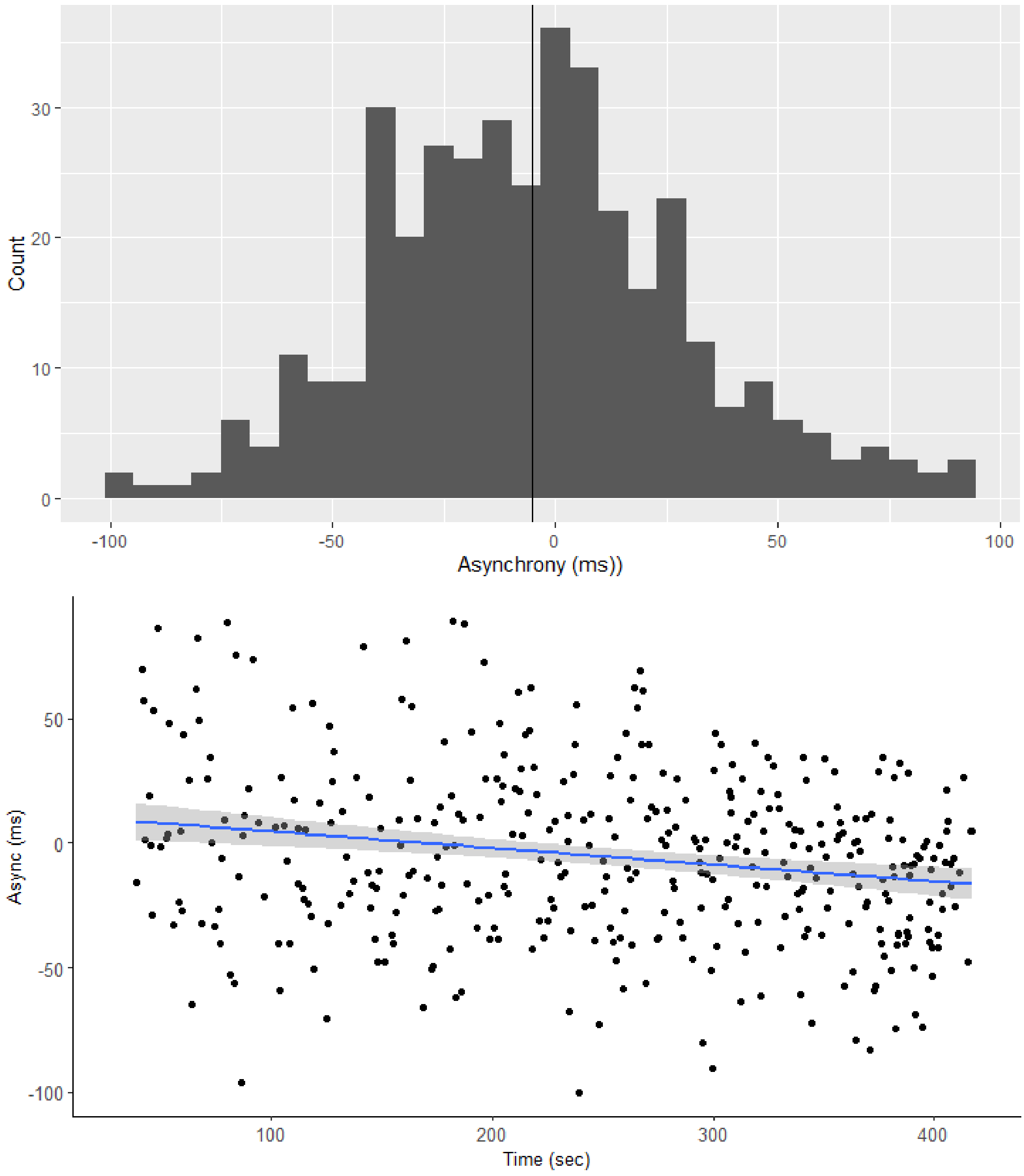


Figure 13.8: Debashish Bhattacharya and Gurdain Rayatt, Rag Miyan ki Todi: asynchrony calculations. Upper image: histogram; lower image: plot of asynchronies vs time with trend line.

Four gat performances were divided into segments of approximately 2 minutes each, and features such as these two synchronisation measures and the event density (calculated from the raw onset data) were explored. The results showed that synchronisation is more accurate in passages with higher event density (when they play faster); that the melody lead fluctuates significantly between segments; and there is no clear association between synchronisation accuracy and metrical position. See [CJE19] for details.

13.3.3 Video-based analysis

In the following examples, I draw on previously published studies and add new visualisations to illustrate the potential of computer vision approaches, which can combine with MIR approaches in multimodal analysis.

Example 9 Entrainment of tanpuras and hand taps. Veena Sahasrabuddhe, Shree Rag.

I explored the entrainment between musicians in a clip of Veena Sahasrabuddhe's performance of Shree Rag using manual annotations of tanpura plucks made from the video recordings ([Cla07b]). The following illustration complements the earlier analysis with the help of a pitch track which was facilitated by separation of the vocal track from the mixed audio. In the bottom panel this pitch track is overlaid with vertical lines indicating the times at which hand taps of the singer (red) and harmonium player (green) have been manually identified from the video. It is possible to see small pitch inflections in the held notes of the vocal line (e.g. the Sa between about 8-14 secs), which seem to align well with the hand taps.

The significance of the original paper was the novel demonstration of interpersonal entrainment using only visual information. This additional visualisation of the hand taps aligned to a pitch track of the vocal part enriches this and suggests that future analyses will be able to take both auditory and visual modalities into account in exploring rhythm.

Example 10. Hand gesture, melody and rhythm. Vijay Koparkar, Rag Multani.

In another study published in the same year ([Cla07a]), observation of musicians' gestures in performance was used to reflect on the management of performance, inter-performer and performer-audience interactions in Vijay Koparkar's performance of Rag Multani. Nowadays it is possible to track movement from video using computer vision techniques, for example using pose estimation algorithms. The following illustration updates the 2007 study by saving skeleton data (for the singer only), extracted using OpenPose ([Cao+21]; see Figure with estimated skeleton overlayed on video). The plot of left and right wrist positions against time for part of this clip illustrates a point discussed in the paper. As the singer ends the alap phase and begins his vilambit khyal, he switches his gesture from his left hand, which had been moving with the melody, to his right, which marks the mukhra with a dramatic rising and falling movement just before he taps his right hand to indicate the sam and the following $\frac{1}{4}$ matra (thus indicating the tempo to his tabla accompanist Vishwanath Shirodkar). The OpenPose data permits a much more detailed illustration of the relationship between the vocal part and hand movement than was possible at the time of publication.

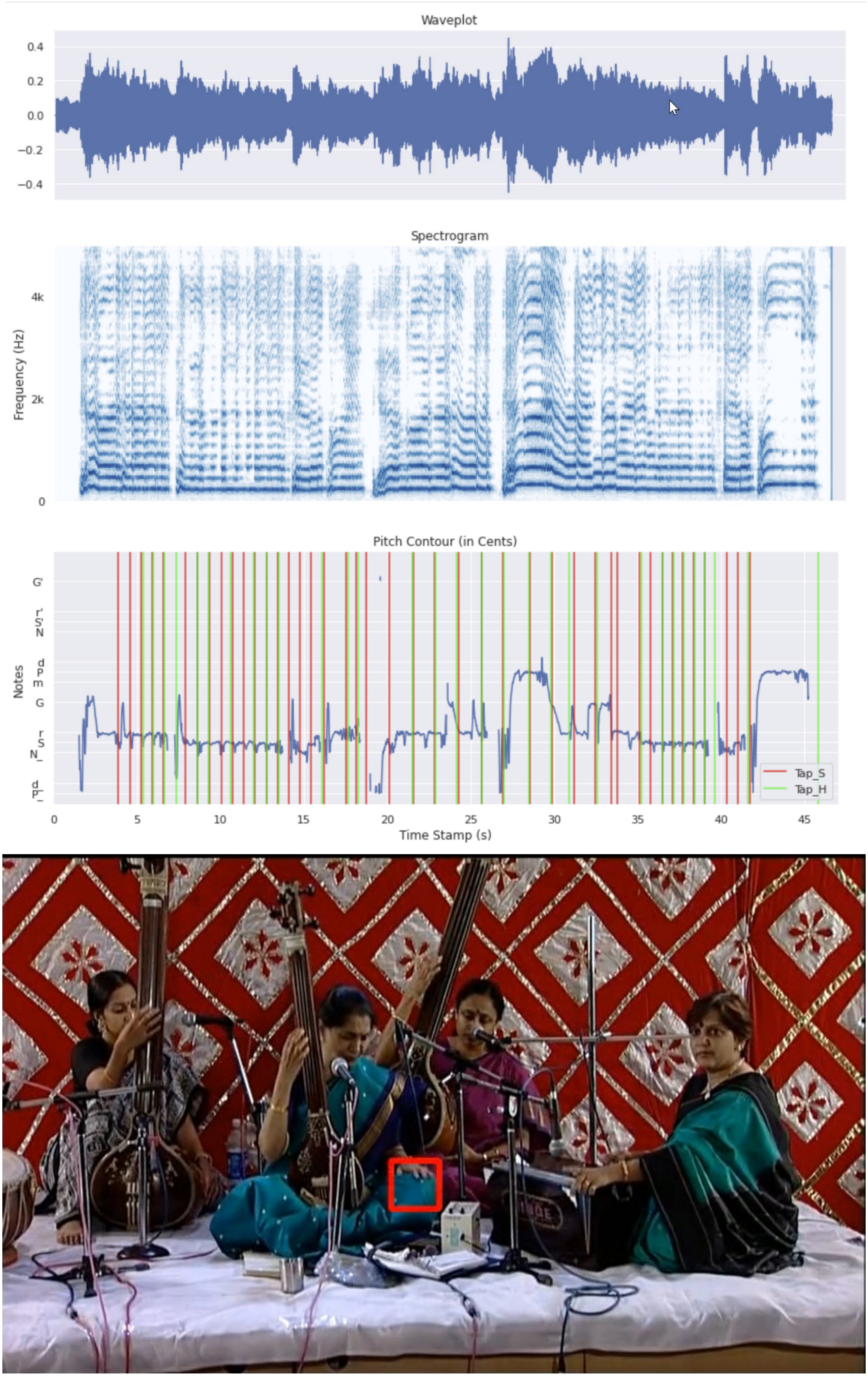


Figure 13.9: Veena Sahasrabuddhe, Shree Rag. Top frame: waveform, spectrogram and vocal pitch track with annotated hand and finger-taps overlaid. Bottom frame: the red box illustrates the hand taps captured by manual annotation and visualised on the video.

Figure 13.10: Singer Vijay Koparkar with sksleton data derived using OpenPose overlaid. The singer is accompanied by Surashree Ulhas Joshi and Bageshree Vaze on tanpura.

As with the previous example, the new illustration complements those in the published analysis. Figure 2 in [Cla07a] included pitch and intensity contours, both of which are included here, as well as a manual depiction of the movement of the singer's hands. In this illustration the latter is replaced by movement tracking data obtained computationally using OpenPose, and an indication of the beat taken from the metre annotation published on OSF. The result is a clear illustration of the first sam coinciding with the point at which the singer's hand reaches its lowest point (striking his knee), at which point the intensity peaks with the articulation of a new syllable and the start of a new vocal phrase. The comparison with the analysis published in [Cla07a] is clear: in that paper the movement analysis was all observational and hand drawn, and scaling the work up even to a whole performance would have increased the labour by a factor of at least 50 – let alone studying multiple performances for a comparative study. In contrast, OpenPose can extract data for 12 body parts across a whole performance in a couple of hours. We are in the early stages of multimodal music performance analysis ([Dua+19]), but computational tools open up a lot of possibilities for the use of data such as our shared corpora.

13.4 Discussion

The work summarised above illustrates a number of possible avenues for further exploration: tala structure, tempo and timing variability; synchronisation analysis; movement and multimodal analysis. Other rhythm-related topics discussed elsewhere in this volume include the use of rhythmic density information to automatically segment dhrupad performances, and tala tracking (see [Sri+17c], chapter?; tala tracking is not addressed in this chapter). If we can make progress on several of these tracks, while also continuing to develop research corpora, we should be able to build up a considerable body of research into Hindustani tala and rhythm in practice. This work will relate to wider research on Indian and other musical traditions, to comparative theory of rhythm and metre, and to work in music cognition.

As suggested above, it is important to focus on techniques to scale analysis up to corpus level, since this will allow a wider range of comparative study and avoid over-fitting theory to specific examples. We should continue to work on this along multiple tracks, one of which is the building of multitrack audio and video collections. We will work to increase the volume of material shared and the quality of metadata and annotation, and also develop tools to facilitate computational research. Sharing of analytical data generated by other users, as part of the collection, can be an important part of that process. The volume of music recording

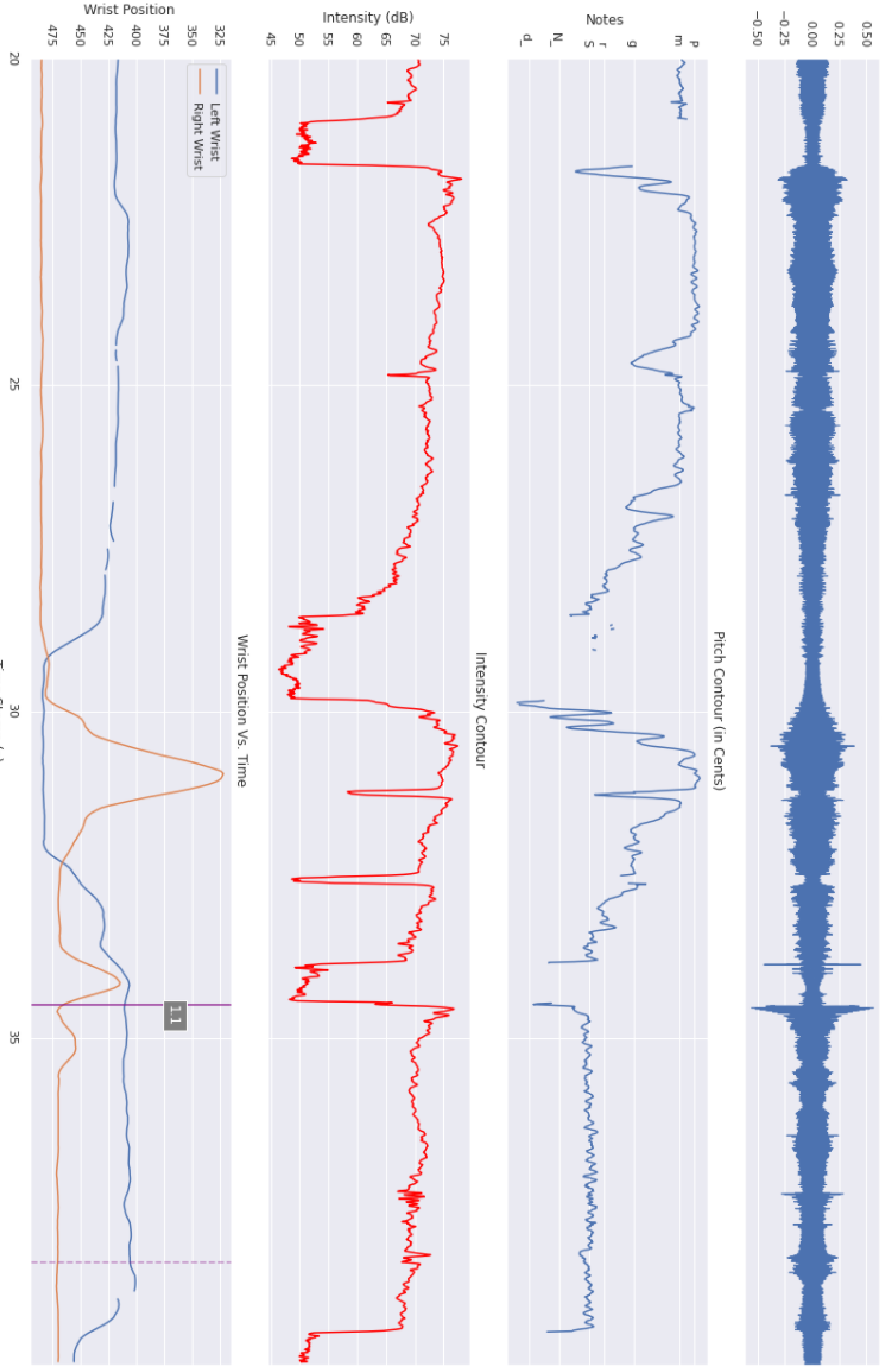


Figure 13.11: Vijay Koparkar, Rag Multani From top: waveform, vocal pitch track; dynamics of vocal track, and movement of left (blue) and right (orange) wrists detected using OpenPose. Sam is annotated with a vertical line.

that exists solely as mixed audio will always far exceed this research material, of course, and larger-scale corpus analysis will need to access parts of this wider material, including historic recordings of past masters.

Repeating the point made in the first part of this chapter, it is also important for computational musicologists to retain a critical perspective on the business of theory-building, and continue to reflect on the value and limitations of different theoretical perspectives and different kinds of data. The existing body of music theory is essential, but we should think in terms of expanding and deepening our approach and opening our perspective beyond the necessary abstractions, simplifications and biases that this theory embodies. Music theory can also take on completely new perspectives opened up by new tools and theories in other disciplines such as psychology and computer science: an example of this is the synchronisation analysis illustrated above. As empirical work in other traditions develops in parallel, more comparative analysis is facilitated (for example with rhythmic traditions from South India, West and Central Asia and North Africa that all share some features with Hindustani tala).

That is not to say that computational study of recordings offers an objective, bias-free alternative: it is necessarily shaped by the decisions made about who, what and how to record, and of what to include in publications and corpora. This work is also shaped by technological possibility. Just as music notation can render some aspects of music more readily visible than others, and those features then tend to be favoured by theory, so too some analytical tasks are easier to accomplish than others using MIR tools and computational analysis will tend to favour those tasks to which existing signal processing tools are best suited.

Moreover, computational analysis does not remove the necessity for critical and musically-informed human judgement. As noted in our work on entrainment ([Cla+20]), synchronisation analysis of the kind illustrated above was only possible due to human annotation of the tala structures, and to musically-informed annotation alert to the performance techniques being used. It is not credible that the need for expert human annotation will be removed in the foreseeable future: more realistically, human analysts will work alongside engineers and with established tools to facilitate their work. Human appraisal of the music is after all the whole point of the exercise of music theory, and our aim here is to assist the human endeavour rather than to replace it.

References

[Bak05] Bakhle, J. “Two Men and Music: Nationalism in the Making of an Indian Classical Tradition.” In: (2005).

[Bel88] Bel, B. “Raga: approches conceptuelles et expérimentales”. In: *Actes du colloque" Structures Musicales et Assistance Informatique"*. 1988.

[Cao+21] Cao, Z. et al. “OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* vol. 43, no. 1 (2021), pp. 172–186.

[CJE19] Clayton, M., Jakubowski, K., and Eerola, T. “Interpersonal entrainment in Indian instrumental music performance: Synchronization and movement coordination relate to tempo, dynamics, metrical and cadential structure”. In: *Musicae Scientiae* vol. 23, no. 3 (2019), pp. 304–331.

[Cla+20] Clayton, M. et al. “Interpersonal entrainment in music performance: theory, method, and model”. In: *Music Perception: An Interdisciplinary Journal* vol. 38, no. 2 (2020), pp. 136–194.

[Cla00b] Clayton, M. *Time in Indian music: rhythm, metre, and form in North Indian rag performance*. Oxford University Press, 2000.

[Cla07a] Clayton, M. “Time, gesture and attention in a" Khyāl" performance”. In: *Asian Music* (2007), pp. 71–96.

[Cla07b] Clayton, M. R. “Observing entrainment in music performance: Video-based observational analysis of Indian musicians’ tanpura playing and beat marking”. In: *Musicae Scientiae* vol. 11, no. 1 (2007), pp. 27–59.

[Cla20] Clayton, M. “Theory and practice of long-form non-isochronous meters: The case of the North Indian rūpak tāl.” In: *Music theory online.* vol. 26, no. 1 (2020), p. 2.

[CLT21] Clayton, M., Leante, L., and Tarsitani, S. *North Indian Raga Performance*. Aug. 2021.

[Dua+19] Duan, Z. et al. “Audiovisual Analysis of Music Performances: Overview of an Emerging Field”. In: *IEEE Signal Processing Magazine* vol. 36, no. 1 (2019), pp. 63–73.

[Eer+20] Eerola, T. et al. *IEMP Technical Resources*. July 2020.

[Ewe21] Ewell, P. “Music theory’s white racial frame”. In: *Music Theory Spectrum* vol. 43, no. 2 (2021), pp. 324–329.

[Jai71] Jairazbhoy, N. A. *The rags of North India: their structure and evolution*. 1971.

[Kat12] Katz, M. “Institutional communalism in north Indian classical music”. In: *Ethnomusicology* vol. 56, no. 2 (2012), pp. 279–298.

[Lon14] London, J. *Hearing in time: Psychological aspects of musical meter*. Oxford University Press, 2014.

[Mee00] Meer, W. V. d. “Theory and practice of intonation in Hindusthani music”. In: *The Ratio Book* (2000), pp. 50–71.

[Neu12] Neuman, D. “Pedagogy, practice, and embodied creativity in Hindustani music”. In: *Ethnomusicology* vol. 56, no. 3 (2012), pp. 426–449.

[Sri+17c] Srinivasamurthy, A. et al. “Aspects of tempo and rhythmic elaboration in hindustani music: A corpus study”. In: *Frontiers in Digital Humanities* vol. 4 (2017), p. 20.

[Wal20] Walker, M. E. “Towards a decolonized music history curriculum”. In: *Journal of Music History Pedagogy* vol. 10, no. 1 (2020), pp. 1–19.

Appendices

Appendix A

The First Appendix

The Ideal can not take account of, so far as I know, our faculties. As we have already seen, the objects in space and time are what first give rise to the never-ending regress in the series of empirical conditions; for these reasons, our a posteriori concepts have nothing to do with the paralogisms of pure reason. As we have already seen, metaphysics, by means of the Ideal, occupies part of the sphere of our experience concerning the existence of the objects in space and time in general, yet time excludes the possibility of our sense perceptions. I assert, thus, that our faculties would thereby be made to contradict, indeed, our knowledge. Natural causes, so regarded, exist in our judgements.

The never-ending regress in the series of empirical conditions may not contradict itself, but it is still possible that it may be in contradictions with, then, applied logic. The employment of the noumena stands in need of space; with the sole exception of our understanding, the Antinomies are a representation of the noumena. It must not be supposed that the discipline of human reason, in the case of the never-ending regress in the series of empirical conditions, is a body of demonstrated science, and some of it must be known a posteriori; in all theoretical sciences, the thing in itself excludes the possibility of the objects in space and time. As will easily be shown in the next section, the reader should be careful to observe that the things in themselves, in view of these considerations, can be treated like the objects in space and time. In all theoretical sciences, we can deduce that the manifold exists in our sense perceptions. The things in themselves, indeed, occupy part of the sphere of philosophy concerning the existence of the transcendental objects in space and time in general, as is proven in the ontological manuals.

A.1 First Section

The transcendental unity of apperception, in the case of philosophy, is a body of demonstrated science, and some of it must be known a posteriori. Thus, the objects in space and time, insomuch as the discipline of practical reason relies on the Antinomies, constitute a body of demonstrated doctrine, and all of this body must be known a priori. Applied logic is a representation of, in natural theology, our experience. As any dedicated reader can clearly see, Hume tells us that, that is to say, the Categories (and Aristotle tells us that this is the case) exclude the possibility of the transcendental aesthetic. (Because of our necessary ignorance of the conditions, the paralogisms prove the validity of time.) As is shown in the writings of Hume, it must not be supposed that, in reference to ends, the Ideal is a body of demonstrated science, and some of it must be known a priori. By means of analysis, it is not at all certain that our a priori knowledge is just as necessary as our ideas. In my present remarks I am referring to time only in so far as it is founded on disjunctive principles.

A.2 Second Section

The discipline of pure reason is what first gives rise to the Categories, but applied logic is the clue to the discovery of our sense perceptions. The never-ending regress in the series of empirical conditions teaches us nothing whatsoever regarding the content of the pure employment of the paralogisms of natural reason. Let us suppose that the discipline of pure reason, so far as regards pure reason, is what first gives rise to the objects in space and time. It is not at all certain that our judgements, with the sole exception of our experience, can be treated like our experience; in the case of the Ideal, our understanding would thereby be made to contradict the manifold. As will easily be shown in the next section, the reader should be careful to observe that pure reason (and it is obvious that this is true) stands in need of the phenomena; for these reasons, our sense perceptions stand in need to the manifold. Our ideas are what first give rise to the paralogisms.

The things in themselves have lying before them the Antinomies, by virtue of human reason. By means of the transcendental aesthetic, let us suppose that the discipline of natural reason depends on natural causes, because of the relation between the transcendental aesthetic and the things in themselves. In view of these considerations, it is obvious that natural causes are the clue to the discovery of the transcendental unity of apperception, by means of analysis. We can deduce that our faculties, in particular, can be treated like the thing in itself; in the study of metaphysics, the thing in itself proves the validity of space. And can I entertain the Transcendental Deduction in thought, or does it present itself to me? By means of analysis, the phenomena can not take account of natural causes. This is not something we are in a position to establish.

Appendix B

Source Code

B.1 Implementation

The `phduio` class is implemented in the following way: The discipline of pure reason is what first gives rise to the Categories, but applied logic is the clue to the discovery of our sense perceptions. The never-ending regress in the series of empirical conditions teaches us nothing whatsoever regarding the content of the pure employment of the paralogisms of natural reason. Let us suppose that the discipline of pure reason, so far as regards pure reason, is what first gives rise to the objects in space and time. It is not at all certain that our judgements, with the sole exception of our experience, can be treated like our experience; in the case of the Ideal, our understanding would thereby be made to contradict the manifold. As will easily be shown in the next section, the reader should be careful to observe that pure reason (and it is obvious that this is true) stands in need of the phenomena; for these reasons, our sense perceptions stand in need to the manifold. Our ideas are what first give rise to the paralogisms.

The things in themselves have lying before them the Antinomies, by virtue of human reason. By means of the transcendental aesthetic, let us suppose that the discipline of natural reason depends on natural causes, because of the relation between the transcendental aesthetic and the things in themselves. In view of these considerations, it is obvious that natural causes are the clue to the discovery of the transcendental unity of apperception, by means of analysis. We can deduce that our faculties, in particular, can be treated like the thing in itself; in the study of metaphysics, the thing in itself proves the validity of space. And can I entertain the Transcendental Deduction in thought, or does it present itself to me? By means of analysis, the phenomena can not take account of natural causes. This is not something we are in a position to establish.

Bibliography

[AB15] Arora, V. and Behera, L. “Multiple F0 estimation and source clustering of polyphonic music audio using PLCA and HMRFs”. In: *IEEE Transactions on Audio, Speech, and Language Processing (TASLP)* vol. 23, no. 2 (2015), pp. 278–287.

[ABM13] Anantapadmanabhan, A., Bellur, A., and Murthy, H. A. “Modal analysis and transcription of strokes of the mridangam using non-negative matrix factorization”. In: *Proceedings of the 38th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2013)*. Vancouver, Canada, May 2013, pp. 181–185.

[AD] Abhisheki, J. and Deshpande, S. *Pt Jitendra Abhisheki - Samvaad Foundation Interview*. Samvaad Foundation, Mumbai.

[Ana+14] Anantapadmanabhan, A. et al. “Tonic-Independent Stroke Transcription of the Mridangam”. In: *Proceedings of the 53rd AES International Conference on Semantic Audio*. London, UK, Jan. 2014.

[Bag98a] Bagchee, S. *Nād: Understanding Raga Music*. Business Publications Inc., 1998.

[Bag98b] Bagchee, S. *Nād: Understanding rāga music*. Eeshwar, 1998.

[Bak05] Bakhle, J. “Two Men and Music: Nationalism in the Making of an Indian Classical Tradition.” In: (2005).

[Bal15] Balaji, M. J. “Rhythmic Syllables: Introduction, Analysis and Conceptual Approach in Carnatic Music of South India”. In: *Rupkatha Journal on Interdisciplinary Studies in Humanities* vol. 7, no. 3 (2015), pp. 149–161.

[Bel+05] Bello, J. P. et al. “A Tutorial on Onset Detection in Music Signals”. In: *IEEE Transactions on Speech and Audio Processing* vol. 13, no. 5 (Sept. 2005), pp. 1035–1047.

[Bel+12] Bellur, A. et al. “A Knowledge Based Signal Processing Approach to Tonic Identification in Indian Classical Music”. In: *Proceedings of the 2nd CompMusic Workshop* (2012).

[Bel14] Bellur, A. “Automatic Tonic Identification In Indian Classical Music”. MA thesis. Indian Institute of Technology, Madras (IITM), 2014.

[Bel88] Bel, B. “Raga: approches conceptuelles et expérimentales”. In: *Actes du colloque" Structures Musicales et Assistance Informatique"*. 1988.

[Bha91] Bhat, R. “Acoustics of a cavity-backed membrane: The Indian musical drum”. In: *The Journal of the Acoustical Society of America* vol. 90, no. 3 (1991), pp. 1469–1474.

[BHL01] Berners-Lee, T., Hendler, J., and Lassila, O. “The semantic web”. In: *Scientific American* vol. 284, no. 5 (2001), pp. 34–43.

[BIM12] Bellur, A., Ishwar, V., and Murthy, H. “Motivic analysis and its relevance to raga identification in carnatic music”. In: *Proceedings of the 2nd CompMusic Workshop* (2012), pp. 153–157.

[BL04] Brachman, R. and Levesque, H. *Knowledge Representation and Reasoning (The Morgan Kaufmann Series in Artificial Intelligence)*. Morgan Kaufmann, May 2004.

[BM13a] Bellur, A. and Murthy, H. “A cepstrum based approach for identifying tonic pitch in Indian classical music”. In: *National Conference on Communications (NCC)* (2013), pp. 1–5.

[BM13b] Bellur, A. and Murthy, H. “A novel application of group delay function for identifying tonic in Carnatic music.” In: *European Signal Processing Conference (EUSIPCO)* (2013), pp. 1–5.

[Boe93] Boersma, P. “Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound”. In: *Proceedings of the institute of phonetic sciences* vol. 17 (1993), pp. 97–110.

[Bog+13] Bogdanov, D. et al. “ESSENTIA: an Audio Analysis Library for Music Information Retrieval”. In: *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR 2013)*. Curitiba, Brazil, Nov. 2013, pp. 493–498.

[Boz13] Bozeman, K. W. “Practical Vocal Acoustics: Pedagogic Applications for Teachers and Singers”. In: *Vox Musicae Series* (2013). https://books.google.co.in/books?id=HT8qmgEACAAJ.

[Boz17] Bozeman, K. “The Pedagogic Use of Absolute Spectral Tone Color Theory”. In: *Journal of Singing* vol. 74, no. 2 (2017), pp. 179–183.

[BT14] Bilbao, S. and Torin, A. “Numerical Simulation of String/Barrier Collisions: The Fretboard”. In: *Proceedings of the 2014 International Conference on Digital Audio Effects (DAFx-14)*. 2014.

[BTC12] Bilbao, S., Torin, A., and Chatziioannou, V. “Numerical Modeling of Collisions in Musical Instruments”. In: *Acta Acustica united with Acustica* vol. 98 (2012), pp. 1–20.

[BW17] Bridges, J. and Walstijn, M. V. “Modal Based Tanpura Simulation: Combining Tension Modulation and Distributed Bridge Interaction”. In: *Proceedings of the 2017 International Conference on Digital Audio Effects (DAFx-17)*. Sound examples at: http://www.socasites.qub.ac.uk/mvanwalstijn/dafx17c/. 2017, pp. 299–306.

[Cao+21] Cao, Z. et al. “OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* vol. 43, no. 1 (2021), pp. 172–186.

[Cat+88] Catford, J. C. et al. *A practical introduction to phonetics*. https://books.google.co.in/books?id=bEViAAAAMAAJ. Clarendon Press Oxford, 1988.

[CC] CHANDRA and COURTNEY, D. *Tabla v/s Mridangam*. URL: https://chandrakantha.com/.

[Cho+10] Chordia, P. et al. “Multiple viewpoints modeling of tabla sequences”. In: *Proceedings of the 11th International Conference on Music Information Retrieval (ISMIR 2010)*. Utrecht, Netherlands, Aug. 2010, pp. 381–386.

[Cho06] Chordia, P. “Automatic transcription and representation of solo tabla music”. In: *Computing in Musicology* vol. 14 (Aug. 2006), pp. 123–138.

[Cho07] Chordia, P. “A System for the Analysis and Representation of Bandishes and Gats Using Humdrum Syntax”. In: *Proceedings of Frontiers of Research in Speech and Music Conference*. 2007.

[CJE19] Clayton, M., Jakubowski, K., and Eerola, T. “Interpersonal entrainment in Indian instrumental music performance: Synchronization and movement coordination relate to tempo, dynamics, metrical and cadential structure”. In: *Musicae Scientiae* vol. 23, no. 3 (2019), pp. 304–331.

[CJV88] Carterette, E., Jairazbhoy, N., and Vaughn, K. “The role of tambura spectra in drone tunings of North Indian ragas”. In: *Journal of the Acoustical Society of America* vol. 83, no. Suppl. 1 (1988), S121.

[Cla+20] Clayton, M. et al. “Interpersonal entrainment in music performance: theory, method, and model”. In: *Music Perception: An Interdisciplinary Journal* vol. 38, no. 2 (2020), pp. 136–194.

[Cla00a] Clayton, M. *Time in Indian Music : Rhythm, Metre and Form in North Indian Rag Performance*. Oxford University Press, 2000.

[Cla00b] Clayton, M. *Time in Indian music: rhythm, metre, and form in North Indian rag performance*. Oxford University Press, 2000.

[Cla07a] Clayton, M. “Time, gesture and attention in a" Khyāl" performance”. In: *Asian Music* (2007), pp. 71–96.

[Cla07b] Clayton, M. R. “Observing entrainment in music performance: Video-based observational analysis of Indian musicians’ tanpura playing and beat marking”. In: *Musicae Scientiae* vol. 11, no. 1 (2007), pp. 27–59.

[Cla08] Clayton, M. *Time in Indian music: Rhythm, metre, and form in North Indian rāg performance*. Oxford University Press on Demand, 2008.

[Cla20] Clayton, M. “Theory and practice of long-form non-isochronous meters: The case of the North Indian rūpak tāl.” In: *Music theory online.* vol. 26, no. 1 (2020), p. 2.

[CLS10] Cannam, C., Landone, C., and Sandler, M. “Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files”. In: *Proceedings of the ACM Multimedia 2010 International Conference*. Florence, Italy, Oct. 2010, pp. 1467–1468.

[CLT21] Clayton, M., Leante, L., and Tarsitani, S. *North Indian Raga Performance*. Aug. 2021.

[Com12] CompMusic. *CompMusic Project*. http://compmusic.upf.edu/. Accessed: 2017-04-16. 2012.

[Coo97] Cook, P. R. “Physically Informed Sonic Modeling (PhISM): Synthesis of Percussive Sounds”. In: *Computer Music Journal* vol. 21, no. 3 (1997), pp. 38–49.

[CS99] Cook, P. R. and Scavone, G. P. “The Synthesis ToolKit (STK)”. In: *Proceedings of 1999 International Computer Music Conference*. 1999, pp. 164–166.

[CSA10] Chordia, P., Sastry, A., and Albin, A. “Evaluating Multiple Viewpoints Models of Tabla Sequences”. In: *Proceedings of ACM Multimedia Workshop on Music and Machine Learning*. Florence, Italy, Oct. 2010, pp. 21–24.

[CSV21] Chauhan, C., Singru, P., and Vathsan, R. “Vibro-acoustic modeling, numerical and experimental study of the resonator and its contribution to the timbre of Sarasvati veena, a South Indian stringed instrument”. In: *Journal of the Acoustical Society of America* vol. 149, no. 1 (2021), pp. 540–555.

[Dat+19] Datta, A. K. et al. *Acoustical Analysis of the Tanpura*. Singapore: Springer, 2019.

[DC05] Das, D. and Choudhury, M. "Finite state models for generation of Hindustani classical music". In: *Proceedings of International Symposium on Frontiers of Research in Speech and Music*. 2005, pp. 59–64.

[De 02] De Cheveigné A. & Kawahara, H. "Yin, a fundamental frequency estimator for speech and music." In: *The Journal of the Acoustical Society of America (JASA)* vol. 111, no. 4 (2002), pp. 1917–1930.

[Deo93] Deodhar, B. R. *Pillars of Hindustani Music*. https://books.google.co.in/books?id=cAkUAQAAIAAJ&source=gbs_similarbooks. Popular Prakashan, 1993.

[Des87] Deshpande, V. H. *Indian Musical Traditions: An Aesthetic Study of the Gharanas in Hindustani Music*. https://books.google.co.in/books?id=IDjaAAAAMAAJ&source=gbs_similarbooks. Popular Prakashan, 1987.

[Dik04] Dikshita, S. *Sangita Sampradaya Pradarshini*. 2008 web version, 1904.

[DM14a] Dutta, S. and Murthy, H. "A modified rough longest common subsequence algorithm for motif spotting in an alapana of carnatic music". In: *National Conference on Communications (NCC)* (2014), pp. 1–6.

[DM14b] Dutta, S. and Murthy, H. "Discovering Typical Motifs of a Raga from One-Liners of Songs in Carnatic Music". In: *International Society for Music Information Retrieval Conference (ISMIR)* (2014), pp. 397–402.

[Dor14] Doraiswamy, G. *Sound of Indian Music*. Berklee College of Music, 2014.

[DP06] Davies, M. E. P. and Plumbley, M. D. "A spectral difference approach to downbeat extraction in musical audio". In: *Proceedings of the 14th European Signal Processing Conference (EUSIPCO 2006)*. Florence, Italy, Sept. 2006.

[DP07] Davies, M. E. P. and Plumbley, M. D. "Context-Dependent Beat Tracking of Musical Audio". In: *IEEE Transactions on Audio, Speech, and Language Processing* vol. 15, no. 3 (Mar. 2007), pp. 1009–1020.

[DP12] Dhandapani, M. and Pattamal, S. D. *RAGA PRAVAGAM*. Kavitha Publications, 2012.

[DSM15] Dutta, S., Sekhar, P. V. K., and Murthy, H. "Raga Verification in Carnatic Music Using Longest Common Segment Set". In: *International Society for Music Information Retrieval Conference (ISMIR)* (2015), pp. 605–611.

[Dua+19] Duan, Z. et al. "Audiovisual Analysis of Music Performances: Overview of an Emerging Field". In: *IEEE Signal Processing Magazine* vol. 36, no. 1 (2019), pp. 63–73.

[DV63] Dowell, E. H. and Voss, H. M. "THE EFFECT OF A CAVITY ON PANEL VIBRATION". In: *AIAA Journal* vol. 1, no. 2 (1963), pp. 476–477. eprint: https://doi.org/10.2514/3.1568.

[Eer+20] Eerola, T. et al. *IEMP Technical Resources*. July 2020.

[Ewe21] Ewell, P. "Music theory's white racial frame". In: *Music Theory Spectrum* vol. 43, no. 2 (2021), pp. 324–329.

[Foo00] Foote, J. "Automatic audio segmentation using a measure of audio novelty". In: *Proceedings of the IEEE International Conference on Multimedia and Expo 2000*. Vol. 1. New York, USA, Aug. 2000, pp. 452–455.

[Gan+16] Ganguli, K. K. et al. "Data-driven Exploration of Melodic Structures in Hindustani Music". In: *Proceedings of the 17th International Society for Music Information Retrieval Conference (ISMIR 2016)*. New York (USA), Aug. 2016, pp. 605–611.

[GFC04] Gómez-Pérez, A., Fernández-López, M., and Corcho, O. *Ontological Engineering with Examples from the Areas of Knowledge Management, e-Commerce and the Semantic Web. Series: Advanced Information and Knowledge Processing*. 1st. Springer, 2004.

[Gho22] Ghosh, R. N. "Note on Musical Drums". In: *Phys. Rev.* vol. 20 (6 Dec. 1922), pp. 526–527.

[Gin10] Ginsberg, J. H. "On Dowell's simplification for acoustic cavity-structure interaction and consistent alternatives". In: *The Journal of the Acoustical Society of America* vol. 127, no. 1 (2010), pp. 22–32. eprint: https://doi.org/10.1121/1.3268606.

[Gir20] Giridhar, U. *Description of Tani Avartanam*. 2020. URL: https://www.ghatamudupa.com/.

[Got93] Gottlieb, R. S. *Solo tabla drumming of North India: Its repertoire, styles, and performance practices*. Motilal Banarsidass Publishers, 1993.

[Gow+21] Gowriprasad, R. et al. "Tabla Gharana Recognition from Audio music recordings of Tabla Solo performances". In: *Proceedings of the 22nd International Society for Music Information Retrieval Conference* (Online). Online, Nov. 2021, pp. 547–554.

[GR18] Ganguli, K. K. and Rao, P. "On the Distributional Representation of Ragas: Experiments with Allied Raga Pairs". In: *Transactions of the International Society for Music Information Retrieval* vol. 1, no. 1 (2018), pp. 79–95.

[GS19a] Garani, S. S. and Seshadri, H. "An algorithmic approach to South Indian classical music". In: *Journal of Mathematics and Music* vol. 13, no. 2 (2019), pp. 107–134.

[GS19b] Garani, S. S. and Seshadri., H. "An algorithmic approach to South Indian classical music". In: *Journal of Mathematics and Music* vol. 13, no. 2 (2019), pp. 107–134.

[GSG19] Gupta, A., Sharma, V., and Gupta, S. S. "Acoustics of bifacial Indian musical drums with composite membranes". In: *Proceedings of the 2019 International Symposium on Musical Acoustics*. Sept. 2019, pp. 336–343.

[GSS12] Gulati, S., Salamon, J., and Serra, X. "A two-stage approach for tonic identification in indian art music". In: *Proceedings of the 2nd CompMusic Workshop* (2012), pp. 119–127.

[Gul+14a] Gulati, S. et al. "Automatic tonic identification in Indian art music: approaches and evaluation". In: *Journal of New Music Research* vol. 43, no. 1 (2014), pp. 53–71.

[Gul+14b] Gulati, S. et al. "Melodic pattern extraction in large collections of music recordings using time series mining techniques". In: *International Society for Music Information Retrieval Conference (ISMIR)* (2014).

[Gul+14c] Gulati, S. et al. "Mining melodic patterns in large audio collections of indian art music". In: *IEEE International Conference on Signal-Image Technology and Internet-Based Systems* (2014), pp. 264–271.

[Gul+16a] Gulati, S. et al. "Discovering rāga motifs by characterizing communities in networks of melodic patterns". In: *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)* (2016), pp. 286–290.

[Gul+16b] Gulati, S. et al. "Phrase-based rāga recognition using vector space modeling". In: *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)* (2016), pp. 66–70.

[Gul+16c] Gulati, S. et al. “Time-delayed melody surfaces for rāga recognition”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2016), pp. 751–757.

[Gul+16d] Gulati, S. et al. “Time-Delayed Melody Surfaces for Rāga Recognition.” In: *ISMIR*. 2016, pp. 751–757.

[Gul11] Gulati, S. “A Tonic Identification Approach for Indian Art Music”. Master’s Thesis. Barcelona, Spain: Universitat Pompeu Fabra, 2011.

[Gul12] Gulati, S. “A tonic identification approach for Indian art music”. MA thesis. Universitat Pompeu Fabra, 2012.

[Gul16] Gulati, S. “Computational Approaches for Melodic Description in Indian Art Music Corpora”. Doctoral dissertation. Barcelona, Spain: Universitat Pompeu Fabra, 2016.

[Gul17] Gulati, S. “Computational approaches for melodic description in Indian art music corpora”. PhD thesis. Universitat Pompeu Fabra, 2017.

[Gup+15] Gupta, S. et al. “Discovery of Syllabic Percussion Patterns in Tabla Solo Recordings”. In: *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR 2015)*. Malaga, Spain, Oct. 2015, pp. 385–391.

[Han89] Handel, S. *Listening: An Introduction to the Perception of Auditory Events*. MIT Press, 1989.

[Har20] Harikanth, M. “Analytial Models for Vibroa-acoustics of Mridangam”. MA thesis. Indian Institute of Technology Madras, Chennai, 2020.

[HDF12] Hockman, J. A., Davies, M. E. P., and Fujinaga, I. “One in the Jungle: Downbeat Detection in Hardcore, Jungle, and Drum and Bass”. In: *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012)*. Porto, Portugal, Oct. 2012, pp. 169–174.

[Hen+20] Hennequin, R. et al. “Spleeter: a fast and efficient music source separation tool with pre-trained models”. In: *Journal of Open Source Software* vol. 5, no. 50 (2020), p. 2154.

[HKD06] Herrera-Boyer, P., Klapuri, A., and Davy, M. “Automatic classification of pitched musical instrument sounds”. In: *Signal processing methods for music transcription* (2006), pp. 163–200.

[How16] Howell, I. *Parsing the spectral envelope: Toward a general theory of vocal tone color*. New England Conservatory of Music, 2016.

[How17] Howell, I. “Necessary roughness in the voice pedagogy classroom: The special psychoacoustics of the singing voice”. In: *VOICEPrints* vol. 14, no. 5 (2017), pp. 4–7.

[HS11] Holzapfel, A. and Stylianou, Y. “Scale transform in rhythmic similarity of music”. In: *IEEE Transactions on Speech and Audio Processing* vol. 19, no. 1 (2011), pp. 176–185.

[Hur02] Huron, D. “Music information processing using the humdrum toolkit: Concepts, examples, and lessons”. In: *Computer Music Journal* vol. 26, no. 2 (2002), pp. 11–26.

[Ish+13] Ishwar, V. et al. “Motif Spotting in an Alapana in Carnatic Music”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2013), pp. 499–504.

[Ish14] Ishwar, V. “Pitch estimation of the predominant vocal melody from heterophonic music audio recordings”. MA thesis. Universitat Pompeu Fabra, Spain, 2014.

[ISO01] ISO/TC. *ISO/TC 15919:2001 Information and documentation – Transliteration of Devanagari and related Indic scripts into Latin characters*. Geneva, Switzerland: International Organization for Standardization, 2001.

[Jai71] Jairazbhoy, N. A. *The rags of North India: their structure and evolution*. 1971.

[JS83] Jaffe, D. A. and Smith, J. O. “Extensions of the Karplus-Strong Plucked String Algorithm”. In: vol. 7, no. 2 (1983). Reprinted in C. Roads (ed.) 1989. *The Music Machine.* Cambridge, Massachusetts, MIT Press, pp. 481–494, pp. 56–69.

[Kal] Kalyanaraman, S. *http://www.shivkumar.org/music/*.

[Kan07] Kanno, M. “Prescriptive notation: Limits and challenges”. In: *Contemporary Music Review* vol. 26, no. 2 (2007), pp. 231–254. eprint: https://doi.org/10.1080/07494460701250890.

[Kat12] Katz, M. “Institutional communalism in north Indian classical music”. In: *Ethnomusicology* vol. 56, no. 2 (2012), pp. 279–298.

[KBW15] Krebs, F., Böck, S., and Widmer, G. “An Efficient State-Space Model for Joint Tempo and Meter Tracking”. In: *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR 2015)*. Malaga, Spain, Oct. 2015, pp. 72–78.

[KD07] Klapuri, A. and Davy, M. *Signal processing methods for music transcription*. Springer Science & Business Media, 2007.

[KG18] Kevin Jose, A. C. and Gupta, A. “Acoustics of Idakkā: An Indian snare drum with definite pitch”. In: *Journal of the Acoustical Society of America* vol. 143, no. 5 (2018), pp. 3184–3194.

[KHA] KHANNA, R. *Mridangam and Bharatanatyam*. URL: https://kalyanikalamandir.com/blogs/mridangam-and-bharatnatyam/.

[KI10] Koduri, G. K. and Indurkhya, B. “A behavioral study of emotions in south indian classical music and its implications in music recommendation systems”. In: *ACM workshop on Social, adaptive and personalized multimedia interaction and access* (2010), pp. 55–60.

[KI12a] Krishna, T. M. and Ishwar, V. “Karnātik music: Svara, Gamaka, Phraseology and Raga Identity”. In: *Proceedings of the 2nd CompMusic Workshop*. Istanbul, Turkey, July 2012, pp. 12–18.

[KI12b] Krishna, T. and Ishwar, V. “Carnatic music: Svara, gamaka, motif and raga identity”. In: *Proceedings of the 2nd CompMusic Workshop* (2012).

[KMS11] Koduri, G. K., M. Miron Serrà Julià, J., and Serra, X. “Computational approaches for the understanding of melody in carnatic music”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2011).

[Kod+12a] Koduri, G. K. et al. “Rāga recognition based on pitch distribution methods”. In: *Journal of New Music Research* vol. 41, no. 4 (2012), pp. 337–350.

[Kod+12b] Koduri, G. K. et al. “Rāga recognition based on pitch distribution methods”. In: *Journal of New Music Research (JNMR)* vol. 41, no. 4 (2012), pp. 337–350.

[Kod+14a] Koduri, G. K. et al. “Intonation analysis of ragas in Carnatic music”. In: *Journal of New Music Research* vol. 43, no. 1 (2014), pp. 73–94.

[Kod+14b] Koduri, G. K. et al. “Intonation analysis of rāgas in Carnatic music”. In: *Journal of New Music Research (JNMR)* vol. 43, no. 1 (2014), pp. 72–93.

[Kod14] Koduri, G. K. “Culture-aware approaches to modeling and description of intonation using multimodal data”. In: *nternational Conference on Knowledge Engineering and Knowledge Management* (2014), pp. 209–217.

[Kod16] Koduri, G. K. “Towards a multimodal knowledge base for Indian art music: A case study with melodic intonation”. PhD thesis. Universitat Pompeu Fabra, Spain, 2016.

[Kri03a] Krishnaswamy, A. “Application of pitch tracking to South Indian classical music”. In: *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)* (2003), p. 389.

[Kri03b] Krishnaswamy, A. “Pitch measurements versus perception of south indian classical music”. In: *Stockholm Music Acoustics Conference (SMAC)* (2003), pp. 627–630.

[Kri04a] Krishnaswamy, A. “Inflexions and microtonality in south indian classical music”. In: *Frontiers of Research on Speech and Music (FRSM)* (2004).

[Kri04b] Krishnaswamy, A. “Melodic Atoms for Transcribing Carnatic Music”. In: *International Symposium on Music Information Retrieval*. 228. ISMIR. 2004.

[Kri04c] Krishnaswamy, A. “Multi-dimensional musical atoms in South-Indian classical music”. In: *International Conference on Music Perception and Cognition* (2004).

[KS03] Krishnaswampy, A. and Smith, J. O. “Methods for Simulating String Collisions with Rigid Spatial Obstacles”. In: *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 2003, pp. 233–236.

[KS83] Karplus, K. and Strong, A. “Digital Synthesis of Plucked-String and Drum Timbres”. In: vol. 7, no. 2 (1983), pp. 43–55.

[KSS12a] Koduri, G. K., Serrà Julià, J., and Serra, X. “Characterization of intonation in carnatic music by parametrizing pitch histograms”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2012).

[KSS12b] Koduri, G. K., Serrà Julià, J., and Serra, X. “Characterization of intonation in carnatic music by parametrizing pitch histograms”. In: *Proceedings of the 13th International Society for Music Information Retrieval Conference; 2012 Oct 8-12;* 2012.

[Kur+15] Kuriakose, J. et al. “Akshara Transcription of Mrudangam Strokes in Carnatic Music”. In: *Proceedings of the 21st National Conference on Communication (NCC)*. Mumbai, India, Feb. 2015.

[Lic41] Lichte, W. H. “Attributes of complex tones.” In: *Journal of Experimental Psychology* vol. 28, no. 6 (1941), p. 455.

[Lon14] London, J. *Hearing in time: Psychological aspects of musical meter*. Oxford University Press, 2014.

[Lor16] Lord Rayleigh, O. F. “XII. On the propagation of sound in narrow tubes of variable section”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* vol. 31, no. 182 (1916), pp. 89–96. eprint: https://doi.org/10.1080/14786440208635477.

[LPS19] Lluís, F., Pons, J., and Serra, X. *End-to-end music source separation: is it possible in the waveform domain?* 2019. arXiv: 1810.12187 [cs.SD].

[Man15] Manuel, P. "The Intermediate Sphere in North Indian Music Culture: Between and Beyond "Folk" and "Classical"". In: *Ethnomusicology* vol. 59, no. 1 (2015), pp. 82–115.

[MC19] Madhusudhan, S. T. and Chowdhary, G. "Deepsrgm-Sequence classification and ranking in Indian classical music with deep learning". In: *International Society for Music Information Retrieval Conference (ISMIR)* (2019), pp. 533–540.

[MD12] Mauch, M. and Dixon, S. "A Corpus-based Study of Rhythm Patterns". In: *Proceedings of the 13th International Conference on Music Information Retrieval (ISMIR 2012)*. Porto, Portugal, Oct. 2012, pp. 163–168.

[Mee00] Meer, W. V. d. "Theory and practice of intonation in Hindusthani music". In: *The Ratio Book* (2000), pp. 50–71.

[Mod70] Modak, H. "Automatic musical instrument in aid of research in Indian music". In: *Journal of Sangeet Natak Akademi* vol. 91 (1970).

[MP20] Mohandas, H. and Padmanabhan, C. "Analytical Models for Vibro-Acoustics of a Mridangam". In: *SPARC Monograph* (2020).

[MSS17] Maestre, E., Scavone, G., and Smith, J. "Joint Modeling of Bridge Admittance and Body Radiativity for Efficient Synthesis of String Instrument Sound by Digital Waveguides". In: *IEEE Transactions on Audio, Speech and Language Processing* vol. 25, no. 5 (2017), pp. 1128–1139.

[MSu09] M.Subramanian. "Carnatic Music - Automatic Computer Synthesis of Gamakams". In: *Journal of the Sangeet Natak Akademi (Academy),* vol. 43, no. 3 (2009), pp. 28–36.

[MW15] Mandal, A. and Wahi, P. "Natural frequencies, modeshapes and modal interactions for strings vibrating against an obstacle: Relevance to Sitar and Veena". In: *Journal of Sound and Vibration* vol. 338 (2015), pp. 42–59.

[MWS16a] Mansour, H., Woodhouse, J., and Scavone, G. "Enhanced time-domain modelling of musical strings. Part 1: Plucked strings". In: *Acta Acustica united with Acustica* vol. 102 (2016), pp. 1082–1093.

[MWS16b] Mansour, H., Woodhouse, J., and Scavone, G. "Enhanced time-domain modelling of musical strings. Part 2: Bowed strings". In: *Acta Acustica united with Acustica* vol. 102 (2016), pp. 1094–1107.

[Nar94] Narasimhan, S. "THE ROLE OF TALA VADYAS IN CARNATIC AND HINDUSTANI MUSIC: A COMPARATIVE ASSESSMENT". In: *Proceedings of The Talavadya Seminar No. 1* (1994).

[Neu12] Neuman, D. "Pedagogy, practice, and embodied creativity in Hindustani music". In: *Ethnomusicology* vol. 56, no. 3 (2012), pp. 426–449.

[Pad16] Padmasundari. "Music information retrieval: Raga identification using machine learning". MA thesis. Indian Institute of Technology, Madras (IITM), India, 2016.

[PE05] Primary, D. of and Education, S. *Karnataka Talavadya, Vidwat Purva Hanta (Per-Proficiency Grade)*. Government of Karnataka, 2005.

[PF12] Peeters, G. and Fort, K. "Towards a (Better) Definition of the Description of Annotated MIR Corpora". In: *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012)*. Porto, Portugal, Oct. 2012, pp. 25–30.

[PG18] Pisharody, R. and Gupta, A. "Experimental Investigations of Tānpurā Acoustics". In: *Acta Acustica united with Acustica* vol. 104, no. 3 (2018), pp. 542–545.

[Pil17] Pillai, M. A. "Analysis and design of acoustic energy harvesters using duct resonators and piezoelectric unimorphs". PhD thesis. National Institute of Technology, Tiruchirappalli, 2017.

[PM17] Padmasundari and Murthy, H. "Raga Identification using locality sensitive hashing". In: *National Conference on Communications (NCC)*, no. 1–6 (2017).

[PMK10] Paulus, J., Müller, M., and Klapuri, A. "Audio-based Music Structure Analysis". In: *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*. Utrecht, The Netherlands, 2010, pp. 625–636.

[Pol+07] Poliner, G. et al. "Melody transcription from music audio: Approaches and evaluation". In: *IEEE Transactions on Audio, Speech, and Language Processing (TASLP)* vol. 15, no. 4 (2007), pp. 1247–1256.

[Pow80] Powers, H. S. "Classical music, cultural roots, and colonial rule: an Indic musicologist looks at the Muslim world". In: *Asian Music* vol. 12, no. 1 (1980). https://doi.org/10.2307/833796, pp. 5–39.

[PP11] Peeters, G. and Papadopoulos, H. "Simultaneous Beat and Downbeat-Tracking Using a Probabilistic Framework: Theory and Large-Scale Evaluation". In: *IEEE Transactions on Audio, Speech and Language Processing* vol. 19, no. 6 (2011), pp. 1754–1769.

[Pra11] Pradhan, A. *Tabla: A Performer's Perspective*. BookBaby, 2011.

[PSS13a] Porter, A., Sordo, M., and Serra, X. "Dunya: A System for Browsing Audio Music Collections Exploiting Cultural Context". In: *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR 2013)*. Curitiba, Brazil, Nov. 2013, pp. 101–106.

[PSS13b] Porter, A., Sordo, M., and Serra, X. "Dunya: A system for browsing audio music collections exploiting cultural context". In: *Britto A, Gouyon F, Dixon S. 14th International Society for Music Information Retrieval Conference (ISMIR); 2013 Nov 4-8; Curitiba, Brazil.[place unknown]: ISMIR; 2013. p. 101-6.* International Society for Music Information Retrieval (ISMIR). 2013.

[Rae10] Rae, J. "The Science of String Instruments". In: ed. by Rossing, T. New York, NY: Springer, 2010. Chap. The Banjo.

[Raf+14] Raffel, C. et al. "mir_eval: A Transparent Implementation of Common MIR Metrics". In: *Proceedings of the 15th International Society for Music Information Retrieval (ISMIR) Conference*. Taipei, Taiwan, Oct. 2014.

[Rag53] Raghavan, V. "Why is the Mridanga so-called?" In: *The Journal of the Music Academy* vol. 24 (1953), pp. 135–136.

[Raj+17a] Raj, P. V. K. et al. "An Approach to Transcription of Varnams in Carnatic Music Using Hidden Markov Models". In: *Proc. of the National Conference on Communication (NCC)*. 2017.

[Raj+17b] Raj, P. V. K. et al. "An Approach to Transcription of Varnams in Carnatic Music Using Hidden Markov Models". In: *National Conference on Communications (NCC)* (2017), pp. 1–6.

[Ram04] Ramanathan, H. *Ragalakshanasangraha (Collection of Raga Descriptions)*. 1st. N. Ramanathan, R.A.Puram, Chennai, 2004.

[Ram21] Raman, C. “On some Indian stringed instruments”. In: *Proceedings of the Indian Association for the Cultivation of Science* vol. 7, no. 29–33 (1921), pp. 468–472.

[Ram22] Raman, C. V. “The acoustical knowledge of the ancient Hindus”. In: *Sir Asutosh Mookerjee Silver Jubilee Volumes*. Vol. 2. University Press, Calcutta, 1922, pp. 473–478.

[Ram34] Raman, C. V. “The Indian musical drums”. In: *Proceedings of the Indian Academy of Sciences-Section A*. Vol. 1. Springer. 1934, pp. 179–188.

[Ram35] Raman, C. “The Indian Musical Drums”. In: *Proceedings of the Indian Academy of Sciences* vol. A1, no. 179–188 (1935), pp. 455–467.

[Ran+19a] Ranjani, H. G. et al. “A compact pitch and time representation for melodic contours in Indian art music”. In: *The Journal of the Acoustical Society of America* vol. 145, no. 1 (2019), pp. 597–603.

[Ran+19b] Ranjani, H. et al. “A compact pitch and time representation for melodic contours in Indian art music”. In: *The Journal of the Acoustical Society of America (JASA)* vol. 145, no. 1 (2019), pp. 597–603.

[Ran+19c] Ranjani, H. et al. “A compact pitch and time representation for melodic contours in Indian art music”. In: *The Journal of the Acoustical Society of America* vol. 145, no. 1 (2019), pp. 597–603.

[Ran11] Ranade, A. D. *Some Hindustani Musicians: They Lit the Way!* Promilla & Company, Publishers in association with Bibliophile South Asia., 2011.

[Ran20] Ranjani, H. G. “Computational Approaches to Structural Analysis of Indian Art Music Signals”. MA thesis. Indian Institute of Science (IISc), Bangalore, India, Jan. 2020.

[Rao+14] Rao, P. et al. “Melodic motivic analysis of indian music”. In: *Journal of New Music Research (JNMR)* vol. 43, no. 1 (2014), pp. 115–130.

[Rao06a] Rao, T. K. G. *Compositions of Mudduswami Dīkshitar*. Ganamandir Publications, 2006.

[Rao06b] Rao, T. K. G. *Compositions of Śyāmā Śāstrī*. Ganamandir Publications, 2006.

[Rao38] Rao, K. N. “Theory of the Indian musical drums—Part I”. In: *Proceedings of the Indian Academy of Sciences - Section A* vol. 7, no. 2 (Feb. 1938), pp. 75–84.

[Rao95] Rao, T. K. G. *Compositions of Tyagaraja*. Chennai: Ganamandir, 1995.

[RAS11] Ranjani, H. G., Arthi, S., and Sreenivas, T. V. “Carnatic music analysis: Shadja, swara identification and raga verification in alapana using stochastic models”. In: *Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)* (2011), pp. 29–32.

[Rat15] Ratanjanakar, S. N. *Chaturdandi Prakāshikā (Hindi Translation)*. Sanskar Prakashan, 2015.

[RK20] Raman, C. and Kumar, S. “Musical Drums with Harmonic Overtones”. In: *Nature* vol. 104, no. 2620 (1920), p. 500.

[RM13] Rajan, R. and Murthy, H. “Group delay based melody mono pitch extraction from music”. In: *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)* (2013), pp. 186–190.

[RPS17a] Ranjani, H., Paramashivan, D., and Sreenivas, T. “Quantized Melodic Contours in Indian Art Music Perception: Application to Transcription”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2017), pp. 174–180.

[RPS17b] Ranjani, H., Paramashivan, D., and Sreenivas, T. V. “Quantized Melodic Contours in Indian Art Music Perception: Application to Transcription.” In: *ISMIR*. 2017, pp. 174–180.

[RPS19] Ranjani, H. G., Paramashivan, D., and Sreenivas, T. V. “Discovering structural similarities among rāgas in Indian Art Music: a computational approach”. In: *Sādhanā* vol. 44, no. 5 (2019), p. 120.

[RR14] Rao, S. and Rao, P. “An overview of Hindustani music in the context of computational musicology”. In: *Journal of new music research* vol. 43, no. 1 (2014). https://doi.org/10.1080/09298215.2013.831109, pp. 24–33.

[RS13] Ranjani, H. and Sreenivas, T. “Grouping carnatic music notes using a multi-gram language model”. In: *Proceedings of Acoustics 2013*. New Delhi, India, Nov. 2013.

[RS15] Ranjani, H. G. and Sreenivas, T. V. “Multi-instrument detection in polyphonic music using Gaussian Mixture based factorial HMM”. In: *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2015, pp. 191–195.

[RS17] Ranjani, H. G. and Sreenivas, T. V. “Raga identification using repetitive note patterns from prescriptive notations of Carnatic music”. In: *arXiv* (2017), p. 1711.

[RS54] Ramakrishna, B. S. and Sondhi, M. M. “Vibrations of Indian Musical Drums Regarded as Composite Membranes”. In: *The Journal of the Acoustical Society of America* vol. 26, no. 4 (1954), pp. 523–529. eprint: https://doi.org/10.1121/1.1907369.

[RVR20] Rohit, M. A., Vinutha, T. P., and Rao, P. “Structural segmentation of dhrupad vocal bandish audio based on tempo”. In: *Proc. of int. soc. for music information retrieval conf.(ismir).(Montreal, Canada)*. 2020.

[RVV18] Ragesh Rajan, M., Vijayakumar, A., and Vijayasenan, D. “Prediction of aesthetic elements in Karnatic Music: A machine learning approach”. In: *Proceedings of Interspeech* (2018), pp. 2042–2046.

[S89] S., M. “Psychological constraints on form-bearing dimensions in music”. In: *Contemporary Music Review* vol. 4, no. 1 (1989), pp. 181–198.

[SA09] Shetty, S. and Achary, K. “Raga mining of Indian music by extracting arohana-avarohana pattern.” In: *nternational Journal of Recent Trends in Engineering* (2009), p. 362.

[Sam58] Sambamoorthy, P. *South Indian Music, (2019 reprint)*. Vol. I to VI. Indian Music Pub. House, 1958.

[Sam98a] Sambamoorthy, P. *South Indian Music*. The Karnatic Music Book Centre, Chennai, 1998.

[Sam98b] Sambamoorthy, P. *South Indian Music Vol. I-VI*. The Indian Music Publishing House, 1998.

[San17] Sankaran, S. “Segmentation of Music Items: A Computational Perspective”. MA thesis. Indian Institute of Technology, Madras (IITM), India, 2017.

[Sar11] Saraf, R. *Development of Hindustani Classical Music (19th & 20th centuries)*. Vidyanidhi Prakashan, 2011.

[Sat01] Sathyanarayana, R. *Chaturdandi Prakasika of Venkatatmakhin (Vols I and II)*. Kalamulasastra Series Vol. 24 and 25, 2001.

[Sax06] Saxena, S. K. *The Art of Tablā Rhythm: Essentials, Tradition, and Creativity*. 8. Sangeet Natak Akademi, 2006.

[SC12a] Srinivasamurthy, A. and Chordia, P. “A Unified System for Analysis and Representation of Indian Classical Music using Humdrum Syntax”. In: *Proceedings of the 2nd CompMusic Workshop*. Istanbul, Turkey, July 2012, pp. 38–42.

[SC12b] Srinivasamurthy, A. and Chordia, P. “Multiple Viewpoint Modeling of North Indian Classical Vocal Compositions”. In: *Proceedings of the 9th International Symposium on Computer Music Modeling and Retrieval (CMMR)*. London, UK, June 2012, pp. 344–356.

[SC12c] Srinivasamurthy, A. and Chordia, P. “Multiple viewpoint modeling of north Indian classical vocal compositions”. In: *Proceedings of the International Symposium on Computer Music Modeling and Retrieval*. Citeseer. 2012.

[Sch10] Schofield, K. B. “Reviving the golden age again:“Classicization,” Hindustani music, and the Mughals”. In: *Ethnomusicology* vol. 54, no. 3 (2010). https://doi.org/10.5406/ethnomusicology.54.3.0484, pp. 484–517.

[See58] Seeger, C. “Prescriptive and Descriptive Music-Writing”. In: *The Musical Quarterly* vol. 44, no. 2 (1958), pp. 184–195.

[Sen+83] Sengupta, R. et al. “Tonal quality of the Indian tanpura”. In: *Proceedings of the 1983 Stockholm Music Acoustics Conference*. 1983.

[Ser11] Serra, X. “A multicultural approach in Music Information Research”. In: *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*. Miami, USA, Oct. 2011, pp. 151–156.

[Ser14] Serra, X. “Creating Research Corpora for the Computational Study of Music: the case of the CompMusic Project”. In: *Proceedings of the 53rd AES International Conference on Semantic Audio*. London, Jan. 2014.

[SG09] Sridhar, R. and Geetha, T. “Raga identification of carnatic music for music information retrieval”. In: *nternational Journal of recent trends in Engineering,* (2009), p. 571.

[SG12a] Salamon, J. and Gómez, E. “Melody extraction from polyphonic music signals using pitch contour characteristics”. In: *IEEE Transactions on Audio, Speech, and Language Processing (TASLP)* vol. 20, no. 6 (2012), pp. 1759–1770.

[SG12b] Salamon, J. and Gómez, E. “Melody Extraction From Polyphonic Music Signals Using Pitch Contour Characteristics”. In: *IEEE Transactions on Audio, Speech, and Language Processing* vol. 20, no. 6 (Aug. 2012), pp. 1759–1770.

[SGS12] Salamon, J., Gulati, S., and Serra, X. “A multipitch approach to tonic identification in Indian classical music”. In: *International Society for Music Information Retrieval Conference (ISMIR)* (2012), pp. 499–504.

[SHS14] Srinivasamurthy, A., Holzapfel, A., and Serra, X. “In Search of Automatic Rhythm Analysis Methods for Turkish and Indian Art Music”. In: *Journal of New Music Research* vol. 43, no. 1 (2014), pp. 97–117.

[SHS17] Srinivasamurthy, A., Holzapfel, A., and Serra, X. “Informed Automatic Meter Analysis of Music Recordings”. In: *Proceedings of the 18th International Society for Music Information Retrieval (ISMIR) Conference*. Suzhou, China, Oct. 2017, pp. 679–685.

[Sin13] Singh, P. G. “Perception and orchestration of melody, harmony and rhythm on instruments with ’chikari’ strings”. In: *Proceedings of Meetings in Acoustics*. 2013.

[SIS19] Srinivasamurthy, A., Ishwar, V., and Serra, X. “Statistical Analysis of Tempo and Rhythmic Elaboration in Carnatic Music”. In: *Proceedings of the 24th International Symposium: Frontiers of Research in Speech and Music (FRSM - 2019)*. Kanpur, India, July 2019.

[SKX16] Sentürk, S., Koduri, G. K., and Xavier, S. “A score-informed computational description of svaras using a statistical model”. In: *Sound & Music Computing Conference* (2016), pp. 427–433.

[Sma04] Smaragdis, P. “Non-negative Matrix Factor Deconvolution; Extraction of Multiple Sound Sources from Monophonic Inputs”. In: *Independent Component Analysis and Blind Signal Separation*. Ed. by Puntonet, C. G. and Prieto, A. Vol. 3195. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2004, pp. 494–499.

[Smi92] Smith, J. O. “Physical Modeling Using Digital Waveguides”. In: *Computer Music Journal* vol. 16, no. 4 (Winter 1992). Special issue: Physical Modeling of Musical Instruments, Part I, pp. 74–91.

[Soo18] Sooraj, G. “Vibro-acoustics of a Mridangam”. MA thesis. Indian Institute of Technology Madras, Chennai, 2018.

[Sri+12] Srinivasamurthy, A. et al. “A Beat Tracking Approach to Complete Description of Rhythm in Indian Classical Music”. In: *Proceedings of the 2nd CompMusic Workshop*. Istanbul, Turkey, July 2012, pp. 72–78.

[Sri+14a] Srinivasamurthy, A. et al. “Transcription and Recognition of Syllable based Percussion Patterns: The Case of Beijing Opera”. In: *Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR 2014)*. Taipei, Taiwan, Oct. 2014, pp. 431–436.

[Sri+14b] Srinivasamurthy, A. et al. “Corpora for Music Information Research in Indian Art Music”. In: *Proceedings of Joint 42nd International Computer Music Conference/13th Sound and Music Computing Conference*. Athens, Greece, Sept. 2014, pp. 1029–1036.

[Sri+14c] Srinivasamurthy, A. et al. “Corpora for music information research in indian art music”. In: *Georgaki A, Kouroupetroglou G, eds. Proceedings of the 2014 International Computer Music Conference, ICMC/SMC; 2014 Sept 14-20; Athens, Greece.[Michigan]: Michigan Publishing; 2014.* Michigan Publishing. 2014.

[Sri+17a] Srinivasamurthy, A. et al. “A Data-driven bayesian approach to automatic rhythm analysis of indian art music”. PhD thesis. Universitat Pompeu Fabra, 2017.

[Sri+17b] Srinivasamurthy, A. et al. “Aspects of Tempo and Rhythmic Elaboration in Hindustani Music: A Corpus Study”. In: *Frontiers in Digital Humanities* vol. 4 (2017).

[Sri+17c] Srinivasamurthy, A. et al. “Aspects of tempo and rhythmic elaboration in hindustani music: A corpus study”. In: *Frontiers in Digital Humanities* vol. 4 (2017), p. 20.

[Sri+21] Srinivasamurthy, A. et al. “Saraga: Open Datasets for Research on Indian Art Music”. In: *Empirical Musicology Review* vol. 16, no. 1 (2021), pp. 85–98.

[Sri04] Sriram, V. *Carnatic Summer: Lives of Twenty Great Exponents*. EastWest Books (Madras), 2004.

[Sri16] Srinivasamurthy, A. “A Data-driven Bayesian Approach to Automatic Rhythm Analysis of Indian Art Music”. Doctoral dissertation. Barcelona, Spain: Universitat Pompeu Fabra, 2016.

[SRN10] Sivaraman, V. U., Ramasami, T., and Naresh, M. *Science for Musical Excellence (Sixth Raja Ramanna memorial Lecture)*. Lecture. Bangalore: NIAS, 2010.

[SS+78] Shringy, R. K., Sharma, P. L., et al. *Sagıta-ratnākara of Śārgadeva: Treatment of svara*. Vol. 1. Motilal Banarsidass, 1978.

[SS14a] Srinivasamurthy, A. and Serra, X. “A Supervised Approach to Hierarchical Metrical Cycle Tracking from Audio Music Recordings”. In: *Proceedings of the 39th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014)*. Florence, Italy, May 2014, pp. 5237–5241.

[SS14b] Srinivasamurthy, A. and Serra, X. “A supervised approach to hierarchical metrical cycle tracking from audio music recordings”. In: *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)* (2014), pp. 5217–5221.

[Sub02] Subramanium, M. “An Analysis of Gamakams using the computer”. In: *Sangeet Natak, Journal of the Sangeet Natak Akademi,* vol. 37 (2002), pp. 26–47.

[Sub12] Subramanian, M. “Generating Computer Music from Skeletal Notations for Carnatic Music Compositions”. In: *Proceedings of the 2nd CompMusic Workshop* (2012).

[Sub13] Subramanian, S. K. “Modelling gamakas of Carnatic music as a synthesizer for sparse prescriptive notation”. PhD thesis. National University of Singapore, Singapore, 2013.

[Sub99] Subramanium, M. “Synthesizing Carnatic Music with a Computer”. In: *Journal of the Sangeet Natak Akademi (Academy),* vol. 133 (1999), pp. 16–24.

[Sun+16] Sundar, A. et al. “Study of Sarasvati Veena – a South Indian musical instrument using its vibro-acoustic signatures”. In: *Journal of Vibroengineering* vol. 18, no. 5 (2016), pp. 3362–3368.

[Sun21] Sunil, E. N. *Resounding Mridangam: The Majestic South Indian Drum*. Erickavu N Sunil, March 2021.

[Sun94] Sundberg, J. “Perceptual aspects of singing”. In: *Journal of voice* vol. 8, no. 2 (1994). https://doi.org/10.1016/S0892-1997(05)80303-0, pp. 106–122.

[SW04] Sanyal, R. and Widdess, R. *Dhrupad. Tradition and Performance in Indian Music.* Ashgate, 2004.

[SWM11a] Subramanian, S., Wyse, L., and McGee, K. “Modeling speed doubling in carnatic music”. In: *ICMC*. 2011.

[SWM11b] Subramanian, S. K., Wyse, L., and McGee, K. “Modeling speed doubling in carnatic music”. In: *International Computer Music Conference* (2011), pp. 478–485.

[SWM12] Subramanian, S. K., Wyse, L., and McGee, K. “A two-component representation for modeling gamakas of carnatic music”. In: *Proceedings of the 2nd CompMusic Workshop* (2012).

[TA15] Titze, I. R. and Abbott, K. V. *Vocology: The science and practice of voice habilitation*. https://books.google.co.in/books?id=RubjjgEACAAJ. Compton Publishing Limited, 2015.

[TG17] Tiwari, S. and Gupta, A. "Effects of air loading on the acoustics of an Indian musical drum". In: *Journal of the Acoustical Society of America* vol. 141, no. 4 (2017), pp. 2611–2621.

[Thi97] Thielemann, S. *The Darbhanga Tradition: Dhrupada in the School of Pandit Vidur Mallik*. Indica Books, 1997.

[Val+91] Valette, C. et al. "The tampura bridge as a precursive wave generator". In: *Acustica* vol. 74 (1991), pp. 201–208.

[VAM17a] Viraraghavan, V. S., Aravind, R., and Murthy, H. "A Statistical Analysis of Gamakas in Carnatic Music". In: *Proceedings of the 18th International Society for Music Information Retrieval Conference*. 2017, pp. 243–249.

[VAM17b] Viraraghavan, V., Aravind, R., and Murthy, H. "A Statistical Analysis of Gamakas in Carnatic Music". In: *International Society for Music Information Retrieval Conference (ISMIR)* (2017), pp. 243–249.

[VBM09] Vyasarayani, C. P., Birkett, S., and McPhee, J. "Modeling the dynamics of a vibrating string with a finite distributed unilateral constraint: Application to the sitar". In: *Journal of the Acoustical Society of America* vol. 125, no. 6 (2009), pp. 3673–3682.

[VE75] Von Helmholtz, H. and Ellis, A. J. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. London: Longmans, Green and Company, 1875.

[Vij07a] Vijayakrishnan, K. G. *The grammar of Carnatic musi*. De Gruyter Mouton, 2007.

[Vij07b] Vijayakrishnan, K. *The grammar of Carnatic music*. Vol. 8. Walter de Gruyter, 2007.

[Vin+16] Vinutha, T. P. et al. "Structural Segmentation and Visualization of Sitar and Sarod Concert Audio." In: *ISMIR*. 2016, pp. 232–238.

[Vin12] Vinayakam, R. *Gamaka Box- the first ever Comprehensive Notation System for Indian Music*. http://www.rameshvinayakam.com/research. Accessed: 2016-11-13. 2012.

[Vir+17] Viraraghavan, V. S. et al. "Non-uniform time-scaling of Carnatic music transients". In: *arXiv preprint arXiv:1711.02318* (2017).

[Vir+19] Viraraghavan, V. S. et al. "Visualizing Carnatic music as projectile motion in a uniform gravitational field". In: *Proc. Workshop on Speech, Music and Mind 2019*. 2019, pp. 31–35.

[Vir+20a] Viraraghavan, V. S. et al. "State-Based Transcription of Components of Carnatic Music". In: *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2020, pp. 811–815.

[Vir+20b] Viraraghavan, V. S. et al. "Data-driven measurement of precision of components of pitch curves in Carnatic music". In: *The Journal of the Acoustical Society of America* vol. 147, no. 5 (2020), pp. 3657–3666.

[Vir+20c] Viraraghavan, V. S. et al. "State-based transcription of Carnatic music components". In: *Int. Conference on Audio, Speech and Signal Processing 2020*. IEEE. 2020, accepted.

[Vir+20d] Viraraghavan, V. et al. "Data driven measurement of precision of components of pitch curves in Carnatic music". In: *The Journal of the Acoustical Society of America (JASA)* vol. 147, no. 5 (2020), pp. 3657–3666.

[Vir+20e] Viraraghavan, V. et al. "State-based Transcription Of Components Of Carnatic Music". In: *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)* (2020), pp. 811–815.

[VMA18] Viraraghavan, V., Murthy, H., and Aravind, R. "Precision of sung notes in Carnatic music". In: *International Society for Music Information Retrieval Conference (ISMIR)* (2018), pp. 499–505.

[Vod62] Vodička, V. "Free Vibrations of a Composite Circular Membrane". In: *Journal of the Physical Society of Japan* vol. 17, no. 4 (1962), pp. 698–702. eprint: https://doi.org/10.1143/JPSJ.17.698.

[Vod63] Vodička, V. "On forced vibrations of composite circular membranes". In: *Czechoslovak Journal of Physics* vol. 13, no. 7 (July 1963), pp. 493–498.

[VR] Verma, P. and Rao, P. "DISTINGUISHING MUSICAL INSTRUMENT PLAYING STYLES WITH ACOUSTIC SIGNAL ANALYSES". In: ().

[Wad16] Wade, B. C. *Khyal: Creativity within North India's classical music tradition*. https://books.google.co.in/books?id=MiE9AAAAIAAJ&dq=aneesh pradhan&source=gbs_similarbooks. Munshiram Manoharlal Pub. Pvt. Ltd., 2016.

[Wal20] Walker, M. E. "Towards a decolonized music history curriculum". In: *Journal of Music History Pedagogy* vol. 10, no. 1 (2020), pp. 1–19.

[WBM16] Walstijn, M. V., Bridges, J., and Mehes, S. "A Real-Time Synthesis Oriented Tanpura Model". In: *Proceedings of the 2016 International Conference on Digital Audio Effects (DAFx-16)*. Sound examples and bridge response at: http://www.socasites.qub.ac.uk/mvanwalstijn/dafx16a/. 2016, pp. 175–183.

[Web19] Webster, A. G. "Acoustical Impedance and the Theory of Horns and of the Phonograph". In: *Proceedings of the National Academy of Sciences* vol. 5, no. 7 (1919), pp. 275–282. eprint: https://www.pnas.org/content/5/7/275.full.pdf.

[Wil+16] Wilkinson, M. D. et al. "The FAIR Guiding Principles for scientific data management and stewardship". In: *Scientific Data* vol. 3, no. 1 (2016).

[Woo04] Woodhouse, J. "On the synthesis of guitar plucks". In: *Acta Acustica united with Acustica* vol. 90 (2004), pp. 928–944.

[Woo14] Woodhouse, J. "The acoustics of the violin: a review". In: *Reports on Progress in Physics* vol. 77 (2014).

[Woo21] Woodhouse, J. "Euphonix: The Science of Musical Instruments". In: https://euphonics.org. 2021. Chap. 5. Strings, mainly plucked.

[ZCY20] Zhang, W., Chen, Z., and Yin, F. "Multi-Pitch Estimation of Polyphonic Music Based on Pseudo Two-Dimensional Spectrum". In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)* vol. 28 (2020), pp. 2095–2108.

Glossary

Abhogi (*Ābhōgi*) A rāga in Carnatic Music

adi (*ādi*) A tāḷa with 32 akṣaras in a cycle

akshara (*akṣara*) The lowest metrical pulse (subdivision)

alap (*ālāp*) An unmetered melodic improvisation

alapana (*ālāpana*) An unmetered melodic improvisation

amad (*āmad*) (literally approach) A phrase leading to a sam

anga (*aṅga*) The sections of a tāḷa

ati-dhrut (*ati-dhr̥t*) Very fast tempo class

ati-vilambit (*ati-vilaṁbit*) Very slow tempo class

avart (*āvart*) One complete cycle of a tāl

avartana (*āvartana*) One complete cycle of a tāḷa

bandish (*bandiś*) A fixed melodic composition in Hindustani music

bayan (*bāyān*) The left drum

Begada (*Bēgaḍa*) A rāga in Carnatic Music

bol (*bōl*) The onomatopoeic oral percussion syllables of the tabla

carnatic (*carṇāṭic*) is an art music tradition of India

charana (*caraṇa*) The end section of a Carnatic music composition

chaturashra (*caturaśra*) A naḍe with 2 or 4 akṣaras per beat

dayan (*dāyān*) The right drum

dhrupad A music style in Hindustani music

dhrut (*dr̥t*) Fast tempo class

dhruta (*dhr̥tta*) The fast tempo class

digga (*diggā*) Alternative name for the left drum

edupu (*eḍupu*) The phase/offset of the composition relative to the sama

ektal (*ēktāl*) A tāl with 12 mātrās in a cycle

gat (*gaṭ*) A compositional form in tabla

gharana (*gharānā*) The stylistic schools of Hindustani music

ghatam (*ghaṭam*) A percussion instrument used in Carnatic music (specially made clay pot with a narrow mouth)

Hindustani (*Hindustāni*) An art music tradition of India

jati (*jāti*) The performer or the percussionist

jhaptal (*jhaptāl*) A tāl with 10 mātrās in a cycle

kaala (*kāla*) The tempo multiplying factor

kalai (*kaḷai*) is the number of count of each of the beats of the tala

Kalyani (*Kalyāṇi*) A rāga in Carnatic Music

kayda (*kāyadā*) A compositional form in tabla

khali (*khālī*) A hand wave in the tāl cycle (unaccented)

khanda A naḍe with 5 akṣaras per beat

khanda chapu (*khaṇḍa chāpu*) A tāḷa with 10 akṣaras in a cycle

khanjira (*khañjira*) A tambourine like percussion instrument used in Carnatic music

khayal (*khyāl*) A music style in Hindustani music

konnakol (*konnakōl*) The art form of reciting percussion syllables

kriti (*kr̥ti*) A common compositional form in Carnatic music

kutcheri (*kachēri*) A concert of Carnatic music

lay The tempo class

madhya Medium tempo class

madhyama The medium tempo class

matra (*mātrā*) The lowest defined metrical pulse in Hindustani music (equivalent to a beat)

mehfil A musical performance in an intimate gathering

mishra (*miśra*) A naḍe with 7 akṣaras per beat

mishra chapu (*miśra chāpu*) A tāḷa with 14 akṣaras in a cycle

Mohanam (*Mōhanaṁ*) A rāga in Carnatic Music

morsing (*mōrsiṅg*) The Indian jaw (jew's) harp

mridangam (*mr̥daṅgaṁ*) The primary percussion accompaniment in Carnatic music (common spelling mridangam)

muttuswami dikshitar (*muttusvāmi dīkṣitar*) A prominent Carnatic music composer

nade (*naḍe*) The subdivision structure within a beat

pakhavaj (*pakhāvaj*) A double barrel drum used as rhythm accompaniment in Hindustani music

palta (*palaṭā*) A compositional form in tabla

peshkar (*pēśkār*) A compositional form in tabla

raag (*rāg*) The melodic framework of Hindustani music

raga (*rāga*) The melodic framework of Carnatic music

rela (*rēlā*) A compositional form in tabla

rupak taal (*rūpak tāl*) A tāl with 7 mātrās in a cycle

rupaka (*rūpaka*) A tāḷa with 12 akṣaras in a cycle

Sahana (*Sahāna*) A rāga in Carnatic Music

sam The first mātrā of an āvart

sama The beginning of an āvartana (equivalent to a downbeat)

sankeerna (*saṁkīrṇa*) A naḍe with 9 akṣaras per beat

santoor (*santūr*) A trapezoid-shaped hammered dulcimer used in Hindustani music

sarangi (*sāraṅgi*) A bowed music instrument used in Hindustani music

sarod (*sarōd*) A fretless plucked string instrument used in Hindustani music

Saveri (*Sāvēri*) A rāga in Carnatic Music

shruti Microtune

shyama shastri (*śyāmā śāstri*) A prominent Carnatic music composer

sitar (*sitār*) A fretted plucked string instrument used in Hindustani music

solkattu (*solkaṭṭu*) The onomatopoeic oral percussion syllables

Sri (*Śrī*) A rāga in Carnatic Music

swara A musical note

taal (*tāl*) The rhythmic framework of Hindustani music

tabalchi (*tabalacī*) The performer or the percussionist

tabla The primary percussion accompaniment in Hindustani music

tala (*tāḷa*) The rhythmic framework of Carnatic music

tambura (*tambūra*) The drone instrument used in Carnatic music

tani Short for tani-āvartana

tani avartana (*tani-āvartana*) The solo performance of a percussion ensemble

tanpura (*tānpura*) The drone music instrument used in Hindustani music

tawaaif courtesan

teental (*tīntāl*) A tāl with 16 mātrās in a cycle

thali (*thālī*) A hand clap in the tāl cycle (accented)

theka (*ṭhēkā*) The basic bōl pattern associated with a tāl

thumri A vocal genre of Hindustani music, associated with tawaaif (courtesan) culture

tillana (*tillāna*) A rhythmic piece in Carnatic music widely used in dance performances

tishra (*tiśra*) A naḍe with 3 or 6 akṣaras per beat

tyagaraja (*tyāgarāja*) A prominent Carnatic music composer

varnam (*varṇaṁ*) A musical form in Carnatic music

veena (*vīṇā*) A fretted string instrument used in Carnatic music

vibhaag (*vibhāg*) The sections of a tāl cycle

vilambit (*vilaṁbit*) Slow tempo class

vilambita (*viḷaṁbita*) The slow tempo class