

ninād



VOLUME 24, DECEMBER 2010
JOURNAL OF ITC SANGEET RESEARCH ACADEMY

ninād

Journal of the ITC Sangeet Research Academy

Vol. 24, December, 2010

Chief Editors

Prof. Asoke Kumar Datta
ITC Sangeet Research Academy, Kolkata

Dr. Ranjan Sengupta,
ITC Sangeet Research Academy, Kolkata

Associate Editors

Prof. Alicja Wieczorkowska,
Polish-Japanese Institute of Information Technology, Poland

Prof. Kristoffer Jensen,
Aalborg University, Denmark

Dr. Soubhik Chakraborty,
BIT, Mesra, Ranchi

Mr. Nityananda Dey,
ITC Sangeet Research Academy, Kolkata

ITC Sangeet Research Academy is a public charitable trust created by ITC Ltd. for the preservation and propagation of Indian classical music tradition.

© ITC Sangeet Research Academy, Kolkata

ISSN 0973-3787 ninād

The Editors of Ninād hold no responsibility for the views / opinions expressed by the individual contributors.

All enquiries to be addressed to the Editor, Ninad
ITC Sangeet Research Academy, 1, Netaji Subhas Chandra Bose
Road, Tollygunge, Kolkata 700 040, India

Published by
ITC Sangeet Research Academy
1, Netaji Subhas Chandra Bose Road,
Tollygunge, Kolkata 700 040, India

Contents

1. A Geometric Model of a musical performer's Ontological Space Guerino Mazzola	1
2. Music Emotion and the Brain Shantala Hegde	16
3. Techne: Revealing Sound, Space and Self in Music for Flute and Electronics Jean penny	28
4. Study of Source Characteristics in Sarod from the Sound Signals R. Sengupta, N. Dey and A. K. Datta	44
5 A note on the problem of finding similarity between Melodies of Unequal Length Soubhik Chakraborty, Swarima Tewari and Gargi Akhoury	52
6. Affine Musical Transformations using Multi-Touch Gestures Florian Thalmann and Guerino Mazzola	58

EDITORIAL

It is heartening to note that scientists from institutions of higher learning in India, particularly those of science and technology, are coming up to objectively investigate the issues pertaining to the domain of music in greater numbers. Of the six articles in the last issue of *Ninad*, five are from such sources in India. It seems that the *Scientific Branch of Musicology* is taking off in India. The need of objective clarity in vexed issues relating to music has always been there and we hope that *Ninad* would continue to promote it zealously.

Of many such issues, two require focused attention. The first one relates to emotion evoked by Indian music. One may note an important difference here from western music. Since Indian music is melodic, intonation is likely to play the most decisive role, leaving timbre far behind. Cognitive studies related to emotion using psycho-perceptual listening tests are gaining ground and are providing interesting insight. However these need to be supplemented by harder objective evidences and deeper understanding through the use of bio-engineering tools of EEG (electroencephalogram), PET (positron emission tomography), fMRI (Functional magnetic resonance imaging) etc. This calls for close collaboration between scientists, medical researchers, musicians and music lovers.

The other one relates to music pedagogy. In India, music teaching follows the oral tradition path, the so-called highly eulogized *Guru-Shishya* parampara. A notable feature of this pedagogy is the deliberate exclusion of all other knowledge sources and apathy to scientific approaches. A steady deterioration of class in performance during the last 100 years is often talked about. It is not surprising when one notices that *Guru-Shishya* parampara is just a chain of communication and the strength of a chain is determined by that of its weakest link. It is now necessary to put the best minds and brains to look afresh at music teaching, using the vast resources of science and technology, so that Indian music may evolve in keeping with the Indian ethos.

Ninad wishes to see more quality articles in these two areas in future.

Ashok Kumar Datta
Ranjan Sengupta
Joint Editors

A NOTE ON THE PROBLEM OF FINDING SIMILARITY BETWEEN MELODIES OF UNEQUAL LENGTH

Soubhik Chakraborty*, Swarima Tewari and Gargi Akhoury
Department of Applied Mathematics
Birla Institute of Technology Mesra, Ranchi 835215, India
*soubhikc@yahoo.co.in

Abstract

The problem of finding similarity between melodies of unequal length has been taken up here. If two melodies are of length w and s where $w < s$ (say), our suggestion is to take all possible segments of length w from the longer melody (which is of length s) and then compare each segment with the shorter melody of length w using correlation coefficient of their shapes. If at least one such comparison reveals similarity, the two melodies would be regarded as similar with degree $(w/s) \times 100\%$. An example is supplied in a raga.

Key words: Melody; similarity; correlation coefficient

1. Introduction

In our previous work (Chakraborty et. al. 2009a), we had used correlation coefficient to measure similarity between melodies of equal length in a North Indian raga. A raga is a melodic structure with fixed notes and a set of rules characterizing a certain mood conveyed by performance. The present note takes up the problem of finding similarity between melodies of unequal length. If two melodies are of length w and s where $w < s$ (say), our suggestion is to take all possible segments of length w from the longer melody of length s (if $s = w+k$, there will be $k+1$ such segments) and then compare each segment with the shorter melody of length w by finding the correlation coefficient of their shapes.

We provide some routine definitions first (definitions 1-9). Next we provide our main definition (definition 10) followed by an illustrative example in raga Bageshree.

Definition 1: A **melody** is a sequence of notes taken from a musical piece that is "complete" as determined by music theory. This means a melody should at least be a complete musical phrase if not a complete musical sentence.

Definition 2: A **segment** is a sequence of notes which is a subset of a melody but is itself incomplete.

For example, [Sa, Sa, Re, Re, Ga, Ga, Ma, Ma, Pa] is a melody in raga *Kafi* as it is complete while [Sa, Sa, Re, Re] and [Sa, Sa, Re, Re, Ga, Ga] are its segments.

Definition 3: **Length** of a melody or its segment refers to the number of notes in it.

For example, the lengths of the melody and its segments in the previous example are 9, 4 and 6.

Definition 4: **Significance** of a melody or its segment (in monophonic music such as Indian classical music) is defined as the product of the length of the melody and the number of times it occurs in the musical piece. Thus both frequency and length are important factors to assess the significance of a melody or its segment. The term has a more technical definition in polyphonic music (Adiloglu, Noll and Obermayer, 2006).

Melody significance has been studied in (Chakraborty et. al. 2009a).

Definition 5: **Shape** of a melody is the difference of successive pitches of the notes in it. See table 1 in section 2.

Definition 6: Two melodies are in **translation** if the correlation coefficient r of their shapes equals +1.

Definition 7: Two melodies are in **inversion** if the correlation coefficient of their shapes equals -1.

Definition 8: Two melodies are called **different** if the correlation coefficient r of their shapes approaches 0 in magnitude. Thus r here is a measure of **similarity** between melodies. We immediately have definition 9.

Definition 9: Two melodies are **similar** if the correlation coefficient r between their shapes is found to be significant. It should be remembered that correlation coefficient measures the degree of linear relationship between two variables (here melody shapes).

Remark: Since correlation coefficient is a number between -1 and +1, fractional values approaching -1 or +1 or even zero can be interpreted likewise based on the above mentioned definitions. Significance of a correlation coefficient can be tested using t test. If r is the value of correlation coefficient and n be the number of pairs of observations (here successive differences), we calculate the statistic $t = rv(n-2) / \sqrt{v(1-r^2)}$. If the absolute value of t , i.e. t exceeds the table value of t at 5% level of significance (say) and $(n-2)$ degrees of freedom, then the value of r is significant at 5% level otherwise insignificant. Here it is assumed that the n pairs are coming from a bivariate normal distribution. The

formula for r is covariance (x, y)/v{sd(x) sd(y)}where sd=standard deviation. Covariance (x,y) can be computed easily as $\text{Sum}(xy)/n - \text{mean}(x) * \text{mean}(y)$. $\text{sd}(x) = \sqrt{[\text{Sum}(x^2)/n - \text{mean}(x)^2]}$ and similarly for sd(y). $\text{Mean}(x) = \text{Sum}(x)/n$ And $\text{Mean}(y) = \text{Sum}(y)/n$.

We propose the following new definition for finding similarity between two melodies of unequal length.

Definition 10: Two melodies of length w and s where $w < s$ are similar if at least one segment of length w from the longer melody (of length s) is found similar with the shorter melody (of length w). The degree of such a similarity is $(w/s) \times 100\%$. The degree is trivially 100% for two similar melodies of equal length for which $w=s$.

2. An analytical illustration in raga Bageshree

We provide an illustration from raga Bageshree. See Table 1 for pitch calculations for structure analysis. Further information on raga Bageshree can be found in (Chakraborty et. al. 2009a). For performance analysis of ragas, see (Chakraborty et. al. 2009b).

Taking the tonic Sa at C conventionally, the twelve notes in the middle octave can be represented by the numbers 0 to 11 respectively (Adiloglu, Noll, Obermayer, 2006). Sa of the next higher octave will be assigned the number 12 etc while Ni (sudh) of the lower octave (before middle) is assigned the number -1 etc.

C	Db	D	Eb	E	F	F#	G	Ab	A	Bb	B	(lower octave)
S	r	R	g	G	M	m	P	d	D	n	N	
-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	
S	r	R	g	G	M	m	P	d	D	n	N	(middle octave)
0	1	2	3	4	5	6	7	8	9	10	11	
S	r	R	g	G	M	m	P	d	D	n	N	(higher octave)
12	13	14	15	16	17	18	19	20	21	22	23	

Table 1: Pitch calculations of musical notes for structure analysis

Abbreviation used in Table 1 :-

The letters S, R, G, M, P, D and N stand for Sa (always Sudh), Sudh Re, Sudh Ga, Sudh Ma, Pa (always Sudh), Sudh Dha and Sudh Ni respectively. The letters r, g, m, d, n represent Komal Re, Komal Ga, Tibra Ma, Komal Dha and Komal Ni respectively. A note in Normal type indicates that it belongs to middle octave; if in italics it is implied that the note belongs to the octave just lower than the middle octave while a bold type indicates it belongs to the octave just higher than the middle octave. Sa is the tonic in Indian music. "Sudh" means natural, "Komal" and "Tibra" means flat and sharp respectively. Corresponding western notations is also provided.

Consider now two melodies in raga *Bageshree* of unequal lengths [0, -2, -3, 0, -2, 0] and [-7, -3, -2, -3, 0]. From Table 1, it can be readily seen that these correspond to melodies [S, n, D, S, n, S] and [M, D, n, D, S]. Consider the first segment of length 5 from the longer melody [0, -2, -3, 0, -2]. Its shape is S1(1) = [-2, -1, 3, -2]. Comparing with the shorter melody [-7, -3, -2, -3, 0] also of length 5, whose shape is S2 = [4, 1, -1, 3], we find correlation coefficient $r = -0.915688$. For this r value, $t = 3.22$ which is insignificant at 5% level of significance but **significant at 10% level of significance**. Here $n = 4$. Table value of t at $4 - 2 = 2$ degrees of freedom at 5% and 10% levels are 4.30 and 2.92 respectively.

The second segment of length 5 from the longer melody is [-2, -3, 0, -2, 0] which has shape S1(2) = [-1, 3, -2, 2]. When compared with the shorter melody, it gives correlation coefficient $r = 0.221028$. For this r value, $t = 0.32$ which is insignificant both at 5% and 10% level of significance, table t values remaining the same as before.

It follows from these results, from definition 10, that *the two unequal melodies are similar at 10% level of significance as at least one segment from the longer melody is found similar to the shorter melody at this level*. The degree of similarity is $5/6 \times 100\% = 83.33\%$. Of course one should sing or play the melodies also to assess their similarity aesthetically. The objective quantified information is only a supplement to, and by no means a substitute of, the subjective aesthetic assessment. Fig. 1 gives the scatter plots of S1(1) and S2 coordinates. Fig. 2 gives the same for S1(2) and S2 coordinates.

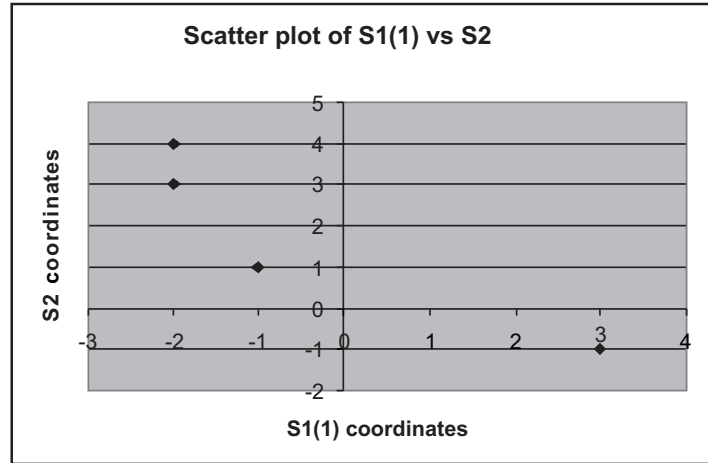


Fig. 1: Scatter plot of S1 (1) and S2 coordinates

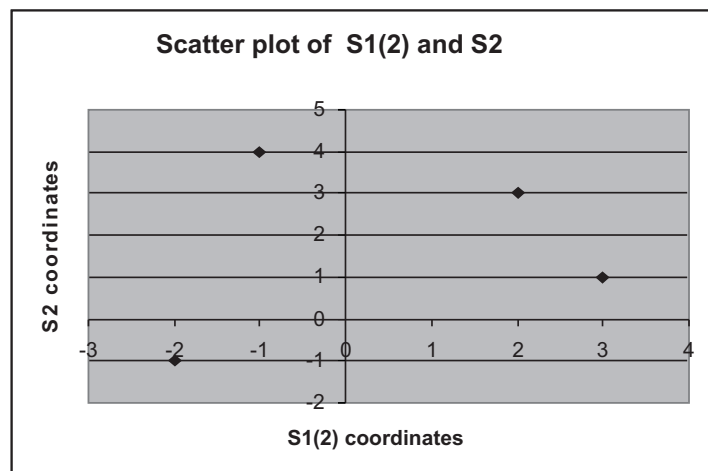


Fig. 2: Scatter plot of S1 (2) and S2 coordinates

3. Conclusion

We have proposed a method of studying the similarity of melodies of unequal lengths by generating segments from longer melody of length equal to that of the shorter melody and calculating correlation coefficient of their shapes and testing its significance. An important finding is that the level of significance at which they are found to be similar, if similarity is detected, is crucial. The p-value (minimum level of significance at which a test statistic, in this case the t-statistic, becomes significant) can be used profitably. Although our illustrative example does not state the result in the form of p-value, there is enough

indication to suggest it. Since we are able to calculate the % of similarity as well, we argue that the strategy can be useful in distinguishing “inspiration” from “theft” in music adaptation.

References

1. K. Adiloglu, T. Noll and K. Obermayer, *A Paradigmatic Approach to Extract the melodic Structure of a Musical Piece*, Journal of New Music Research, 2006, Vol. 35, no. 3, 221-236
2. S. Chakraborty, R. K. Shukla, M. Kumari, K. Krishnapryia, Loveleen, S. Chauhan, S. S. Solanki and K. Mahto, *A Statistical Analysis of Raga Bageshree*, Proceedings of the International Symposium on Frontiers of Research in Speech and Music (FRSM) 2009 held on Dec 15-16 at ABV IITM in Gwalior, India, 148-157 [a]
3. S. Chakraborty, R. Ranganayakulu, S. Chauhan, S. S. Solanki and K. Mahto, *A Statistical Analysis of Raga Ahir Bhairav*, Journal of Music and Meaning, Vol. 8, sec. 4, Winter 2009 <http://www.musicandmeaning.net/issues/showArticle.php?artID =8.4>

TECHNE: REVEALING SOUND, SPACE AND SELF IN MUSIC FOR FLUTE AND ELECTRONICS

Jean Penny
Independent Researcher
5 Owen St, Kallista
Victoria, 3791, Australia
mjpenny@bigpond.net.au

Abstract

The expansion of electroacoustic music over the last half century has changed the flautist's sonic and performative environment. New equipment and systems, new technical demands, new synthesis and interpretative strands have evolved, generating new questions and responses to performance. In this setting, the traditional idea of the flautist has transformed into a meta-instrument entity: a collaborative symbiosis of instrumentalist, technologist, hardware, software, virtual and real performance space, and sound. There is a radical shift in performance experience when playing in this multi-dimensional performance zone. This paper traces an historical progression of flute sonority and extended techniques, and the development of electronic spatialisation techniques and their influence on flute performance. These influences include the extended sonic capacities of the flute, new physical responses in the performer, a renewed sense of performance space, the synergy of newly nuanced interconnections and relationships, and a transforming sense of musical identity.

The flute sound extended

The flute's resonance and timbral qualities have attracted composers and performers to explore ways of capturing and expanding its translucent sonic qualities into evocative expression. This sonic ethos stretches across the centuries, from the architectural structures of Baroque works, such as J. S. Bach's *Partita in A minor*, to more recent works of composers such as Toru Takemitsu and Kaija Saariaho in which the resonance and malleability of the tone and breath are a critical element of composition. The influence of composers such as Varèse and Berio on flute playing and the subsequent changes to perception of the flute's sonic capabilities stimulated an expansion in acoustic techniques, including incorporation of techniques from antiquity and world music, that paved the way for diverse electronic interventions. The convergence of these paths has created a genre that exploits and explores resonance within the

music's compositional structure, within the expanded acoustic sounds of the flute and through electronic transformation.

Marcel Moyse (1889 - 1984) taught the modern flute world to listen in extreme detail to nuance and colour, to achieve homogeneity of tone throughout the range, and to project tonal variety and individuality into performances. Moyse, and his French predecessors, including Paul Taffanel and Philippe Gaubert, are largely credited with a resurgence of interest in playing and composing for the flute in the twentieth century, for establishing a rigorous approach to technique and sound, for teaching and for publishing teaching and practice material. It was in this context that Edgard Varèse's seminal composition for solo flute, *Density 21.5* (1936), made its appearance. This work took these beautiful resonant sounds and extended, challenged and redirected them towards new tonal expression through the use of key clicks, dynamics as sound manipulation and extremes of tessitura. As practitioners began to explore and extend these new paths, texts began to appear examining such techniques as microtones, whistle-tones, breath tones, tongue clicks, key slaps, multiphonics, harmonics, glissandi, new embouchure techniques, vocalization and varieties of articulation (for example, the texts of Bruno Bartolozzi [1], Pierre-Yves Artaud [2], and Robert Dick [3]).

Extended flute techniques became the source material of composers, scrutinized for their sonic properties and expanded out into the musical milieu through amplification and spatial projection, provoking the evolution of expanded capacities and creative potential in a genre driven by invention and the re-positioning of sound and instrumentalist. This renewed approach continues to impel performers to analyse their own practice, to develop new capacities and processes in their work, to incorporate new technologies and techniques and to embrace new performance situations.

Activating space: Sound, technology and performance

From the echo-vases of Vitruvius [4] in the first century AD, to the expanse of cyberspace today, performers have sought to transform the 'empty space' (the theatre, concert stage, outdoor or virtual arena) into a vibrant performance space, to engage the space as collaborator and enhancer, to draw in the audience and empower the performance. This space, the performer's habitat, is influenced by many elements of historical context. The location of music (buildings, outdoors), the forms of presentation (formal, intimate), the position of the audience (separate, amongst) all have an impact on the performer, often

This article is closely based on Chapter 3 of the author's DMA dissertation: The Extended Flautist: Techniques, Technologies and Performer Perceptions in the Performance of Music for Flute and Electronics (2009).

stimulating and sometimes challenging the quest to project and synthesize. Spaces may be integrated, manipulated, and developed through a blend of virtual (composed) and real (performance) elements, new flexibilities and illusions, intersections and separations.

The influence of spatial acoustics has long been recognized as part of the flautist's tools of trade: the eighteenth-century flautist / composer, Johann Joachim Quantz, for example, advises the flautist on resonant projection of tone in public performances [5], and developments in construction, such as the Boehm flute of 1849, [6] and more recent refinements, have chiefly been generated by the quest for increased tonal capacities. The venues where these flutes are heard, the reverberation characteristics, the diffraction and absorption of materials within them, create acoustics that are unfixed and unpredictable to a degree, but remain largely static, single dimension spaces. In the contemporary arena, new modes of spatial representations produced from technological means have changed the listening space to something more malleable and adjustable, subject to the forces of imagination and electricity; transformations dependent on the effective combination of specific designs and sonic properties.

As the microphone developed as a musical tool, the construction of virtual acoustic space through loudspeaker placements and diffusion technologies became possible. This development evolved through the pioneering work of early generations of engineers and composers [7], and the important influence of John Cage's amplification explorations. Douglas Kahn describes silence as "Cage's emblem" [8], but this silence incorporated every audible, potentially audible, and mythically audible sound. New ideas and attitudes in music that generated a sweeping shift in musical thinking were consequently opened up. New ways of listening, and new sonic expectations emerged as composers explored the full spectrum of sound as music, stimulated new performance techniques and began to reconfigure acoustic spaces.

The Phillips Pavillion (designed by Le Corbusier with Xenakis) (1958), Clozier and Barrière's Gmebaphone (1973), Francois Bayle's Acousmonium, a loudspeaker orchestra (1974), Annae Lockwood's SoundBall, a flying loudspeaker with built-in amplification, 6 loudspeakers, and an antenna (1984), Denis Smalley's multidirectional sound environments and Jonty Harrison's BEAST, a moveable construction of loudspeakers: The Birmingham ElectroAcoustic Sound Theatre, are just a few of the manifestations of spatial music projection and research occurring through the later part of the twentieth century. Technologies such as these inspired a focus towards the systems and processes, the possibilities to change sound timbre and behaviour, and the study of the microcosms of musical elements and perceptual connections.

Stockhausen realized his spatialisation ideas at the Osaka World Fair in 1970. His earlier work, *Kontakte* (1959/60 version), uses spatial elements with striking innovation. This work explores connections between live instrumental music (piano and percussion) and electronic music. The instrumentalists represent immobile sound sources, and the electronics represent continually changing sources through rotation, looping movement, alternation, fixed and isolated space points and movements. Hearing *Kontakte* live is an astounding and absorbing experience, as performance gestures as well as the electric and acoustic spatialisation reveal the 'contacts' or relationships between the various elements of this work: the performers, instruments, sound sources and the sounds themselves. Stockhausen's later work *Mikrophonie* (1964), for 6 players (2 percussionists, 2 'microphonists' and 2 musicians working the technology filters with potentiometers for volume, timbre, pitch and rhythm controls) [9] creates an electronically activated spatial sensation, where the sounds of a tam-tam brushed with 'household goods' are amplified by the two microphones moved in rhythms towards and away from the tam-tam. This work is among the first electronic performance works. Thirty years later, Italian composer Marco Stroppa's explorations of spatial contexts and techniques include the space within sounds as well as physical and virtual space through sound structures of, for example, double bass and metallic percussion amplification up to a billion times of the recorded start of a sound, as used in *Zwielicht* (1994-99) [10].

Pauline Oliveros described virtual acoustic space as “a perceptual phenomenon ... created with electronic processing within an actual physical space” ([11] p. 19). This space creation uses amplification, diffusion, sound movement and timbre manipulation techniques to create multiple effects or distortions: including expansion and contraction of the fixed acoustic space, alteration of tonal quality, proximity illusions, magnification of sonic intensity and to set up new relationships, dramatic discourse and immersive or remote sensations. Marshall McLuhan describes auditory space as 'without fixed boundaries . . . always in flux' [12], emphasizing the organic, contingent qualities of electronically driven auditory experience. The virtual spaces of sound art found in works of such composers as Francis Dhomont (for example, *Sous le regard d'un soleil noir* (1979-81)) and Hildegard Westerkamp (for example, *Into India* (1997-2002)) create vibrant and dramatic use of spatialisation technology that often depicts a defined but pliable space. These spaces are equally powerful elements in performance-based works, where the dynamics of the instrumentalist within the space become important elements of expression.

Mixed virtual and real elements create a powerful performative situation: the field of play in which the music can occur, where perceptions are

challenged and expectations adjusted. Often the technology will reveal layers and aspects of the music otherwise hidden such as spaces within sounds and contexts or act as a mask, only revealing certain facets of the sound or performer, or positioning the performer forward or back. Sounds can be completely out of proportion to one's expectations, as in Gordon Mumma's *Hornpipe* (1967), in which impossible breathing patterns are created through electronic illusions, and the human appears transformed [13]. Reverberation strongly influences spatial listening perceptions, giving a sound a context and positioning; harmonization effects add layers and relationships between voices, and moving sound cues can encircle and submerge, or dislocate and disassociate.

Through the expansion of instrumental sound via amplification two major performative outcomes are instantly apparent: an expanded sound creates a set of new freedoms and powers, and microsounds become audible and thus viable musical elements. Responses from creators and performers can be intense: Westerkamp, in her speech *Speaking from inside the soundscape* describes the microphone as a seductive tool that gives an impression of close contact but is in reality a separation, an illusion [14]. Flautist, Anne La Berge states that “amplification has allowed me to be heard in many more situations than before. It has also allowed me to use some very special, wonderful, intimate sounds in settings where they would be inaudible if not amplified.” [15]

This hyper-magnification also creates an expanded capacity to draw the listener into the sound world of the performer. The soundscape is greatly expanded, and listening expectations alter: the flute sound can be stunningly large, or coloured by minute breath or percussion sounds, expanded out into the room; the sounds can be processed to barely resemble a flute at all, or collect layers or mirroring reverberations to increase density or texture; they can dart from speaker to speaker, create an immersion, a disconnection or illusions of intimacy or distance through panning or delay. These expansions and contractions invoke changed expectations and impose new meanings.

The diffusion of sound through speakers to articulate or create a spatial musical experience can range from stereo (usually two speakers facing the audience) to spatial orchestra (many speakers) and the use of ambisonic techniques. The technologies can allow the creation of auditory space as a set of flexible environments, defining the sound world through creating precise virtual spaces, with the configuration of speakers or computer generated outputs enabling, for example, geometric sonic patterns, circles of immersion, or circles of exclusion with the audience observing from a distance. Strategies for diffusion underline this multiplicitousness: sound placement and movement, distance, volume, spectrum, intensity, clarity and relationships

between sounds.

The techniques underline the significance of the manipulated or moving sounds, give clarity to the projection of the structure and layers of the composition, and create an interface that may suggest musical ideas and representations such as ambiguity, disembodiment and illusions of dialogue or presence. Denis Smalley describes spatial movement as unidirectional, bi-directional, reciprocal, centric/cyclic and multidirectional, and motion styles as synchrony and asynchrony, continuity and discontinuity, conjunction and disjunction, periodic and aperiodic (cited in [16]). He terms the dynamics and interchanges of composed space and listening space spatiomorphology. Smalley's spectromorphology concerns the shaping of sound partials and frequencies that imply spatial experience; spatiomorphology explores spatial properties and spatial change [17]. Spatial illusion, allusion and implied spatial occupation is further discussed by composer Natasha Barrett [18]. Approaching spatial issues in music as either intrinsic (the spectrum, its morphology and structural organisation) or extrinsic (the sound's capacity to imply, refer or associate), she stresses the importance of multiple elements in the implication of place, characteristics of a performance space, sound movement and reverberation. Composers across the spectrum have delved into the potential applications of these fields, of the melding of techniques and richness of musical outcomes.

New spaces – New balances

The artistic application of these developments has created new dimensions for performative discourse, expression and research. Altered sonic environments such as immersion or diffraction, illusion or magnification have become a vibrant element of composition, influencing the location of performer and focus of performance. The shifting relationships of the performer to the space, to the sound, and to the audience become part of the presentation, elements that are vital to expression.

Conventional concert configuration of the stage facing an audience in straight lines can suggest a formal separation. A more interactive and intimate set up can be achieved with the audience placed closer to the performer, more in the style of the ancient Greek's cosmic circle [19]. The inner and outer fields of live electroacoustic performance have been discussed by Simon Emmerson [20], describing the balance of the instrumentalist's control (or local control) with control of an environment (or field). These fields are linked through technological transformations, which can blur the definitions of space as well as create it. If the flautist is the inner (or local) field, this would encompass the physical body of the performer, the sound source, the launch of musical

explorations and interpretation, perceptions and creative responses, the visual focus in other words, the performance centre. Sound transformations created via the sound technician's operation of electronic equipment, and technological synthesis, project the music into an outer field, or audience, where evocations are received, experiences shared and responses invoked.

That the dynamics between live performer and recorded or manipulated sound have become an important expressive field is reflected in significant contemporary writing. Katharine Norman's discourse about Luigi Ceccarelli's *Birds*, for bass clarinet and tape, describes the different sense of location and emanation produced by a live instrument (physical) and electronic sound (invisible), and the creation of a sonic universe for the performer on the stage. In the case of this work, the clarinet sounds are extended, but essentially remain clarinet sounds, blurring the distinctions through placement, and creating a sense of space which encompasses up, down, and 'aboveness' all at once [21]. Marco Stroppa describes the relationship between instrumentalist and electronics as a "dialectic relationship, where each realm remains what it is, yet, interacts with the others . . ." [22]. Several of his works, for example *Traiettorie...deviate* (for piano and electronics), *Auras* (for percussion and electronics) and *little i* (for flutes and electronic room) explore this concept through spatial configurations on stage, amplification and loudspeaker placements. The influences of the projection of the instrument and the electronics on each other, and the quest for symbiosis or morphology are integral to his compositional approach in these works (Ibid).

Through varying balances and emphases, the focus of performance can be diffracted, creating situations where the responses and permissions granted to audiences may change. Simon Emmerson discusses the importance of electroacoustic music in reviewing the quality and concentration of listening, and the revitalizing of performance spaces and performance practice. The standard concert space, he states, is challenged to allow for flexibility of media and audience intersection [23]. In less traditionally structured situations, different permissions can be granted in environments without specific audience placement, where the performers may be scattered, and the audience invited to move amongst them. In these situations, blurring of the position of the performer, whether conceived as foreground or background, can occur as the balance between sonic, spatial and visual elements intersect. In formal settings the hardware required can additionally set up distinct barriers that emphasize separateness. Often one sees a conglomeration of black boxes, wires, and devices spread all over the stage area, creating a mini obstacle course through which the players carefully step and play. The visual aesthetics strongly influence the dramatic style of the performer and synaesthetic responses of an audience, and the separations that occur are equally forceful.

Flute compositions: Extensions, expansions and integration

Spatialised sound manipulation technologies in extant flute works include amplification, delay, filters, panning, reverberation, multi tracking and digital signal processing (modifying digital representations of audio material, both live and through sound files, samples or recorded sounds). Source material may be live instrument, pre-recorded or computer generated sound, and real-time control over sound levels, equalization and placement of sound can occur. Multiple sonic choices and altered expectations generated by amplification, in particular, have led on to new freedoms, new sound discoveries and new understandings of performance and creative solutions.

The desire for an expanded flute sonority and the simultaneous development of applicable electronic procedures have incited a new set of sonic goals, far removed from the traditional resonance of the flute. Microsounds are a major part of this approach to sound, invoking new expression and an altered sense of sonic identity in the player. Multiphonics create depth and colour, strands of voices and even instability; microtones blur conventional note distinctions and scalar structures; glissandi smudge and meld, giving a new sense of placement and movement of flute tone. Works that exploit these techniques abound, for example: Mary Finsterer's *Ether* for amplified solo flute (1998), written almost entirely for amplified whistle tones; Kaija Saariaho's works (for example *NoaNoa* (1992)) amplify minutiae including whispers and voice sounds; Drake Mabry's *5.4.88* (1988), for amplified C flute uses muffled micro-sound throughout in a percussive soundscape; Nicola Sani's *I binari del tempo* (1998) for flute and tape uses electronic sounds sourced from different flutes, transformed by digital sound processing to create an expanded world of flute and air sounds. Other works use spatialisation to compliment musical ideas inherent in the composition, for example, Andrew Ford's piccolo piece, *...les debris d'un reve* (1992), which uses massive reverb and its removable, and Chris Dench's *Caught Breath of Time* (1981/2004). The electronic version of Dench's work pans the sound from the flautist towards the audience, past and back to the performer again as a radiant metaphor for the passing of time.

Examples of the broader application of electronic techniques expanding the territory of flute performance include Karlheinz Stockhausen's *Solo* (1965), and Brian Ferneyhough's *Mnemosyne* (1986), which explore the spatial relationships between the soloist and prerecorded flute. Stockhausen's *Solo* for instrument and multi-channel tape delay creates spatialisation effects through sensations of sound movement, by recording the flute at various delay intervals, and playing these back with the live flute. *Mnemosyne*, the last part of *Carceri d'Invenzione*, for bass flute and pre-recorded 8-track tape, creates chordal patterns which unfold as a backdrop of pitches surrounding the flautist who

plays material derived from these. The multiplication of bass lines in the tape part from 4 to 8 creates the space for the soloist, magnifying the sense of the bass flute within progressively decreasing and 'imprisoning' material [24]. Pierre Boulez' *Dialogue de l'ombre double* (version for flute, 2002) was originally written for clarinet in 1985. The concept of a double shadow is explored, with live instrumental passages moving against an electronically spatialised double. Spatial relationships between the soloist and the shadow sound of the prerecorded flute, often the alto flute, are developed.

Real time spatialisation of sound occurs in numerous works, including Laurie Radford's *A Florus Exchange* for flute and DSP (2004), where spatial movement of sound is teamed with radical timbre alteration. The 'exchange' is manifest in the treatment of the musical material of the flute part by the computer (through Max/MSP), and the consequent responses and dialogue between both performers. Predetermined spatialisation of the sound through the Max patches creates a spread of sound, and specific movements from speaker to speaker. The solo flute line itself provokes multiple opportunities for timbral variation by the instrumentalist, and these are augmented to new dimensions through digital signal processing techniques: Delay Pitch Shift (flute signal is delayed and shifted in frequency), Pitch Shift/RM (flute signal is doubled by two harmonizing voices and then ring modulated), Multilooper (audio is captured/recorded and then looped; controls for loop direction, length, frequency) and Granular Synthesis (audio is captured/recorded and then granulised with controls for various parameters). The sound effects follow the flute sound, changing and developing it, instantly or through improvisation, commenting and manipulating.

Important developments have been influenced by the work of flautists in the field, as significant players have collaborated with composers and sound technicians to develop ideas and resolve problems. The collaboration between Roberto Fabbriciani and Luigi Nono, for example, is well known. *Walking with Gigi* [25] describes the explorations of new flute sounds and electronic treatments, giving a vibrant picture of this important relationship.

Nono would ask me to realise extraordinary sounds on the very edge of the audible/in audible. Within these sonorities he would then look for stimuli to new ways of thinking and listening . . . The executant proposed and the composer chose . . . His aim was in fact to produce a more conscious listening. . . and to generate strong emotions against any established, traditional form . . . Technically, the live electronics consisted of only a few sound treatments . . . including: amplification, spatial projection, delay, pitch shift (harmoniser), filtering and mixing. The novelty of being able to take advantage of these techniques in real time, generated new ideas and opened the way to numerous

innovations.

Transformative influences: A performer's view

The transformative elements of performance with electronics, the influence on the listener, and the role of familiarity in reception and perception of this music are responses that underline the performer's world. Overlapping layers of instrumental performance and extended practice evolve into a complex, yet compelling search for interpretative cohesion. Observation of these transitions adopts a hierarchy based on personal priorities and musical aesthetic, summarized here.

Sonority

The new sonic possibilities created by the combination of flute with electronic spatialisation and manipulation generate a re-evaluation of soundscape. A great diversity of expression is achieved as sounds become imbued with new meanings, change and meld into interpretative forms, are separated from the player, treated as sonic units, dislocated within the space and from past associations. The power of the player is magnified through a projection of sounds impossible at the acoustic level, giving a hugely expanded range of expression choices to the player. Amplification of very soft, intimate sounds can magnify effects or change emphases; amplified breath sounds can create a breathless, tense and dissonant sensation in the listener. Vivid effects can be created through amplified multiphonics, often unstable and swinging, evoking uncertainty in player and listener, at other times assertive and declamatory. Amplified whistle tones suggest distance and ethereality, often used to depict a distant character or thought. Changes in vibrato intensity and speed can give shimmering colour variation with amplification, especially in combination with reverberation. Combined flute tone and voice or breath can introduce a grainy, indistinct tone that is quite malleable with amplification, and varied emphasis of these colours can have very differing effects. Amplified microtones and overtones can be mixed together effectively to distort pitch; and magnified percussive sounds bring completely contrasting sonic worlds into play with, for example, sharp, metallic key slaps or muffled articulations. Techniques such as tongue rams (which sound lower harmonics) expand the sonic range of the flute, and can disorient a listener, especially if the sound is given a broadened spatiality.

Physicality

The emergence of performance identity through gesture and shaping of musical material is founded on physical actions (posture, breathing, throat and mouth

shaping, arm and hand movement, emotionally reflective body language) learnt over many years of practice. This physicality of the performer in the space is significantly impacted by electronics: new techniques have emerged, such as microphone and pedal triggering; new playing emphases are demanded, such as refined embouchure adjustments or variations of tone and intensity; new breath techniques and muscular demands with new finger, arm and hand movements are frequently required; constraints may be created by equipment set up and the demands of controlling and managing computer and hardware, and a sense of enclosure or openness in the space. Physical aspects of aural adjustments are equally critical, with the hearing ability of the performer in a re-configured sound space subject to challenging alterations. A detachment from reality can occur, with blowing and muscular movement opposing electronic forces.

Performance space

The distinctive properties of acoustic and virtual spaces generate new rules of sonic behaviour and response. The virtual space becomes a player in the performance, controlled in the main by the technologist. It is an element of expression and communication, changing the relationships and positioning of performer and sound. Sound sources can become entirely mixed up through live and pre-recorded sound sources, creating ambiguity, the sense of invisible presence and unexpected interconnections. A sense of enclosure, or disorientation, uncertain proximity and passing of time can result, or dialogues with spatialised voices. The instrumentalist can be magnified or obscured, empowered or reduced, thus altering the listener's perception of the performer and the performer's own relationship with the space.

Immersive diffusion techniques can draw both listener and player into the space, heightening the sense of participation, and intimacy. The 5.1 or 7.1 surround sound systems of cinematic origin, and virtual environment simulations with real-time multichannel spatialised sound processing that create virtual reality are well established, recognizable listening conditions. Spatial environments such as these create specific reactions in listeners: to be immersed is to become completely involved; the performer feels a sense of dissolve in the music, a merging of person with sound, a virtual acquiescence [21]. The intimacy of some techniques such as whispering can enhance this effect, with a sense of closeness and inner connection. The 'feeling' in the room can be manipulated through this expansion and contraction, immersion and emergence, the discreet and the extrovert. Electronic reverberation can add to a sense of immersion, through allowing the build up of textures, or sense of enlargement of the tonal space, the contrapuntal or harmonic layers giving a sense of support to the player, and dialogue with the musical material or

implied hidden persona. This situation is a feature in much of Saariaho's flute music, where the influence of reverberation on perception of space and placement is vibrantly demonstrated.

Interactions

In the context of spatialised flute performance, interactions occur between the sound technologist and flautist through an exchange of responses, such as software processing, recording and play back, amplification, diffusion, and techniques such as delay and timbre manipulation. This performance interaction has developed in a somewhat *ad hoc* manner, with changing roles, changing demands and input balancing working around the needs of individual performances and the vagaries of electronic equipment. The balance, projection, and coherence of the instrumental sound are frequently in the hands of the technologist, as the flautist is unable to judge the full effect from the playing position because of the placement of speakers behind, to the side and in front. It becomes a shared production, where the aural judgment and ensemble decisions depend on trust and common artistic goals. The human direction for successful balance and cohesion is a search for a fruitful alliance with machinery that is not necessarily built with a particular performance in mind. At times it can be challenging, as a performer, to 'stay in the moment' when working with technology functionality, the wider activity schema often diverting attention from pure musical concerns.

Identity

The performer's identity is a combination of multiple interior and exterior factors, and is strongly bound up with being a particular type of musician, making specific musical sounds. The outward identity (style, sound, interface, instrument) and inward identity (self awareness, interpretation, transformative potentials, musical background) blend to form the performing persona. The sonic self "comes into being with sound", Naomi Cumming states, "and the tone is central to creating the impression of musical personality" [26]. As an extension of the self, the instrument acts as an amplifier [27] displaying inward and outward identity through musical performance, revealing aspects of inner processes and thoughts, responses and emotional connections through interpretational signals. This self-projection creates a perception in both player and listener, imprinting both visual and aural impressions on a listener, and a sense of positioning and potential to the player.

Jean-Claude Risset has written: 'What matters in music is not intrinsic complexity, but the complexity we perceive, complexity in us' [28]. This includes the perception of influences on the self, generated by spatialisation

techniques. Simply enlarging the sound gives an immediate sense of power of projection and flexibility to the flute player; adding effects such as reverberation or delay alters the location of self and ensemble symbiosis; exposure of small body sounds out into the hall can give a sense of self discovery and also unveiling. Crucially, these influences provide a vastly increased range of musical elements to activate, thus enlarging the personality in sound. "In the process of changing my relationship to space, I discover a new possibility of 'self': . . . a new construction of my embodied position and relative dominance . . ." [26].

Taina Riikonen argues that reverberation affects the embodied identities of flautists, sets up ambiguities, creates new power relationships and experiences of self, as identities implied through voice, of live and recorded flautists, become difficult to distinguish [29]. The use of voice with flute introduces a vibrant strand: the personal, human perspective of the performer is projected directly, and often surprisingly, through these techniques. The impact of melding voice or whispering, or of using the voice as narrative in disembodied projection immediately changes the expression: it becomes a recognizable human, located within, or distant from the player. The position of the player immediately becomes subject to distance perception, to reality challenges and identity questioning.

The melding of techniques, an enhanced freedom of expression, sonic empowerment and flexibility creates an expanded sense of identity and energy source in the player expressed through gestures, breath, and interpretational vibrancy. There is a new sense of ensemble, and at times a spatialisation of solo playing into layers and illusions of other presences, a spatialisation of the self. There is a magnification of the micro, intimate elements of playing, which can move the audience into the performer's sound space, the proximity of the sound source, and the possibility to mask and confuse with links to the known and unknown sound source. The new sound worlds and new techniques become elements integral to the self, part of one's own 'vocabulary', a representation of personality in sound. In Stockhausen's words: "New means change the method; new methods change the experience; and new experiences change man" [30].

Explorations and intentions

This paper has traced developmental strands of contemporary flute techniques, music technology applications, performer and composer contributions and musicology texts. As the microscope is taken to the experience of performance, a disclosure of inner perspectives and clarification of minutiae occurs. The meta-instrument emerges, reflective of an expanded

identity, an ecology of sounds, processes, machines and people.

The symbiosis of the meta-instrument creates intense enrichment in my personal performance practice. I relish the opportunity to explore unusual sounds and performative responses. Invisible presences and dramaturgy elements enhance and deepen my performance as new partnerships with sound transpire and new works are explored. New procedures and adapted techniques provide an exciting, opulent artistic setting from which to encounter and realize music as performance.

These elements feed this flautist/researcher's curiosity, and impel the construction of multiple settings from which to further investigate the nexus of instrument, performer and digital technology.

References

1. Bartolozzi B, "New Sounds for Woodwind" (2nd ed.), Oxford University Press, London, 1967/1982.
2. Artaud P, "Present Day Flutes", Billaudot, Paris, 1995.
3. Dick R, "The Other Flute", Multiple Breath Music Company, St Louis, 1989.
4. Godman R, "The enigma of vitruvian resonating vases and the relevance of the concept for today", eContact! 7(2), 2004. Retrieved 22 March 2009, from http://cec.concordia.ca/econtact/Harvest_Moon/files/Vitruvius_Godman.pdf
5. Quantz J J, "On Playing the Flute", (E. Reilly, Trans.), Faber and Faber, London, 1966.
6. Powell A, "The Flute", Yale University Press, New Haven, 2002.
7. Emmerson S, "Living Electronic Music", Ashgate, Aldershot, 2007.
8. Kahn D, (2001), "Noise, Water, Meat", MIT Press, Cambridge, MA, 2001.
9. Moritz A, "Mikrophonie I. In Stockhausen: Essays on the Works" 2007, Retrieved 25 October 2008, from <http://home.earthlink.net/~almoritz/mikrophonie1.htm>
10. Zlender M, "Marco Stroppa: A visionary of sound architecture", The Living Music Foundation, Inc., 2001. Retrieved 22 August 2004, from <http://www.dwightwinenger.net/journalx3b.htm>
11. Oliveros P, "Acoustic and virtual space as a dynamic element of music", Leonardo Music Journal 5, 19-22, MIT Press, Cambridge, MA, 1995.
12. Gow G, "Spatial metaphor in the work of Marshall McLuhan".

- Canadian Journal of Communication, 26(4). 2001. Retrieved 16 March 2009, from <http://www.cjc-online.ca/index.php/journal/article/view/Article/1254/1251>
13. Reynolds R, "Mind Models: New Forms of Musical Experience" (2nd ed.), Routledge, New York, 2005.
 14. Westerkamp H, "Speaking from inside the soundscape". Paper presented at Hûr Upp! Stockholm Hey Listen! Conference on Acoustic Ecology, June 813, 1998. Retrieved 7 May 2007, from <http://www.sfu.ca/~westerka/writings/speakingsound.html>
 15. Penny J, "The Extended Flautist: Techniques, Technologies and Performance Perceptions in the Performance of music for Flute and Electronics", DMA thesis, 2009. Available online at http://www.griffith.edu.au/__data/assets/pdf_file/0003/184782/penny_the_extended_flautist.pdf
 16. Landy L, "Understanding the Art of Sound Organisation", MIT Press Cambridge, MA, 2007.
 17. Smalley D, "Spectromorphology Explaining sound shapes". Organised Sound, 2(2), 107-126, Cambridge, 1997.
 18. Barrett N, "Spatio-musical composition strategies", Organised Sound, 7(3), 313-323. Cambridge, 2002.
 19. Wiles D, "A Short History of Western Performance Space", Cambridge University Press, Cambridge, MA, 2003
 20. Emmerson S, "Local/field: Towards a typology of live electronics music". Journal of Electroacoustic Music 9, 10-13, Sonic Arts Network, London, 1996.
 21. Norman K, "Sounding Art: Eight Literary Excursions Through Electronic Music", Ashgate, Aldershot, UK, 2004.
 22. Trigueros F C & Tascon R, (n.d.). An interview with Marco Stroppa. Composer's Notebook Quarterly. [Online]. Retrieved on 1 November 2008, from http://web.mac.com/cnqonline/CNQ/CNQ_Interview_stroppa.html
 23. Emmerson S, "New spaces/ new places: A sound house for the performance of electroacoustic music and sonic art". Organised Sound, 6(2), 103-105, Cambridge, 2001.
 24. Toop R, "Interview with Richard Toop" in Boros and Toop (Eds.), "Brian Ferneyhough: Collected Writings (Contemporary Music studies)", Harwood Academic Publishers, The Netherlands, 1995.
 25. Fabbriani R, "Walking with Gigi", Contemporary Music Review, 18(1),

- 7-15, 1999. Retrieved 7 January 2008, from <http://www.informaworld.com/smpp/content~db=all?content=10.1080/07494469900640031>
26. Cumming N, 'The Sonic Self: Musical Subjectivity and Signification', Indiana University Press, Bloomington, 2000.
 27. Emmerson S (Ed.), "Music, Electronic Media and Culture", Ashgate, Aldershot, UK, 2000
 28. Risset J C, "The perception of musical sound. Computer music: Why?" France-University of Texas Institute for Interdisciplinary Studies. University of Texas at Austin, 2003. Retrieved 27 January 2008, from <http://www.utexas.edu/cola/insts/france-ut/archives/risset.pdf>
 29. Riikonen T, "Shaken or stirred virtual reverberation spaces and transformative gender identities in Kaija Saariaho's NoaNoa (1992) for flute and electronics", *Organised Sound*, 8(1), 109-115, Cambridge, 2003.
 30. Ball M, "Licht aus Stockhausen", *Avant Magazine* 5, December 1997. Retrieved 15 March 2009, from http://www.stockhausen.org/licht_by_malcolm_ball.html

STUDY OF SOURCE CHARACTERISTICS IN SAROD FROM THE SOUND SIGNALS

R. Sengupta^{1,2}, N. Dey², A. K. Datta^{3,2}

¹Sir C V Raman Centre for Physics and Music, Jadavpur University
Kolkata 700 032, India

²Scientific Research Department, ITC Sangeet Research Academy
1, N S C Bose Road, Tollygunge, Kolkata 700 040, India

³BOM Charitable Trust, Kolkata 700 035, India

Abstract

Sarod is one of the most important traditional Indian string musical instruments used for playing classical music. The large number of strings, including drone strings, sympathetic strings and the mode of attachment as well as the peculiar structure of bridge attached to a pliable membrane makes modeling the production mechanism using usual mathematical tools extremely complex. Moreover the system appears to be a non-linear complex one.

Fractal dimensional analysis is an effective tool in understanding the nature of the produced sound in such cases. In the present paper D_0 (fractal dimension), D_2 (correlation dimension) and also the higher moments are studied for acoustic signals obtained from actual performances. Four ragas performed by eminent sarodists are used. The obeying of the power law indicates the non-linearity. The presence of multifractality is noticed.

Introduction

Sarod is perhaps one of the most popular string instruments in Indian classical music. This instrument evolved out of Rabab played in Afganistan as early as the seventeenth century. The modern sarod uses about 23 strings made of steel and bronze. The sarod body is hollowed out of Toon wood or Teak wood, to form an almost hemispherical open tumba at the lower end, distinctly unlike that of sitar and smaller in size. The circular opening of tumba is covered with stretched goat skin, secured at the sides with glue (figure 1). The bridge rests on this. Eight strings, that are played, go through holes in this bridge, to be clamped under a metal piece, secured under the tumba. There are also three chikari strings and fifteen taraf strings [1]. There is no fret and the finger-board is of metal. The mode of attachment of strings and interaction between themselves create the acoustic environment, which includes the global resonances of the instrument. There are, however, regular predictable natures of complexity variation, which may have functional values other than providing a rich melodious sound. The complexity of the sound thus created may be modeled using fractal dimensional approach.

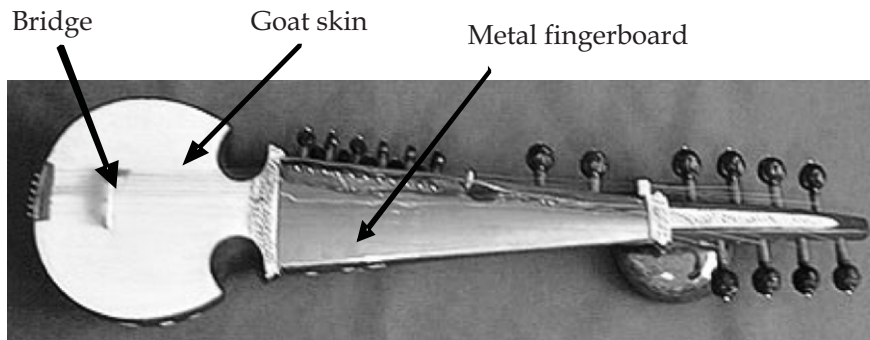


Figure 1. Picture of a Sarod

The fractal dimension is an effective tool for analysis in such cases. Fractal dimensions of time series generally reveal the presence of non-linearity in the production mechanism.

There are, however, regular predictable natures of complexity variation, which may have functional values other than providing a rich melodious sound. The source of origin seems to be related to a large number of non-linear systems associated with the strings, bridge and the resonant chamber. Fractal dimensions of time series generally reveal the presence of non-linearity in the production mechanism. Time series data is a quantitative record of variations of a particular quality over a period of time. One way of analysing it is to look for the geometric features to categorise data in terms of concept [2]. Study of fractal dimensions might be a technique to analyse this behaviour. Non-linear dynamical modeling for source clearly indicates the relevance of non-deterministic approaches in understanding these types of signals [3, 4]. Some form of regular variation of complexity in tanpura signals has already been reported [5]. In the present paper D_0 (fractal dimension) and D_2 (correlation dimension) and also the higher moments for sarod signals are studied. The obeying of power law indicates the presence of non-linearity in the production system. The presence of multifractality is studied through an examination of relationship between q and D_q and the functional relationship between the D_q 's.

Methodology

Fractal Dimension

In general music signals are single valued continuous time function of pressure. Digitisation converts them to discrete point series. Depending upon the sampling rate, the sharpness of variations of the original signal is retained in the time series. Music signals contain relevant information up to approximately

10 kHz that necessitates a sampling rate of around 20 kHz. For the purpose of fractal analysis the lowest meaningful box size can only be what corresponds to the sampling rate. With this limiting box size no box can contain more than one point because the original function is a single-valued function of time. Also the number of boxes required to cover the point set in this limiting case would be equal to the total number of samples. The object of analysis in the present case is a map of single valued continuous function and, therefore, the geometric objects are not closed figures but rather are only open segments.

We have used amplitude-normalised form of the pre-processed acoustical signal i.e., the pressure signal as a function V_t of time. The values of V_t would form a digital curve in the two dimensional plane. The length of this curve may be estimated by counting the number $N(r)$ of boxes of size r , which would be required to cover the full length of the curve. According to Mandelbrot [6] $N(r)$ should obey the power law,

$$N(r) \sim r^{-D} \quad \dots \quad (1)$$

Where D is a positive number. From (1)

$$D = \ln N(r) / \ln(1/r) \quad \dots \quad (2)$$

The purpose of covering the curve with similar squares of small size is to estimate the length of the curve. If we decrease the size of the squares gradually it would be possible to take care of finer and finer details of the curve and more length would be exposed. We shall observe that as r decreases, $N(r)$ increases and D varies within a short range. In some limiting situation D attains a constant value. This is the fractal dimension D of V_t . D is given as

$$D = \lim_{1/r \rightarrow 0} \{ \ln N(r) / \ln(1/r) \} \quad (3)$$

As already mentioned, r can be reduced only to the value which corresponds to the sampling rate. It should be noted that in this limiting case each column would generally contain one box that contain a sample. It therefore may not properly reveal the complexity of the signal. If we examine the $\ln N(r)$ vs. $\ln(1/r)$ curve (figure 2) we shall see that there are two areas where the curve is approximately linear implying the obeying of power law (eqn. 1). While D , normally used for time series, at $1/r \rightarrow 0$ is < 1 , D at the small values of $1/r$ will also give a fractal dimension which is > 1 and will reflect the nature of the complexity of the signal with which we are interested.

In general, these curves are almost linear at the beginning up to a value of

$\ln(1/r)$ corresponding to the box size of 128, starting from box size of 2048. Thereafter they start to bend towards the x-axis.

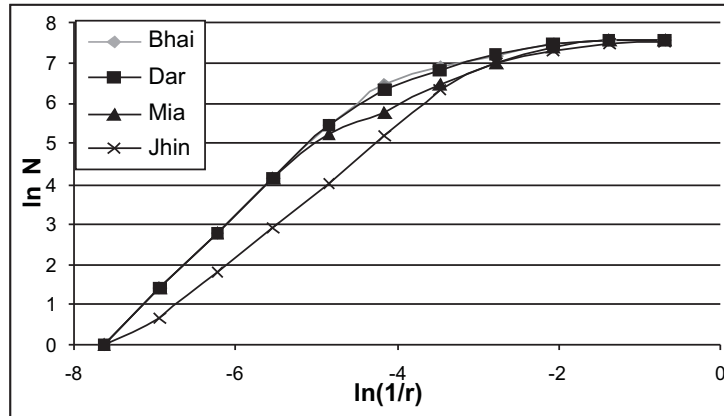


Figure 2. $\ln P - \ln(1/r)$ plot for four ragas

Generalised Dimension

The gross features of the fractal set can be observed by calculating the fractal dimension D . Now, to see whether the distribution of points on the set is uniform or not, the set is again covered by small boxes of size r .

With the i_{th} box we now associate a measure p_i , the probability of the map visiting the i_{th} box. The generalised dimension is now given by [7].

$$D_q = - \lim_{r \rightarrow 0} \left[\frac{1}{q-1} \frac{\ln \sum_{i=1}^N p_i^q}{\ln(1/r)} \right] \quad \dots (4)$$

Where N = total number of boxes and p_i = probability of the map visiting the i_{th} box, which is the total number of points n_i of the time series that have fallen in the i_{th} box normalised to the total number of points n , i.e., $p_i = n_i/n$. When $q = 0$, D_0 becomes the fractal dimension.

If the distribution is uniform, p_i becomes p independent of i .

$$p_i^q = p^q = (1/N)p^q = p^{q-1} \quad (5)$$

and $D_q = D_0$ for all values of q .

Experimental Details

Approximately 10 minutes of alap were taken from the recordings of sitar signals played by Ali Akbar Khan (ragas- Bhairav, Darbari Kanada, Mian ki Mallhar) and Verma (Jhinjoti). These were recorded directly on Pentium IV PC. The signals were digitized at a sampling rate of 22050/sec (16 bit per sample). Each signal consists of the alap part of the above mentioned ragas. Each signal space was divided into non-overlapping windows of 2048 sample points. For dividing the image window of 2048 x 2048 into smaller box sizes the concept of a floating origin, which minimises the number of boxes for a given size at each time slot has been used [4]. Each window was first divided into vertical slots of the required box size. Then in each vertical slot the lowest sample point was detected. The ordinate value of this point was considered as the origin for that slot. The upper part of this slot was then divided into square boxes. All counting relate to these boxes.

Fractal and Correlation dimensions were extracted using the methods outlined in the earlier section. Fractal and Correlation dimensions for each window of each raga were obtained by box counting method [6]. It may be noted that this approach allows some optimality in covering the map according to the distribution of the clusters without introducing any scaling. D_0 and D_{10} were calculated by the aforesaid method. In the present study, behaviour of the fractal and correlation dimensions is presented separately for these signals. Each signal is represented by the mean value of the corresponding dimensions for all the windows in that state.

Results And Discussions

Table 1 presents the mean and standard deviations of the values of D_q for all

D_q	Bhai		Dar		Mia		Jhin	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
D_0	1.9	0.03	1.95	0.09	1.98	0.06	1.95	0.08
D_2	0.98	0.32	1.18	0.8	1.24	0.68	0.53	0.87
D_3	1.16	0.24	1.33	0.64	1.39	0.55	0.85	0.7
D_4	1.2	0.22	1.37	0.6	1.42	0.51	0.94	0.64
D_5	1.21	0.21	1.37	0.58	1.43	0.49	0.97	0.61
D_6	1.21	0.2	1.37	0.57	1.43	0.48	0.98	0.6
D_7	1.21	0.2	1.36	0.56	1.42	0.48	0.98	0.59
D_8	1.2	0.19	1.36	0.55	1.42	0.47	0.98	0.58
D_9	1.19	0.19	1.35	0.55	1.41	0.47	0.98	0.57
D_{10}	1.19	0.19	1.34	0.55	1.4	0.47	0.98	0.57

Table 1: Mean and SD of D_0 to D_{10} for four ragas

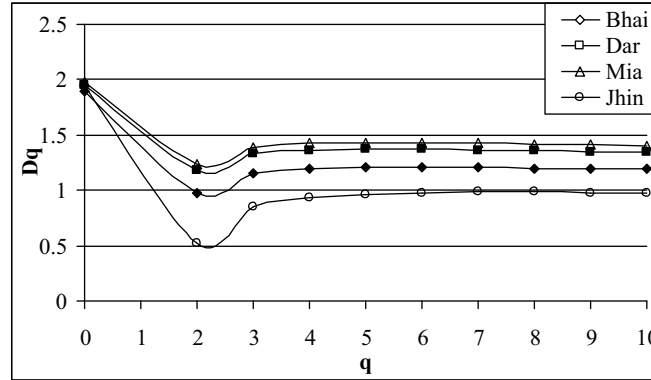


Figure 3. Variation of D_q with q

	a_0	a_1	a_2	a_3	a_4	a_5	a_6	R^2
Bhai	1.8995	-1.7793	1.1646	-0.3367	0.0491	-0.0035	1.00E-04	0.9979
Dar	1.9525	-1.4956	0.9811	-0.2841	0.0414	-0.003	8.00E-05	0.9979
Mia	1.9838	-1.4561	0.9591	-0.2781	0.0406	-0.0029	8.00E-05	0.9978
Jhin	1.9521	-2.8557	1.8989	-0.5512	0.0804	-0.0058	0.0002	0.9974

Table 2. The coefficients of the 6th degree polynomial with value of R^2

the windows covering the whole of the signal covering separately all the four ragas. Standard deviations for D_0 are significantly lower than the higher dimensions for all the ragas. This indicates that the estimate of D_0 by the mean over all windows is a good one. While D_0 values appear to be non-distinctive for different ragas the higher dimensions appear to be quite distinctive.

Figure 3 shows the variation of fractal dimensions averaged over all windows for a raga for different moments. The curves look quite similar in form but shifted along the D_q axis. It indicates that they be represented by the same functional form. In fact the high values of R^2 indicate that a sixth degree polynomial fit quite satisfactorily. This good fit indicates a functional form to represent the behaviour of D_q s with q and in turn points towards the multifractality of the signals. Table 2 gives the values of the coefficients of the polynomial fitted to the D_q values for each raga. It is interesting to note that the average values of the dimensions saturates at $q=4$, same as that obtained for sitar signals [8].

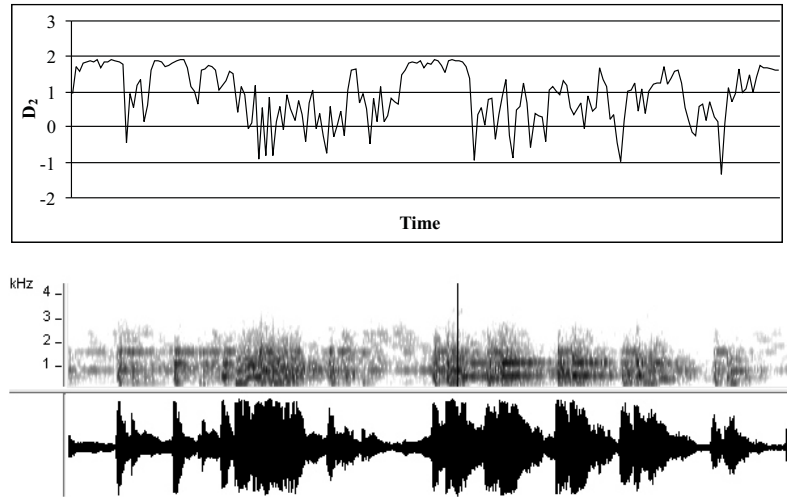


Figure 4. D_2 , 3D spectra and waveform of a portion of Sitar composition

Figure 4 presents variation of D_2 (the top curve), 3D spectrum (middle part) and waveform (the bottom part) of a portion of a sarod signal. An examination of the figure reveals that D_2 decreases with increase of complexity. One measure of the complexity may be the spectral energy calculated by summing the amplitudes of the components of the Fourier spectrum [8]. The shaded area in figure 5 shows the spectral energy of a part of the signal. Table 3 gives the correlation of D_2 with spectral energy for signals corresponding to different ragas. Both negative and positive correlations are observed. In most of the cases the correlation value is large. It seems that, at least for these sarod signals, the relation between spectral complexity and D_2 is ambiguous.

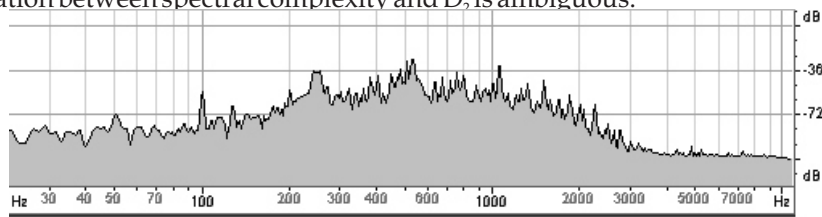


Figure 5. Spectral energy for a window of 2048 points of a signal

Bhai	Dar	Mia	Jhin
-0.80778	-0.88339	0.460047	0.975467

Table 3. Correlation between spectral energy and D_2

Conclusions

1. Sarod signals are multifractal in nature.
2. D_0 for all samples are consistent.
3. Correlation between D_2 and spectral complexity is not consistent.

Reference

1. B Roychoudhury, "Bharatiya Sangeet Kosh", Katha Silpa Prakash, Calcutta 12, E S Perera, "The Origin and Development of Dhrupad and its bearing on instrumental music", K P Bagchi and Co., Calcutta 12
B M Banerjee et.al., "Spectral analysis of the acoustic signals from Sarod played by an expert", J. Acoust. Soc. Ind., Vol.26 (3&4), 1998
2. R.Devaney, "Introduction to chaotic dynamical systems", Addison-Wesley, NY, 1986.
3. R.Sengupta, N.Dey, D.Nag and A.K.Datta, "A study on fractal analysis of vowel sounds", J. Acoust. Soc. India, Vol. XXVII, 1999.
4. R.Sengupta, et. al., "Comparative study of Fractal Behaviour in Quasi-Random and Quasi-Periodic Speech Wave Map", FRACTALS - Complex Geometry, Patterns and Scaling in Nature and Society, Vol.9, No.4, pp.403, 2001.
5. R.Sengupta, et. al., "Fractal analysis of Tanpura signal", J. Acoust. Soc. India, Vol.XXVII, 2000. R.Sengupta, et. al., "Study on Shimmer, Jitter & Complexity Perturbations in Voices of Singers & Non- Singers", J. Acoust. Soc. India, Vol. XXVIII, 2000. R. Sengupta, et. al., "Fractal and Correlation Dimension Analysis in Tanpura Sound Signals", Proc. Indian National Science Congress, January, 2002, Lucknow, India. R. Sengupta et.al., "Assessment of Musical Quality of Tanpura by Fractal Dimensional Analysis", FRACTALS - Complex Geometry, Patterns and Scaling in Nature and Society, Vol.13, No.3, 2005.
6. B.Mandelbrot, "The Fractal Geometry of Nature", Freeman & Co., San Fransisco, 1982.
7. G.P.Williams, "Chaos Theory Tamed", Taylor & Francis, London, pp.321-328, 1997. T Vicsek, "Fractal Growth Phenomena", 2nd. Ed., World Scientific, Singapore, pp.55, 1992.
8. A.K.Datta et.al., "Analysis of the signal complexity in Sitar performances", J. Acoust. Soc. India, Vol. 32, 2004.

MUSIC EMOTION AND THE BRAIN

Shantala Hegde
Cognitive Psychology Unit
Center for Cognition & Human Excellence
Department of M H & S P
National Institute of Mental Health & Neuroscience (NIMHANS)
Bangalore 560029, India
shantalah@nimhans.kar.nic.in

Abstract:

This paper provides an overview of current understanding on music and emotions from the cognitive neuroscience perspective. It is a fact that music has the power to induce and modulate emotional experience. Over past thirty years, there has been increased interest among cognitive neuroscientists to understand the very purpose of music in our lives, the role that it is playing and its neurobiological basis. The associated emotional experience seems to be key reason for ubiquity of music. Musical processing demands a fine level of coordination and involvement of several of the brain areas. Better understanding of how our brain processes music is imperative in the light of growing popularity of music as a therapeutic method. There is growing need for evidenced based research and cross-cultural studies. In this paper an attempt has been made to highlight the need for research in neuromusicology in our country using our own music which has certain unique features not observed in other forms of music in the world.

Key Words: Music, Emotion, Cognition, Neuromusicology

Music and emotional experience share an inseparable association. The intimate relation between music and emotional experience has been the topic of scholarly debate in various disciplines such as philosophy, musicology, anthropology, literature, psychology and neuroscience. Trying to understand the meaning of music and its effect is not a recent endeavour [1]. Effect of music on our system has been appraised from time unknown. Aristoxenus an Aristotelian philosopher (4th century BC) proposed that musical intervals should be classified by their '*effects on listeners*' as opposed to merely examining their mathematical ratios which was until then a much entertained viewpoint [2]. For example in Indian Classical system (ICM), the artist aims at expressing the *rasa* (emotional flavour) associated with the ragas. There is a well established system of ragas and talas that form its edifice. Ragas can be described a melodic framework with specific tones with a well defined ascending (*aroh*) and descending (*avaroh*) structure as well as unique phrases (*pakkad*). The very word 'Raga' has its origin from the Sanskrit word (**ranjayiti iti raagah**) meaning 'that which entertains or colours our mind'. Each raga in the

ICM is conventionally assigned to a corresponding rasa/ emotion is known to consistently evoke certain emotions. The artist exploits ones creativity and elaborates the raga well within the melodic framework to bring out the rasa or the emotion. In the Hindustani or the North Indian classical music system (HCM) ragas are also assigned to a particular time of the day and particular season. The intricate relation between musical compositions, the tempo and rhythm is closely associated with various emotional experiences. There are hundreds of varieties of talas or rhythmic cycles. Compositions are set to certain ragas and taala with the aesthetic sense to enhance the emotional experience.

Not surprisingly, music is often considered as language of emotion [3]. The very emotional experience we derive from indulging in musical behaviour is perhaps the obvious reason for the ubiquity of music in our lives [4]. Despite widespread presence of musical behaviour, music was only occasionally studied as a basic and distinctive cognitive ability in the area of cognitive neuroscience and cognitive psychology. Within the area of cognitive neuroscience the topic of music and emotion was rather neglected for a very long time. The abstract and complex nature of music was perhaps one of the reasons for this neglect. Secondly, musical disability following an injury to the brain would have far less impact on an individuals day to day functioning as compared to deficits in language ability. Thirdly, the very existence of music in our lives has been focus of much debate. Music was by several scientists considered as mere spandrel or a function that supported development of language and a by product of mechanisms underlying language[5]. Steven Pinker, a cognitive scientist at the Massachusetts Institute of Technology in his book 'How mind works' termed music as mere “auditory cheese cake” and deemed music as a non-adaptive phenomenon that activates other truly adaptive functions such as rhythmic bodily movements, walking, running, natural cadences of speech and our ability to make sense of cacophony of sounds[6].

These discussions and debates served as an impetus to the burgeoning interest in understanding the biological basis of music and led to the growth 'Neuromusicology' and 'Music cognition' as a specialized branch in cognitive neuroscience and psychology. This area of research owes a lot to the advances in technology and methods of investigations such as newer brain imaging techniques like functional magnetic resonance imaging (fMRI) which provides information on the areas of brain involved while performing a given task. These advances imaging techniques was not envisaged few decades ago. Over the past thirty studies on music and cognition has grown exponentially. The simplistic view of brain areas implicated in language and music processing to the left and right hemispheres has been challenged and proved inadequate to

explain all facets of musical behaviour.

Music is ubiquitous. We all indulge in musical behaviour, actively or passively. Actively we indulge in playing an instrument or singing and passively we may listen to music playing in the background while performing other day to day activities. Many of us are good music critics as well; we are able to identify if a melody goes off tune; able to judge if an artist has presented a composition elegantly; have the aesthetic sense which enables us to differentiate a maestro from an amateur performer by mere listening and able to easily recognize timbre i.e., the sound of a particular instrument or voice of the singer. From a cognitive neuroscience perspective all the above functions that we carry out pretty much automatically in matter of few seconds or fraction of a second involves an array of cognitive functions such as information processing, attention, learning, memory, decision. Indulging in creating music makes significant demands on several cognitive functions. One has to keep several features of music in mind- for example the pitch, the rhythm, memory for a composition, emotional content etc. Music composition and improvisation can be seen as a preparatory activity for training a host of higher cognitive functions such as focused attention, fluency, working memory, planning etc. From an evolutionary perspective indulging in musical behaviour would seem like a display or an indication of mental as well as emotional flexibility and fitness [7]. In other words ability to produce and express via music would indicate presence of superior intellect and a creative mind thereby increasing the survival value of the individual. Music-dance often coexists and this would indicate physical fitness and motor coordination along with creative mind. Emotions too play a very important role in modulating and controlling our higher cognitive functions such as planning, decision making, memory so on and so forth [8-9]. Sandra Trehub, psychologist from the University of Toronto suggests that singing is the key to mother-child bonding. Maternal speech differs from normal speech and has several musical features such as higher pitch, slower tempo which are associated with happiness and tenderness. Infants are known to prefer maternal singing over adult speech as enhances the survival value by emotional and arousal regulation of the infant[10].

Studies have aimed to understand musical behaviour such as listening, remembering, performing, learning, composing as well as to certain extent movement and dancing among musically trained as well as untrained individuals. There has also been an increased interest in understanding the relation between music and language, to examine brain network that are overlap and unique to these two functions[11] as well in the area of music and emotion[1]. Understanding musical emotion is all the more important for the fact that unlike other stimuli in our everyday life, music is intentionally

designed to induce emotions! Many of us seek out for certain musical pieces in order to alter already existing mood. Music is also known to alter physiological conditions such as heart rate and respiratory changes[12]. Music can be the source of a variety of pleasant bodily sensations, such as chills, thrills and tingles running down the spine[13-14]. Probably humans indulge in music and experience the emotion, with or without the obvious knowledge of the therapeutic aspects of music.

Researchers have used direct methods like interviewing and behavioural ratings by the listeners, as well as advanced tools that help measure the brain electrical activity and blood flow in the regions of the brain by using methods like magnetoencephalography (MEG) electro-encephalography and event related potentials (EEG/ERP), the positron emission tomography (PET) or functional magnetic resonance imaging (fMRI). Methods like polygraph have been used to measure heart rate variability, respiration, sweating which are considered to be key indicators of emotional processing. The other method which is employed by cognitive neuroscientists and neuropsychologists is the classical method i.e., to examine clinical cases with brain lesions following conditions like head injury, stroke or tumour and study the deficits causing variations in musical behaviour. The behavioural and cognitive changes following injury to the brain would help us understand the brain areas involved in processing the cognitive function under investigation.

Science helps us understand the functioning of various phenomena present in nature. It has enabled us to explore and understand the process and mechanism of how we process music, how and why does it have the effect on us, the way it does. Understanding how our brain perceives and makes meaning out of music would yield valuable insights into the brain itself as well as help us understand manner in which music effects on our biological system. This pursuit is ongoing and in this paper I only aim to provide an overview of research work that has been carried out to understand neurocognitive basis of musical emotion. I believe the findings of research will also benefit musicians in appreciating the manner in which certain very obvious or naturally occurring nuances of music that they perform has varying effect on their audiences. Better understanding of this intimate relation between music and the emotional experience and its biological basis is crucial with the growing popularity of music as a therapeutic method.

Before dwelling further upon the concept of musical emotion I shall introduce the reader to some of the commonly used terms to refer to various emotional experiences such as affect, emotion, mood, emotion and musical emotion. Musical emotion encompasses the various emotional states

experienced due to music. *Affect* is an umbrella term that covers all evaluative or valenced states (i.e., positive/negative) such as emotion, mood, and preference. *Moods* are affective states that feature a lower felt intensity than emotions that do not have a clear object but remain for several hours to days[15]. *Emotions* are relatively brief, though intense, affective reactions to potentially important events or changes in the external or internal environment. Emotions are adaptive responses that have positive effect on survival value. It has six subcomponents. a) Cognitive appraisal (appraisal of a situation as dangerous) b) subjective feelings (one feels afraid) c) physiological arousal (heart starts pounding d) expression (scream or call out for help) e) action tendency (strongly feel like running away from the situation f) regulation (try to calm oneself) [16]. Music can cause mood changes as well as emotional changes[17].

Musical emotion and everyday emotion:

The question which logically arises is whether emotion induced via music has the same effect on ourselves similar to other 'real' emotions experienced by us? For example while listening to a musical piece one might say 'this is a sad song or a happy song' without actually experiencing the sadness or happiness. 'Real' emotion would be actually experiencing sadness due to a personal loss or happiness due to some joyful event. The psychological approach to music and emotion seeks an explanation for this process of how and why we express and experience emotional via music[1]. The neuroscientific approach has aimed at understanding brain basis of musical emotion.

There are evidences from many different strands of research showing how musical emotion does produce changes in all the above mentioned six subcomponents of emotion. Experiments using questionnaire methods and qualitative interviews have shown that listeners report that experienced various emotions while listening to music [16-18-19]. Similar to findings on emotional appraisal via facial expressions people are able to appraise happy and sad emotions quite easily and consistently in music from another culture . Autonomic responses such as heart rate, respiratory changes, bodily temperature are known to vary according to the reports of valence (positive or negative) and arousal/intensity (low or high), which are considered as two dimensions of emotion[20] including musical emotions[12-21]. Music listening brings about expressive changes in form of making people cry, laugh, make facial grimaces[22]. Music is known to influence action tendencies such as enhancing tendency to help others, affect consumer behaviour[21] Listeners attempt to regulate their own emotional reactions to music, e.g., with regard to what are deemed appropriate responses in a social context[23-24]. Evidence has led us to believe that musical emotions are indeed similar to emotional responses to 'real or everyday emotions'.

How does music induce emotion?

The central question is how does music induce emotion? What feature or aspect of music does a composer or an artist makes use of to convey a particular emotion and what enables the listener to experience this emotion? Several psychological mechanisms have been implicated in induction of musical emotion. Musical emotion may be induced by its association with an event in the listener's life. For example an individual may experience feeling of warmth by listening to a particular song that was often sung by his mother during his/her childhood. Emotion may be induced due to the association formed between particular music or a song due to repeated paring with other positive or negative stimuli. The listener may not be aware of this unconsciously formed association. A particular song played at a friend's place over a number of occasions may evoke the same feeling of happiness even when heard in a different occasion or place. Emotion may also be induced in a listener because he or she conjures up visual images, a beautiful landscape or a death scene for example. There would be a close interaction between the music and the images. The above mentioned psychological mechanisms explain the subjective experience and variability between listeners. Musical expectations have been shown play a very important role in induction of musical emotion. Musical expectancy involves syntactical relationships between different parts of the musical structure similar to those in language [25]. This aspect therefore would explain why listeners from different cultures react differently to the same composition. Musical expectations do not depend on explicit musical training. Musically untrained individuals internalize it through passive exposure[26-27]. An emotion is induced in the listener because a specific feature of music violates or confirms or merely differs from the listener's expectations about the continuation of the music. For example listeners of HCM experience a positive emotion when the artist after elaborating on a composition comes back to 'sam' or the first point of the rhythmic cycle. Here a positive emotion is often induced when there is a confirmation of one's expectation and the actual musical rendition. Several studies have shown that listeners are able to pinpoint the particular moment in the musical excerpt when they experienced a strong emotion and these coincided with points in the music where an expectation of some kind was violated. Different physiological reactions such as 'chill' or 'shiver down the spine' were produced by different kind of structural changes and violations in the music[13-22]. The structural aspects of music are known to induce emotional experiences that are consistent across listeners. Musical structure functions as a code for communicating emotion[17]. Structural aspects of music would include features like tempo, mode (major or minor or other variations), micro-intonation, contour, rhythm, sound level, timbre, accents on specific notes, tone attacks and decays etc. Faster tempi, major mode,

consonant intervals is considered to be more positive (happy) and slower tempi and minor mode, dissonant intervals is considered to be more negative (sad)[28]. Musical instrument, the quality of voice or timbre is also highly relevant to expression of musical emotion[19-23]. Composers select particular instruments (either consciously or intuitively) to convey specific emotional qualities.

In a study carried out by the author six raga excerpts of HCM were taken. The six raagas were chosen keeping in mind Indian Music theory. In tune with the raaga-rasa theory three raagas with high number of major notes (shuddh swaras) and three raagas with minor notes (komal swaras) were chosen to induce happy (positively valenced) or sad (negatively valenced) emotion. By presenting all the excerpts in flute, timbre was kept constant. The raaga excerpts were varied at two levels. One was with nature of the notes which was taken care by the raga chosen. Second variation made was with tempo. Two excerpts from each raaga, one from the raaga elaboration section (Alaap) without tempo and the other from the raaga elaboration with tempo (Jod- Jhala). Despite being unfamiliar with the musical tradition, ratings given by participants were in tune with raaga-rasa aspect of Indian music theory. Participants rated raagas with major pitches with tempo as having positive valence and raaga excerpts with minor pitches with no tempo as having negative valence and all participants showed similarity in their ratings. Variation in tempo influenced and enhanced the intensity of the experience of positive emotion[29]. Cross cultural studies have shown that people use the above mentioned features as measure to evaluate the emotional content even in music that they are unfamiliar with[19-30]. Interestingly research has shown there are specific areas in the brain that are specialized for perception of various features of music such as melody, rhythm and timbre. Clinical cases have been recorded where an individual following injury to the brain has the ability to process only certain features of music and not the others [31].

Sensitivity to emotional meaning in music emerges early in development, as early as infancy and improves with age. By the age of three, children show ability to recognise happiness in music from the culture they belong to and by the age of 6 years they show adult like abilities to identify happiness and sadness in music based on specific musical features such as tempo and mode[32-36]. Studies on children in the area musical emotion have shown that sensitivity to tempo variations precedes sensitivity to mode[33]. Infants as young as two months prefer to listen to consonant over dissonant musical intervals [37].

Empirical studies examining individuals following injury to the brain as well using imaging methods like the fMRI have shown that music elicits intense emotional responses that activate brain regions thought to be involved in reward/motivation, known as limbic system and the para limbic areas which

are buried deep within the cerebral cortex are implicated in emotional processing [14 38]. The activations in these areas varied in their degree depending on intensity of pleasantness and unpleasantness as perceived by the listener.

Brain has the ability to change its environment and the environment in turn changes the brain functions. The changes in brain functioning and structure following exposure to certain kind of stimuli or environment is described in the neuroscience field as 'brain plasticity or 'neuronal plasticity'. Music, a rich art form is a product of complex brain functions and indulging in this product in turn as we know now has immense ability to restore and alter brain functions. Modulation of emotion via music plays a central role in bringing about psychological changes in individuals.

As mentioned previously in the text, appreciation of therapeutic effect of music on health has been known since ages. During the World War II musicians were brought into veterans' hospitals as a way to boost the morale of soldiers requiring long-term stays. It was observed that these performances not only lifted soldiers' spirits but also made significant contributions to their physical and mental rehabilitation[39]. Neurocognitive research as a basic research approach has the potential to identify and explain the relevant effects of music therapy using experimental research designs and neurophysiological investigation methods. Scientists are yet to fathom the role musical emotion is playing in our lives and how it affects our system in myriad ways. Music as a therapeutic technique is growing in popularity. There is need to carry out systematic research work to understand the neurocognitive basis of Indian music and carry out evidence based research work to ascertain the therapeutic effect.

Despite the world's diversity of musical cultures, the majority of research in cognitive psychology and the cognitive neuroscience of music have been conducted on subjects and stimuli from Western music cultures. From the standpoint of cognitive neuroscience, identification of fundamental cognitive and neurological processes associated with music requires ascertaining that such processes are demonstrated by listeners from various cultural backgrounds and music across cultural traditions. It is unlikely that individuals from different cultural backgrounds employ different cognitive systems in the processing of musical information. It is more likely that different systems of music make different cognitive demands[40].

Western classical music which is based on harmonic relation between notes versus the melodic mode (raaga) structures in the Indian classical music system (ICM) with in the rhythmic cycle music may demand qualitatively different

cognitive engagement. To elaborate this point further, ICM music is monophonic or quasi monophonic. Unlike the western classical system, there is no special notation system for music. Instead letters from the colloquial languages are used to write music. For instance notations of the ICM such as 'Sa, Re, Ga" may be written in Hindi, Kannada or Tamil where as the Western classical system music includes a unique visuo-spatial representation. It emphasizes on reading the exact position of symbol indicating a whole tone or a semitone on the treble or a bass clef. The scale systems (Raagas) are quite elaborate and complex provides a strict framework within which the artist is expected bring out maximum creativity. Although specific emotion (rasa) is associated with particular raaga, it is well known that the same raaga may evoke more than one emotions[41-42]. Well trained artists are able to highlight a particular rasa by altering the structures of musical presentations such as stressing on specific notes, accents, slurs, gamakas or taans varying in tempo etc. Musicians as well as ardent connoisseurs of music would agree that every single note has the ability to convey an emotion. Many experience a 'chill' or 'shiver down the spine' when a musician touches certain note or sustains of a note. The meter system is again quite complex. Indian rhythm & metre system is one of the most complex systems compared to other meters used in world music. Film music, which has been influenced by music from all over the world, is much more popular in the current times. Therefore implicit of knowledge of the Western chord system is perhaps present in our population. ICM is chiefly an oral tradition with importance given on memorizing compositions and raaga structures and differences exist in the methods of training even within the two traditional systems of ICM. Semantics of Indian music would differ from that of the western classical music system or other forms of musical system. More often than not music in Indian culture is intimately associated with religious and spiritual practices. Hypothetically these differences in the musical systems perhaps makes qualitatively different demand on the cognitive functions involved and thereby qualitatively varying degree of involvement of the specialized neural networks implicated in musical processing. Research endeavours are yet to be carried out in this direction.

To summarise and conclude musical emotion as an experience has been considered as a psychological construct all the more seriously in the recent past. Musical emotions have shown to induce the same psychological and physiological changes similar to 'real life emotions'. It engages the areas of the brain involved in motivation, reward appraisal. Evaluation of musical emotion like other cognitive process is based upon feature analysis, syntactical and analysis of semantic similar to language. Humans are able to appraise some of the basic emotion via music and these are shown to be universal phenomenon. Musical expectations which play an important role in appraisal of emotion

however is known to be influenced by explicit or implicit exposure to music from ones own culture. There still remains a great lag in empirical endeavours examining cross cultural variations. There is a need for better understanding of musical emotions itself as well as evidence based research in understanding the neuroscientific basis of music as mode of rehabilitation and therapy. One can expect significant advances in near future, in this area of science. Apart from the growing interest in the area of music and emotion, music therapy and the advances would be due to newer developments in methods of investigation such as the brain-imaging methods. With more and more evidences proving the biological basis of music and effect of music on various aspects of human functioning including other higher order cognitive functions, neuroscientists have accepted music as a key element in the human behavioural repertoire and not just as abstract and adjunct phenomena.

References:

1. Juslin PN, Sloboda JA. *Music and Emotion: Theory and Research*. London: Oxford University Press, 2001.
2. Griffiths P. *New Penguin Dictionary of Music*. London: Penguin Books, 2004
3. Scherer KR. Which emotions can be induced by music? What are the underlying: Mechanisms? And how can we measure them? *Journal of New Music Research* 2004;33:1-13.
4. Panksepp J. The emotional sources of "chills" induced by music. *Music Perception* 1995;13:171-207.
5. Darwin C. *The descent of man and selection in relation to sex*. London: John Murray., 1871.
6. Pinker S. *How the mind works*. New York: Norton, 1997.
7. Sluming VA, Manning JT. Second to fourth digit ratio in elite musicians: Evidence for musical ability as an honest signal of male fitness. *Evol. Hum. Behav* 2000 21:19.
8. Rains GD. *Principles of human neuropsychology*: McGraw Hill Companies, 2002.
9. LeDoux J. *The emotional brain*. London: Phoenix, 1998.
10. Trehub SE. The developmental origins of musicality. *Nat Neurosci* 2003;6(7):669-73.
11. Levitin DJ, Tirovolas AK. Current advances in the cognitive neuroscience of music. *Ann N Y Acad Sci* 2009;1156:211-31.
12. Krumhansl CL. An exploratory study of musical emotions and psychophysiology. *Can J Exp Psychol* 1997;51(4):336-53.

13. Blood AJ, Zatorre RJ, Bermudez P, Evans AC. Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat Neurosci* 1999;2(4):382-7.
14. Menon V, Levitin DJ. The rewards of music listening: response and physiological connectivity of the mesolimbic system. *Neuroimage* 2005;28(1):175-84.
15. Juslin PN, Vastfjall D. Emotional responses to music: the need to consider underlying mechanisms. *Behav Brain Sci* 2008;31(5):559-75; discussion 75-621.
16. Juslin PN, Laukka P. Expression, Perception, and Induction of Musical Emotions: A Review and a Questionnaire Study of Everyday Listening. *Journal of New Music Research* 2004;33(3):217-38.
17. Bharucha JJ, Curtis M, Paroo K. Varieties of musical experience. *Cognition* 2006;100(1):131-72.
18. Balkwill L, Thompson WF, Matsunaga R. Recognition of emotion in Japanese, Western and Hindustani music by Japanese listeners. *Japanese Psychological Research* 2004;46(4):337-49.
19. Balkwill L, Thompson WF. A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues. *Music Perception* 1999;17(1):43-64.
20. Lang PJ, Bradley MM, Cuthbert BN. Emotion and motivation: measuring affective perception. *J Clin Neurophysiol* 1998;15(5):397-408.
21. North AC, Hargreaves DJ. Liking, arousal potential, and the emotions expressed by music. *Scand J Psychol* 1997;38(1):45-53.
22. Sloboda JA. Music structure and emotional response: Some empirical findings. *Psychology of Music* 1991;19:110-20.
23. Gabrielsson A, Juslin PN. Emotional expression in music performance: between the performer's intention and the listener's experience. *Psychology of Music* 1996;24:68-91.
24. Gabrielsson A. Emotions in strong experiences with music. In: Juslin PN, Sloboda JA, editors. *Music and Emotion: Theory and Research*. New York: Oxford University Press, 2001:431-49.
25. Patel AD. Language, music, syntax, and the brain. *Nature Neuroscience* 2003;6:674-81.
26. Meyer LB. *Emotion and meaning in music*: University of Chicago Press, 1956.
27. Krumhansl CL, Louhivuori J, Toiviainen P, Järvinen T, Eerola T. Melodic expectation in Finnish folk hymns: Convergence of statistical, behavioral, and computational approaches. *Music Perception* 1999;17:151-95.

28. Peretz I, Gagnon L. Mode and tempo relative contributions to "happy-sad" judgements in equitone melodies. *Cognition and Emotion* 2003;17(1):25-40.
29. Hegde S. Appraisal of Happy & Sad Musical Emotion: Influence of pitch distribution and tempo. Bangalore, 2010.
30. Gregory AH, Varnay N. Cross cultural comparison in the affective response to music. *Psychology of Music* 1996;24:47-52.
31. Peretz I, Gagnon L, Bouchard B. Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition* 1998;68(2):111-41.
32. Cunningham JG, Sterling RS. Developmental change in the understanding of affective meaning in music. *Motivation and Emotion* 1988;12:399-413.
33. Dalla Bella S, Peretz I, Rousseau L, Gosselin N. A developmental study of the affective value of tempo and mode in music. *Cognition* 2001;80(3):B1-10.
34. Kratus J. A developmental study of children's interpretation of emotion in music. *Psychology of Music* 1993;21:3-19.
35. Terwogt MM, Van Grinsven F. Recognition of emotions in music by children and adults. *Percept Mot Skills* 1988;67(3):697-8.
36. Terwogt M, Van Grinsven F. Musical expression of mood states. *Psychology of Music* 1991;19:99-109.
37. Schellenberg EG, Trainor LJ. Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: tests of adult and infant listeners. *Journal of Acoust. Soc. Am* 1996;100:3321-8.
38. Blood AJ, Zatorre RJ. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc Natl Acad Sci USA* 2001;98(20):11818-23.
39. Davis WB, Gfeller KE. Music therapy: Historical perspective. In: Davis WB, Gfeller KE, Thaut MH, editors. *An introduction to music therapy: Theory and practice* 3rd ed: Silver Spring, MD: The American Music Therapy Association, 2008:17-39.
40. Koelsch S. A neuroscientific perspective on music therapy. *The Neurosciences and Music III Disorders and Plasticity: Ann. N.Y. Acad. Sci.* 2009;1169:37484.
41. Jairazbhoy NA. *The rags of North Indian music: Their structure and evolution.* London: Faber & Faber., 1971.
42. Bhatkande V. *Hindusthani Sangeet Paddhati.*: Sangeet Karyalaya., 1934.

Acknowledgement:

Sincere thanks to Shri. Laxmikant Hegde, Ms Arya and Ms Bhargavi for their helpful comments and feedback during the preparation of the manuscript.

A GEOMETRIC MODEL OF A MUSICAL PERFORMER'S ONTOLOGICAL SPACE

Guerino Mazzola^{1,2}

¹University of Minnesota, School of Music

²University of Zürich, Institut für Informatik
guerino@mazzola.ch

Abstract

Musical performance is above all complex by its ramified ontological richness. We therefore want to give an ontologically complete picture of music that will enable us to locate performance with respect to all those coordinates known to be determinants of the overall phenomenology of music. This approach will give us the necessary conceptual architecture to unfold a presentation of this paper's subject, namely the performer's ontology. We refer to complex time as introduced by Stephen Hawking and Itzhak Bars to describe the space of a performer's presence.

We shall first describe the general setup of musical ontology and then, in a second movement, investigate the nature of performance within this general context. Although general musical ontology has been exposed in a concise way in [10] and in [9], we shall recapitulate it here in order to offer a self-contained text, but also since certain aspects must be stressed in a way as to be more adapted to the context of performance.

The first movement presents what we have coined musical “onontology”. It is the classical musical ontology that was introduced in [8], comprising the dimensions of realities, communication, and semiotics, but enriched by a fourth dimension, namely embodiment. Since this fourth dimension splits into three layers, the classical ontology is viewed as a triply layered ontology, whence the somewhat fancy name of an “onontology”.

1 Onontology

This onontology presents a topographic landscape of musical ontology, it is a geographic display of localities determined by coordinates as specified from the four dimensions of ontology. In other words, musical onontology is a conceptual space, on which phenomena of musical existence are distributed.

This spatial display enables in the second movement to interpret performance as a dynamical process that moves around in the topography, it is an ontological trajectory rather than a constant spot of musical existence. So let us get off ground with the description of this space of music.

1.1 Realities

This dimension describes the three fundamental values of reality involved in music: Physical reality, psychological reality, and mental or symbolic reality. So, acoustical phenomena relate to physics, emotional effects to psychology, and symbolic structures (e.g. mathematical descriptions in music theory) to the mental reality. Observe that the mental reality is not conceived as being a part of the psychological one.

1.2 Communication

Following the famous scheme of Jean Molino and Paul Valéry [8], music deals with communication from the first value, the poetical position of the composer, or creator, to the creator's work, which is the material essence and output of the second value, called neutral level by Molino. The communication, as encoded in the work, targets at the third value: the aesthetic position of the listener, the addressee of the composer's message. Valéry coined the word "aesthetic" to differentiate it from the aesthetical understanding. Aesthesis means perception and can be acoustical, psychological, analytical, and need not relate to aesthetical evaluation. The aesthetic instance could even be a computer software that takes a MIDI file as input and processes an analytical task thereof.

1.3 Semiotics

This axis comprises all sign-theoretic aspects of music. It is articulated in the three classical constituents of a sign: Expression, content, and signification. Expression, the first value on this axis of reality, relates to the surface of a sign, that something, which stands for the sign's meaning or content. The latter is the second value, it is the "aliquo" in the classical definition "aliquid stat pro aliquo" ("something stands for something else") of a sign. The third value is the signification part of a sign. It refers to the middle word "stat pro" of the classical definition and explains the way or process engaged for the transfer of the surface value of expression to the "hidden" value of contents. For example, when reading the musical expression for a fermata, the reader must invoke a complex machinery to understand the expression, i.e., produce the symbol's content. The classical three-dimensional cube of musical topography is shown in figure 1.

1.4 Embodiment

The very making of art is a level which was not articulated in that three-dimensional cube of musical ontology. No one of its 27 ($3 \times 3 \times 3$) positions did grasp the gestural aspect of making art (and science). The cube does, strictly speaking, only deal with the ontology of facts, of what is the case in Ludwig Wittgenstein's sense. It does not, however, include the processual level of ontology.

Formally speaking, processes are the diagrams of spaces and transformations, which describe the interaction of components of a complex system. We have to differentiate between processes and their products, the output of processual dynamics. Processes are a kind of factory for facts, but not the facts themselves. The processual level is fundamentally different from its output products. Processes and facts are instances of different ontologies.

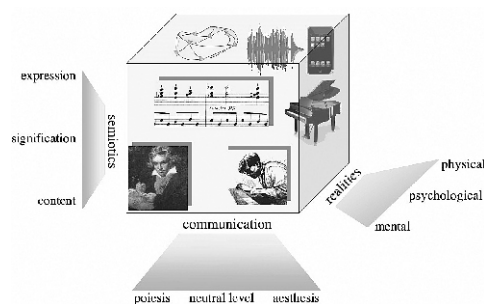


Figure 1:
The classical three-dimensional cube

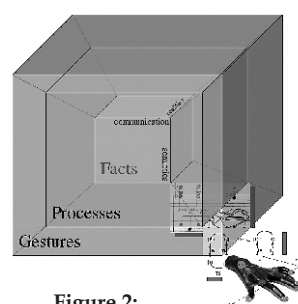


Figure 2:
The hypercube of musical ontology adding the dimension of embodiment.

Still going farther in the initiated direction, processes are also an abstraction from a more basic layer, namely the gestural layer, where all processes and their facts are initiated. Processes are disembodied gestures, reduced to their referential system of transformations.

This entails that a new dimension must be added to the cube of musical ontology. This fourth dimension is coined dimension of embodiment. Its three values are: facts, processes, and gestures. They deal with, respectively, these activities: "what is the case", "to refer to", and "to make". In this scheme, the transition from gesture to process is dominated by disembodiment and schematization, whereas the transition from process to facts is dominated by evaluation and dissection (from the relating transformations).

Together with the previous three-dimensional cube of ontology, this fourth dimension creates a four-dimensional cube, which we coin the hypercube of musical ontology. It takes the form of a three-layered onion of gestural, processual, and factual levels of ontology, as shown in Figure 2.

1.5 The Baboushka Principle

The above dimensions do not mean that musical ontology is indecomposably inscribed in such coordinates. It mostly happens that the $3 \times 3 \times 3 \times 3$ coordinates are themselves encapsulated sub- systems of the same nature. This reiteration of the hypercube's structure is called the Baboushka principle. It does not mean that new dimensions are generated, but that each position in the hyper- cube can recursively be the compactified representation of still a finer hypercube of the same type. For example, in semiotics, this is known as double articulation in language from Louis Hjelmslev's investigations [6]. And in the dimension of musical communication, the overall poietic position may be seen as articulated in composer, work (score), and interpreter, whereas this entire communicative unit is the poiesis that generates the neutral level of a performed work (on the acoustical level, say, when the performance is taking place in a concert), which in turn reaches the aesthetic level of the audience.

2 Topography of Performance

Performance can now be defined as a transformation \wp of the symbolic/mental level of the score into a set of sounding/physical events. It is crucial to understand this concept as excluding other types of performance not because they are not relevant, but because the chosen type is the per- spective which has been given the most intense and elaborate scientific investigations. However, performance involves all ontological dimensions of music. Above all, the intermediate gestural realization of score symbols, their “thawing to gestures” that act on the instrumental interface and thusly generate sounds, plays a major role, but this is unfortunately not yet a relevant topic of performance theory. Only the performance transformation \wp from the score to sounds is. But performance theory and practice is of course not focused to the mere fact of the transformation \wp , the central topic is the investigation and understanding of the transformation's “backstage” structures. In performance research these are addressed by the catchword “expressive performance”. This somewhat ambiguous concept refers to the communicative process giving rise to \wp . As such it starts from the poietic side of the composer and interpreter and is targeted at the aesthetic side of the audience and analyst. This movement is mediated by the performed acoustical and gestural rendition of music. Such expressivity has two significations: 1. It relates to a message that must be transmitted. It expresses a semiotically specified meaning or content. This expressive activity answers to what is expressed in performance, see figure 3. 2. It relates to the means and strategies used to transmit the message to the audience. This is a rhetoric activity and answers the question of how communication is shaped, see figure 4. Both, semiotic and rhetoric expressivity

have to take place and to correspond to each other in order to qualify performance as being successful: Good performance communicates contents in an adequate way.

2.1 Semiotic Expressivity

If a successful expressive shaping of performance works it has to deal first with the semiotic anatomy of the message. Let us first consider the message as it is built from the complex poietic communication unit defined by the trias Composer? Score? Interpreter, which we abbreviate by "CSI". This anatomy is centered around the score's sign expressions and points to a variety of contents distributed among the axes of embodiment and realities.

Figure 3 shows a selection of possible assignments of score-related contents to the nine positions defined by the two axes of realities and embodiment. This variety makes clear that the semiotic specification of expressivity is a crucial factor that may have strong implications on the rhetorical expressivity since its different characters will require different (if adequate) communicative strategies to be successful qua communication to the audience.

Let us shortly comment on these examples. In the left column indexed by the embodiment coordinate "facts", we find the three examples emotions, sound, and harmonic values for psychological, physical, and symbolic/mental facts, respectively. They are just given information, things that are or are not the case. They have no generative infrastructure as such. In the middle column, defined by the embodiment coordinate "processes", we find three examples, one for each coordinate of reality, which carry the character of processes. A psychological drama involves or may involve emotions, but it has also the dynamics of a movement, a diagram of mutual influences and forces

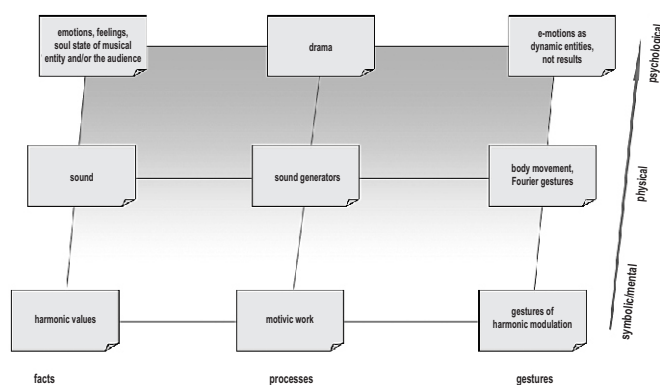


Figure 3: A set of possible message contents relating to the topographic position in the plane of realities and embodiment.

acting between different dramatic characters. Similarly, sound generators display a generative process involving different stages and components, such as envelopes, overtones, onset and duration information (see, for example, the classical description of a sound generator by Max Mathew's Music N processes). On the symbolic level, the well-known motivic work, typically realized by Beethoven, is not a fact, but a complex diagram of motives and their relations of similarity or transformations. On the level of the gestural coordinate, emotions may appear as movements, etymologically relating to emotion: moving out. Also, Fourier's decomposition of sounds can be expressed as a complex rotational movement of gestures that embody the abstract circle rotation giving rise to sinoidal waves, see [10, chapter 11] for details. Finally, harmonic modulations may be shaped by rich gestural note movemets in order to express the dramatic dynamics of tonal transformation, see also [11].

2.2 Rhetorical Expressivity

Rhetorical expressivity relates to the quality of the neutral acoustical content generated from the CSI sign by the \wp transformation in order to enable optimal perceptive absorbtion of CSI contents. This means that what is to be communicated arrives in an optimal (though in general not unique) way at the audience level. This quality is a strong function of the part of CSI contents (from the nine above positions) that is being communicated. It is not only a difficult task to realize such a communication for certain parts, given their very nature, but it is also not trivial to give a proof of the communicative success since often, there is no additional (meta)communication between audience and artists, either because the composer/interpreter is dead, or because social circumstances prevent such an information exchange, either voluntarily or by case.

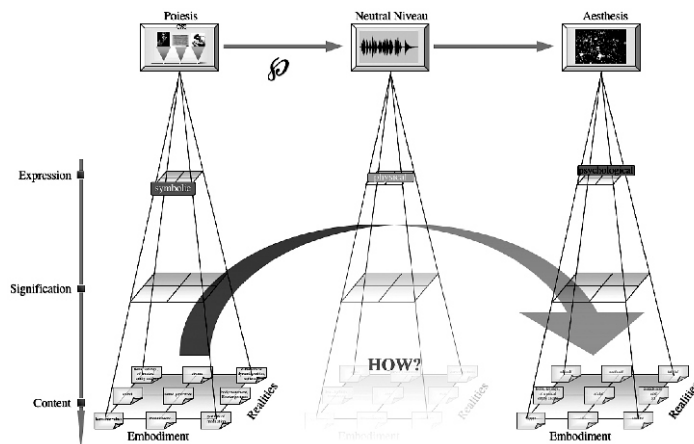


Figure 4: The rhetorical expressivity is a complex communicative process with the bottleneck in the physical (acoustical) level of the performed composition on the neutral niveau.

The composer, the score, and the interpreter (CSI) are a communicative complex in its own, and one, whose semiotic dimension is far from homogeneous. This is due to a number of semiotic determinants, in particular those stemming from dia- and synchronic distances between composer and interpreter. These extensions relate to the general fact that a semiotic system expands in space and time: for example, our language has a spatial extension over the territory where it is spoken. Regional differences at the same time refer to this spatialization. Phenomena in a semiotic system that happen simultaneously are called synchronic. The second axis of a semiotic system is time: given a spatial localization, the system's structure may depend upon the moment of time when we observe it. For example, a word has its history, also known as its etymology. The time axis of a semiotic system is called the diachronic axis.

Typically, diachronic distance (also called transversal ethnology) is significant if the composer is dead, when his/her score is being interpreted, e.g. Beethoven being interpreted by Maurizio Pollini. It is also significant via a number of noise factors blurring the transmission through time of the composer's poietic position. Often, such information is missing from the beginning since many composers do not communicate their technical, emotional, philosophical, or religious secrets of composition.

Synchronic distance (referring to ethnology) is due to simple socio-cultural separators. The performer is usually not part of the same socio-cultural region as the composer, a fact that might create different understandings of one and the same composition. But it is also due to the freedom of interpretation. The latter relates to the very nature of communication, which places the aesthetic position symmetrically to the poietic one, i.e., the composer is only understood to be the first interpreter. The interpreter has his/her own rights to understand a composition. This perspective is known in the theory of painting.

It is not mandatory to follow the composer's hints and preferences, in particular if there are fields, where such information is simply absent; for example with Bach's missing dynamic signs. Or Beethoven's missing gestural determinants, so dramatically misinterpreted by Glenn Gould in op 57. If cultural separators include distant ethnics and oral traditions, interpretation may become a dramatic distortion of the composer's intention. For example, if a score is produced that complies with extracultural standards (e.g. fixed pitch as opposed to variable pitch), the interpretation may become unacceptable to the creators.

In the communicative body of CSI, the score is a semiotic bottleneck. It is the neutral reference, but it is a poor information repertory in many regards. The

construction of the nine semiotic positions discussed above cannot be completed upon exclusively score-based analysis.

Even with support of the composer (if that one is asked for), the composition traced on the score is so rich in details of performative work that the interpreter is forced to recur to knowledge external to the composer's poietics (in particular Kofi Agawu's paratextual attributes: style etc., see [1]) and the score's neutral data. For example, the microtiming (agogics) is virtually never made explicit and has to be shaped by the interpreter alone, using a number of rationales related to the semantic fields under consideration.

The audience is the aesthetic instance “opposed to” CSI, i.e. lying on the other side of the central position of the work's neutral level. It is however somewhat underestimated in its complexity. In fact, aesthetic processing of the neutral level data is everything but elementary and, in particular, involves a symmetric construction with regard to CSI. It is therefore reasonable to call this symmetric communicative configuration ISC= Impression/Schematization/Comprehension. Here is what we intend by these three communicative positions: (1) Impression is perceptual as opposed to constructive interpretation, (2) the score is substituted by a scheme representing the perceptual body's organization, and (3) the poietic creator, the composer, is substituted by the comprehensive force rebuilding the ideas from perception, possibly, but not necessarily, in accordance with the composer's intentions.

This ISC articulation in the audience position is necessary also to explain the difference between the understanding of a naive audience as opposed to expert listeners. The latter is typically invoked by Ray Jackendoff and Fred Lerdahl in [7]. ISC is critical in the judgment of performances. A dramatic situation may, for example, arise from prejudices imported from previous performances to qualify the present one. These prejudices are typically not founded in any logic, but are just there and prevent the audience from dealing with a new type of performance in a fair way.

3 Multi-agent Communication

The above architecture of communication in performance is detailed in its poietic and aesthetic dimensions, but it is not, however, detailed with respect to the concrete embodiment of communication by humans and/or machines. It is in fact virtually never the case that the communicative agents are in the singular: Several if not many participants are involved in the performative communication of music. The communicative stream is not between two individual instances, but is distributed among a number of composers,

musicians, audience members, and even and progressively more so machines such as computers with interactive programs that may intervene in the shaping of audio output, the algorithmic transformation of compositional input, or the redistribution of musical objects by global social multi-agent music making, such as it has been proposed by Ali Momeni and coworkers [15]. The musical agents may have multiple communicative interactions, reaching from score writing composer(s) to performers, improvisers (real-time composers), listeners, and machines that share a variety of roles. It may also happen that such agents are simultaneously involved in different roles, i.e. not only composing, but also listening, reacting, and producing new compositional components.

We therefore propose to enrich Molino's communicative scheme by introducing a multiplicity of poietic, neutral, and aesthetic agents involved in the sonic realization of a musical composition. We therefore introduce a multi-agent communication matrix, see also figure 5. It consists of a series of poietic agents $P(1), P(2), \dots, P(S)$ and a series of aesthetic agents $A(1), A(2), \dots, A(T)$, which are connected to each other by neutral niveaus $N(k,l)$ from $P(k)$ to $A(l)$ for certain pairs. It is not excluded that $P(k) = A(l)$, which is the case for improvisers, for example: The same agent may be poietic and aesthetic! But we position any such agent in the poietic row position or the aesthetic column position, according to its communicative roles.

We may exhibit different functions of such agents, for example poietic composers or musicians and aesthetic composers, musicians or audience. So, for example, a poietic composer communicates to an aesthetic musician via the neutral niveau of the written score. This is one of the classical relations. But a musician may also act poietically upon a composer, when an improvised musical structure is inserted to the composition, which a composer is writing. And here, the composer might be identical to the musician in the sense that the composer acts as a musician and then processes the played music in his/her compositional creation. This is a frequent relation in jazz, but also in classical composition, where the composer switches roles during the creative process. In improvised contexts, the communicative relation from musician to musician is standard, and in computer-aided composition the computer might act upon the human composer in a poietic role.

The overall image of this network is a global field of agents interacting in possibly changing specific roles and thereby blurring the rigid classical image of a unidirectional communication from creative composers to passive performers and listeners. The musical creation in performance is a network with loops and distributed roles.

4 The Performer's Balanced Dancing Presence

Although we may develop the technical details of structure and rationales of performance later in a more detailed study, it seems important to draw a compact image of the existential focus, which is embraced by an engaged performer. We do so for two reasons: first, we would like to sketch an aspect of performance which is vital to the understanding of a performer as a living artist, and second, we want to give a first hint to those performers who rightly cannot be happy with a merely technical understanding of their artistic life. It is only a germinal hint, since the existential shape of a performing artist is far from being understood, and it is not the target of this book to unfold this deep topic. Let us shortly digress on the very difficulty of such an enterprise and have a look at the philosophy of dance as it has been addressed by Paul Valéry in [16].

A dancer is a performer par excellence since the dancer's score, written in Laban notation or any other dance score language, is known to be a poor reference to what dance is when it happens. This has been analyzed in detail in [5]. Valéry, in his treatise, recognizes that dance is more action than every other art, that it is an art of time, and one, where the artist's life is taken in its full extension as the dancer's body is fully engaged in the unfolding of dance. But then, instead of proceeding to a valid definition of dance (what he explicitly wants to achieve), he looks into the philosopher's mirror and recalls that Aristotle, Nietzsche, and in fact all philosophers are dancing with their words and thoughts. What a pirouette of thought: Valéry, instead of writing a philosophy of dance, makes philosophy dance. This dance of philosophy seems to miss the point: you usually do not represent knowledge about an object by making it. Valéry's approach does not give the expected definition of the concept of dance. Valéry however admits that dance, as an art of time, pertains to a fundamental quality of human existence, namely time. And he recalls that already Saint Augustin admitted that he knows what is time, but when asked about time, does not define it. Dance for Valéry is a similar phenomenon: Impossible to define it, you have to live it, and he lets us know that thinking is a way of dancing.

Although this insight is precious, we cannot accept the answer "Let's dance, and you will know!" to the question "What is dance?". Of course, the answer confirms that famous saying that dancing is a way of thinking. And for our concerns it is also true that performing music is a way of thinking. But here lies the problem: If it is a way of thinking, in what sense is it a special way, and different from other ways of thinking? The performer thinks music in a very specific way, radically different from the way a music theorist thinks music, see Martin Puttke's paper "Learning to dance means learning to think!" in [12].

So let us try to describe those characteristic coordinates of this way of thinking music.

It is a logical necessity to locate the performer's (oni)ontology in the framework of multi-agent communication since a performer should focus on communication among all agents. A performing musician cannot limit his/her interaction to a unidirectional messaging from composer to audience. This is an outdated casting of performative creativity to slavish service for the ingenious composer, an all too narrow perspective propagated by Arnold Schönberg and resonating in a casted performance education that produces only robots, not musicians. It is a sad fact that even in jazz school education, the slavish messaging of jazz on the basis of lead sheet changes replaces creative interaction of improvisers and disseminates that “whitened jazz” catechism, against which apart from more general social and political motives free jazz in the sixties of the 20th century was also rebelling.

Within the three dimensions of musical ontology, which complement the communicative dimension, namely realities, embodiment, and semiotics, the performer realizes a crossing of a singular type. In what follows we stress the characteristic features (which does not mean that other aspects are absent).

The most characteristic feature on the axis of realities is the interaction of two bodies: the musician's body and the body of physical sounds. Their interaction is generated on the interface of the musical instrument, whose bodily manipulation produces the music's sounds. For an acting performer, this coupling of bodies is the core neutral niveau. All other levels of neutrality might be implied or subsumed, but this one is the manifest neutral building block.

On the axis of embodiment, corresponding to the reality of instrumentally interacting bodies, the performer's focus is on gestures. It is these gestures that are communicating musical formulas or processes. It is the highest quality of musical expressivity to deploy compactified musical formulas into gestures. Gestural embodiment does not populate given spaces, but creates them, defines their extension and thereby enables fellow musicians' gestures to resonate with one's gestures in these shared spaces. “Understanding is catching the gesture and being able to continue,” this deep insight by the French philosopher and mathematician Jean Cavallès [3] is what happens in the gestural interaction among performers. Their gesturally spaced vibration is what drives their bodies to move and to shape the “body of time”. This mimetic approach is perfectly matched by the current neuro-cognitive action theory of human culture as induced by the cortical mirror neurons, see [14], [13].

On the axis of semiotics, the (successful) performer organizes the future of the

music being performed with reference to the music's past. The meaning of the music played to this present moment is connected to the shaping of the meaning of the next musical signification in a flow of thoughts. This creative transfer is performed by the body's gestural utterance. In order to achieve this in a coherent and persuasive way, we have to identify an environment, where such a strong shaping activity is executed. On the level of physical events, we cannot realize such a program, since physical time presence is a real number t_0 , past is the interval $t < t_0$, and future is $t > t_0$. Therefore classically conceived physical presence reduces to a single time point, and nothing can really happen in such a vanishing point.

This perspective is not satisfactory from the performer's point of view since the concept of presence in the time-sensitive arts cannot be reduced to zero. We do not embark in a neurophysiological model of performer's presence, also because to our knowledge there is no such a model. We rather want to postulate such a reality of artistic presence independently of a neurophysiological modeling. The fact is that the time-space of presence in artistic creation and shaping of structures is a huge environment, where all the logical decisions upon performative actions are made, be it gestural strategies, receiving and processing the structures of past musical events, the contributions from other agents in the multi-agent network, the knowledge from the symbolic score and its pre-fabricated analyses, etc. We need such a time-space that conceptually is independent of the physical time point of presence. Let us call this time-space "imaginary" for two reasons. On the one hand, it is an environment that pertains to the psychological reality and as such is imagined. On the other, and this is a speculative thought, it is known that modern physics, above all by the research of Stephen Hawking [4] in Big Bang cosmology, but also by contributions of Itzhak Bars [2] to the unification of gravitation and quantum mechanics, has introduced a second time dimension.

In Hawking's model, time is a complex number $t + i.s$, where t is the traditional time value, whereas $i.s$ is the imaginary component. It is not clear at all, whether and how this imaginary component could be a part of human consciousness. But we conjecture that our presence of consciousness, where thoughts are built and processed, could happen in that imaginary direction, so that the classical real space-time is complemented by an imaginary time-space defined by imaginary time. If physicists are entitled to introduce new time dimensions there is no reason to prevent artists to also do so and to claim that creative human consciousness is hosted in such a time-space that is "orthogonal" to the physical one. This means that at any classical physical time t , we would have an entire time-space defined by an imaginary time $i.s$ plus some space coordinates attached to that time.

Based upon such arguments for an extra time dimension, we argue that the concept of presence in time-critical arts requires such an imaginary time and space. It is this realm, where the transitional processing of past music to future music, the planning of gestural strategies, the body-instrument-sound interface are all displayed and organized. It is a quite dramatic change of understanding of what happens in the artistic presence, since it eliminates the mystification of spontaneity and imprevisibility in creative performance. These attributes have had a great influence on the non-understanding of creative performance, from classical music to free jazz: Improvisation, creative performance, all of that has been boiled down to these negative concepts: Spontaneity and imprevisibility are just negations of any sort of positively defined artistic shaping, they are the “emergent properties” of creativity mysticism, creativity by negation. Telling a musician to be spontaneous is of no help whatsoever: it is just a recommendation to rely on nothing that could be conceived of in the artistic shaping activity.

Offering to the artist the concept of imaginary time-space is a completely different affair: It opens a huge environment where the artist's consciousness can evolve and construe complex architectures and highways of shaping musical structures. The gestural complex is driven in this time-space, the flow from past to future music that is being played is driven and lived in this imaginary realm. The overall image is drawn in Figure 5.

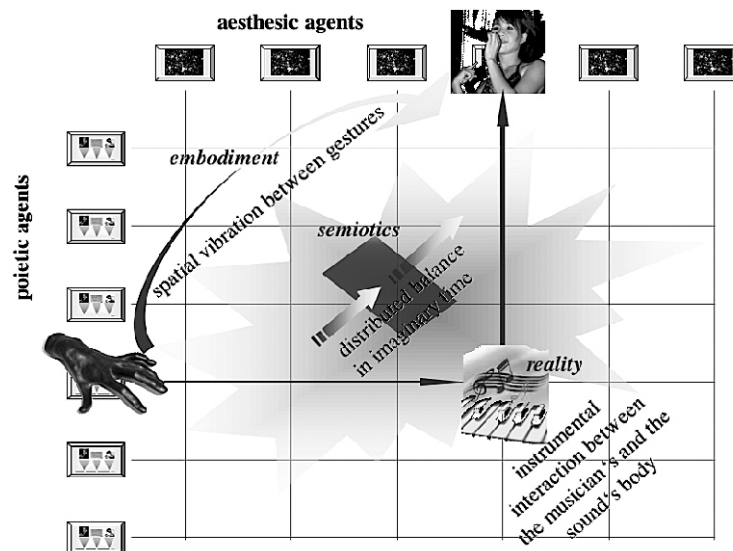


Figure 5: One could conceive musical performance as being the balanced dancing of presence, an existential intensity in three directions, embodiment, reality, and semiotics, of communicative activity.

Let us focus on the semiotic flow process connecting past music instances to the shaping of future ones. This aspect can take place in very different ways. To begin with, the performers may rely on different extensions of the past music's structure. An improviser might only listen to one other fellow-musician or he/she might listen to many of them, when shaping the next sounds. But these future sound might also be shaped by a more or less strong reference to preconceived structures, independently of the actual performance, such as, typically, the reference to a given score that tells me which notes to play next independently of how and what has been played to the moment. These factors in the shaping of future sounds are more or less distributed processes, they define an identity of the performed piece of music that is distributed among the multiplicity of agents. This is why we call this process one of a distributed identity. And why we would define the quality of the present performance as being defined by the coherence and strength of this flow as a distributed identity.

Putting the three components of performance together, we see that the pairing of body-instrument-body and of gestural space vibration could be conceived as the aspect of dance. In short: dance would be viewed as a synthesis of body and gesture. A second pairing: body-instrument-body and the flow of structural unfolding would then be seen as the balance in a bodily realm, a concept that is akin to what classical Greek aesthetics called "kairos": the perfect balance in the body's dynamics of presence. And finally, the pairing of gesture and structural flow would be understood as a shaping of the body of time, the dynamics within the imaginary time-space that defines our imaginary body of time; we might call this the presence in performance. Putting all these three pairings together, one could then conceive musical performance as being the balanced dancing of presence, an existential intensity in three directions of communicative activity.

References

1. Agawu V K: *Playing with Signs*. Princeton University Press, Princeton 1991
2. Bars I: *Extra Dimensions in Space and Time (Multiversal Journeys)*. Springer, Heidelberg 2009
3. Cavailles J: *Méthode axiomatique et formalisme*, Hermann 1938
4. Hawking S: *A Brief History of Time: From the Big Bang to Black Holes*. Bantam Books, New York 1988
5. Hecquet S and A Prokhoris: *Fabriques de la danse*. PUF, Paris 2007
6. Hjelmslev L: *La Stratification du Langage*. Minuit, Paris 1954

7. Jackendoff R and F Lerdahl: A Generative Theory of Tonal Music. MIT Press, Cambridge MA, 1983
8. Mazzola G: Geometrie der Töne. Birkhäuser, Basel 1990
9. Mazzola G: La vérité du beau dans la musique. Delatour 2007
10. Mazzola G and P B Cherlin: Flow, Gesture, and Spaces in Free Jazz Towards a Theory of Collaboration. Springer Series Computational Music Science, Heidelberg et al. 2009
11. Mazzola G: Categorical Gestures, the Diamond Conjecture, Lewin's Question, and the Hammerklavier Sonata. Journal of Mathematics and Music Vol 3, Nr. 1, 2009
12. Puttka M: Learning to Dance means Learning to Think! In: Bläsing B, Puttke M, Schack Th. (Eds.): The Neurocognition of Dance. Psychology Press, Hove and New York 2010
13. Ramachandran, Vilayanur S.: Mirror Neurons and imitation learning as the driving force behind the "great leap forward" in human evolution. <http://www.edge.org/3rd culture/ramachandran/ramachandran p6.html>
14. Rizzolatti G and L Craighero: The mirror-neuron system. Annual Review of Neuroscience 27: 169-192, 2004
15. Tanaka A, N Tokui N, and A Momeni: Proceedings of ACM Multimedia, Facilitating Collective Musical Creativity, 2005
16. Valéry P: Philosophie de la danse. Oeuvres, tome I, Variété, "Théorie poétique et esthétique", pp. 1390-1403, Gallimard, 1956

AFFINE MUSICAL TRANSFORMATIONS USING MULTI-TOUCH GESTURES

Florian Thalmann
School of Music, University of Minnesota
thalm007@umn.edu
and
Guerino Mazzola
School of Music, University of Minnesota
guerino@mazzola.ch

Abstract

We propose an elegant extension of the common vocabulary of multi-touch gestures so that they can be applied to define general affine transformations in a two-dimensional space. This new set of simple gestures is currently being implemented in a new version of the Big Bang Rubette, a gestural music composition module for the Rubato Composer music software.

1. Introduction

Any kind of musical composition can typically be situated between the two extremes of human *free improvisation* and computer-aided *algorithmic composition*. The former is characterized by gestural immediacy and spontaneous reaction to sound, whereas the latter fully delegates music creation to a computational data flow and is entirely detached from human intervention. In computer-aided composition, as opposed to the directness of improvisation, there is a commonly experienced *conceptional and temporal distance* between compositional activity and the resulting musical output (as for instance in serial composition or sound synthesis). Composers working with computer software often have to specify what they wish to do and then wait for the corresponding outcome to be calculated, before they exactly know how the result is going to sound or look like.

We already addressed this problem in previous research projects and publications [1, 2, 3] and managed to overcome significant parts of this distance. On one hand, the conceptual distance was significantly shortened by providing means for clear and versatile visualizations of musical data types (scores, notes, hierarchical scores) and navigation through these visualizations. On the other

hand, the temporal distance was eliminated by visualizing and sonifying complex musical transformations while they are being carried out. This way composers can instantly understand musically what they are doing. This paradigm of immediate behavior is what we call *gestural manipulation*; it connects the most intuitive and the most abstract thoughts in music. It is inspired by the theory of gestures, which are known to be the turning point between thinking and making [4].

The software developed in that context is the so-called *Big Bang Rubette* module (see Fig. 1) for the Rubato Composer music software [2]. It provides an intuitive graphical interface for drawing and creating hierarchies of musical notes and manipulating them with geometrical and more complex compound mathematical transformations, as for instance wallpapers or alteration [5]. The Big Bang Rubette has already been used by various composers and one of its compositions has been discussed in [6].

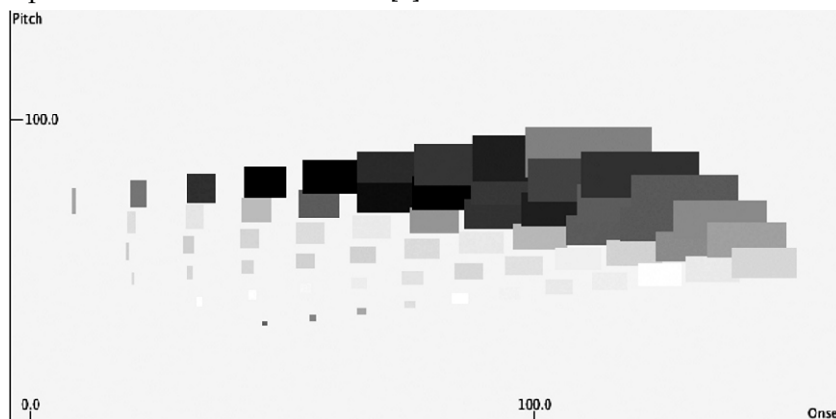


Figure 1.

The Big Bang Rubette's multidimensional visualization of musical objects.

Given the mentioned improvements, one kind of distance still remained. For complex mathematical actions to be easily understandable and controllable, the interface used to perform them has to be as intuitive as possible. The computer mouse supported in previous versions was subject to significant restrictions. For a truly gestural way of composition we do not only expect the software to behave gesturally, but also the interface. This way composers can, analogous to performers in freely improvised music, keep in touch with the algorithmic processes as these are literally made “tangible”.

A change towards such interfaces is taking place right now in the world of everyday information technology. The recent and rapidly increasing commercialization and popularization of so-called *multi-touch interfaces*

(iPhone/iPad [7], Multi-Touch Wall [8], multi-touch tables [9], etc.) brings about a paradigm shift in our daily lives. Previously abstracted data becomes suddenly accessible in a most intuitive way with controls simply using simple finger gestures such as pinches, flicks, or drags. This not only reduces the operational distance experienced when interacting with the device, but also dramatically increases the pace of one's work. Yet, we still find ourselves at the beginning of this development. In fact, the gestural controls introduced so far are merely used for simple tasks such as multi-finger drawing, distributing and sharing objects on a screen, enlarging or rotating pictures, or executing simple gesturally-linked commands such as flips and swipes. More complex operations are generally still controlled by laborious concatenations of elementary mouse-imitating controls.

In this article, we propose an extension of the standard multi-touch gestures for simple transformations of objects so that they can be applied for any type of two-dimensional affine transformations. First we will look at the most common definitions of transformational finger gestures. We will then present our extensions on a theoretical level and discuss practical issues concerning a preferably easy-to-use implementation. We will conclude our paper by describing our realization of these gestures specifically for the Big Bang Rubette.

2. Common Transformational Multi-touch Gestures

In recent years, numerous industrial parties have been designing and patenting ways to interact with multi-touch interfaces (among them Apple [10], Synaptics [11], or Elan Microelectronics [12]). These developments led to a more or less standard vocabulary of finger gestures, which consist of local simultaneous movements of multiple fingers across a multi-touch surface. Many of these gestures are used to trigger discrete events, such as Apple's three-finger flick used to navigate pages or pictures, or their two-finger twists used to rotate objects by 90 degrees.

However, in other applications, such as Jeff Han's Multi-Touch Wall [13], gestures control continuous movements to handle and transform screen objects. In the context of this article, these are the gestures that we are interested in, since we are intending to use them to control continuous transformations of musical objects. All of these gestures are usually executed by placing one or more fingers on an object in a two-dimensional plane, then moving them across the surface and finally lifting them from the surface. The most common of these gestures are, as summarized in [14],

- **drag**, where one finger is placed on an object or a group of objects and moved across the surface to perform a *translation* (Fig. 2(a)),

- **pinch**, where two fingers are placed on an object or a group of objects and moved towards or away from each other to perform a *dilation* (Fig. 2(b)), and
- **twist**, where two fingers are placed on an object or a group of objects and moved in circular clockwise or anti-clockwise motion to perform a *rotation* (Fig. 2(c)).

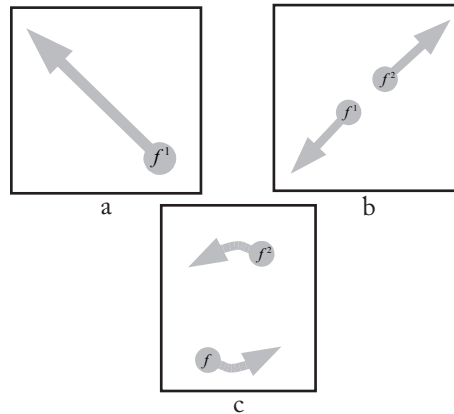


Figure 2. The three most common transformational multi-touch gestures

A similar set of gestures is typically used for *navigation* across the surface, where the fingers have to be placed on background areas (i.e. not containing any objects). Pinch and twist gestures can be used to zoom in or out of and rotate the surface with all the objects. The standard gesture for background scrolling is a two-finger drag, where two fingers move in parallel motion.

In many applications these gestures are perceived by a gesture detector, which once recognizing the intention of the user after an initial defining motion, uniquely forwards the coordinates for a specific gesture. In other words, drags, pinches, and twists often cannot be executed at the same time. In our context, however, it is essential that they can. For the maximal compatibility of these three gestures we add a two-finger object drag with the same functionality as the one-finger drag to the vocabulary.

We thus obtain a set of three two-finger object manipulation gestures that may be applied simultaneously and therefore be defined as a single, more general gesture: **two-finger transform**, where two fingers are placed on an object or a group of objects and moved around. For two starting points $p_s^i = (x_s^i, y_s^i)$ on the surface plane and two intermediate or ending points $p_e^i = (x_e^i, y_e^i)$ for $i = 1, 2$ corresponding to the locations of fingers 1 and 2, and the two vectors $v_j = p_j^e - p_j^s$ for $j = s, e$, the following transformation components are obtained:

the difference $p_e^1 - p_s^1$ defines the translation component,

$\frac{v_e}{v_s}$ forms the dilation component, and
the angle $\arccos\left(\frac{v_e}{|v_e|} \cdot \frac{v_s}{|v_s|}\right)$ forms the rotation component.

Please note that for a translation, only the points of finger 1 are necessary, which is analogous to the above-defined one-finger *drag* operation. However, for all three gestures to be united in a single two-finger gesture, we need to be able to drag an object with two fingers.

3. Requirements For An Extension

The transformations that can be realized with the above gestures form a significant part of the set of two-dimensional geometrical transformations. Since for many applications in computer science we only need to be able to change position, orientation, and size of screen objects, such as images or windows, they are sufficient in most cases. However, in other situations, we might need to be able to perform other kinds of transformations.

In the Big Bang Rubette, for instance, we want to be able to perform all possible geometric transformations with musical objects. These so-called *affine transformations* can basically be represented as a concatenation of five transformation types: *translation*, *dilation*, *rotation*, *reflection*, and *shearing*. However, the gestures we defined above though can handle neither reflections nor shearing. To be able to perform arbitrary affine transformation we thus need to extend or redefine the basic vocabulary of multi-touch gestures.

It is our goal to do this in a way as simple as possible while ensuring that the present gestures, which are doubtlessly maximally intuitive, form a special case of our general definition, so that they can be maintained as they are. It can easily be proven that all the parameters that can be taken from a two-finger gesture (central point, distance and angle) are used for the above-defined general transformation gesture. Therefore, it is necessary to add at least a third finger to be able to get more parameters. Traditionally, gestures of three or more fingers are merely used in a static way, e.g. as a parallel motion of all fingers. However, since the movement of a third finger can practically be independent from the other two, we obtain an additional degree of freedom which will prove to be sufficient for our extension.

4. Gestures For 2d Affine Transformations

For the definition of a set of gestures for general affine transformations, we would like to reconsider the *two-finger transform* definition from Section 2 and

add an additional distinction that leads to an equivalent set of gestures providing slightly more control. Let us distinguish fingers $f^1 = (p_s^1, p_e^1)$ and $f^2 = (p_s^2, p_e^2)$ as the first and the second finger to be placed on the multi-touch surface. According to the definition above, f^1 defines the translation component, while f^2 defines the vectors v_s, v_e , where v_e is the image of v_s . If we now add a third finger $f^3 = (p_s^3, p_e^3)$, we obtain two additional vectors $w_j = p_j^3 - p_j^1, j=s,e$ (see Figure 3), where again w_e can be seen as the image of w_s .

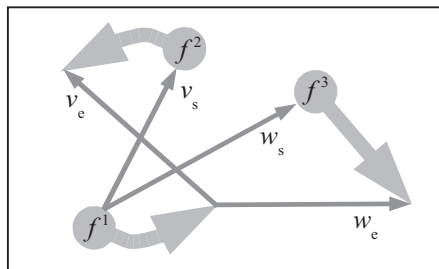


Figure 3. The vectors resulting from a three-finger gesture.

With three fingers employed this way, we can perform both missing geometrical transformations:

- if fingers f^1, f^2 remain in constant position and we move f^3 in a parallel motion to, we get a *shearing* (see Figure 4(a),
- if fingers f^1, f^2 remain in constant position and we move f^3 in perpendicular motion to v_s until we reach the point beyond with the same distance from v_s , we obtain a *reflection* (see Figure 4(b).

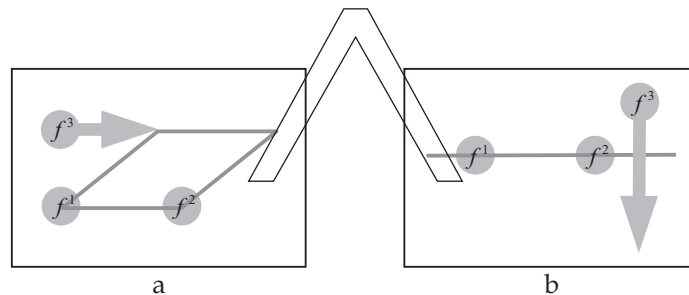


Figure 4. The two new three-finger gestures for (a) shearing and (b) reflection.

This completes the collection of transformations necessary and sufficient to perform general affine transformations by means of concatenation. However, if we move f^3 in an arbitrary motion and even move f^1 and f^2 simultaneously, we are able to define complicated affine transformations in a remarkably simple way.

We conclude this section by defining a corresponding new gesture, **three-**

finger transform, having the following transformational components, where $d^i = p_e^i - p_s^i, i=1,2,3$ and $v_j = \frac{v_j}{|v_j|}$

the difference $d1$ defines the translation component,

$\frac{|v_e|}{|v_s|}$ forms the dilation component,

the angle $\arccos(\hat{v}_e \cdot \hat{v}_s)$ forms the rotation component,

the projection length $|(d^3 \cdot \hat{v}_e) \cdot \hat{v}_e|$ determines the shearing component, and

the projection length of $|(d^3 \cdot \hat{u}_e) \cdot \hat{u}_e|$, where u_e is a vector perpendicular to, defines the reflection component.

It remains to be mentioned that there is another party suggesting the use of three-finger gestures to access additional geometric transformations. In their open source gesture library, *Gestureworks* [15] assign two locked fingers with a third finger moving horizontally to a *roll* operation (in 3D-graphics) and the same constellation with a third finger moving vertically to the corresponding *yaw* operation. This case is comparable to our three-finger gesture in smart mode but with the difference that the two locked fingers do not have an influence on the final result. Here, f^1 and f^2 define the vectors v_j which form the orientation of the shearing or the axis of reflection.

5. Practical And Usability Issues

In a previous version of the Big Bang Rubette that was still using mouse controls, we defined a number of *application modes* between which users could switch by clicking on a number of on-screen buttons. For affine transformations there was a total of five modes for each basic transformation including translation, dilation, rotation, reflection, and shearing. In each mode, a predefined minimal number of mouse clicks and drags was necessary to define transformations. A rotation, for instance, could be performed by a single click to define the center of the rotation, followed by a click and drag gesture to define the angle. During this an on-screen tool visualized the progress (see Figure 5). Now that we use multi-touch gestures, we are able to perform all these transformations within a single mode.

Nevertheless, to ensure maximal usability, we define two modes: *general mode* and *smart mode*. The former can be used to perform affine transformations without any restriction, i.e. the vectors v_j, w_j can be defined freely. The latter, however, automatically detects which transformation a user wants to perform and locks all other parameters on their respective values. Let us now look at the exact functionality more closely.

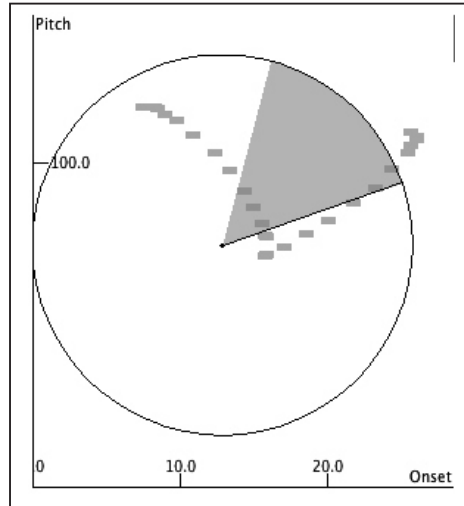


Figure 5. A rotation in the former Big Bang Rubette.

A gesture in the general mode consists of placing and dragging one to three fingers on the multi-touch surface and is not complete until the first finger f^1 is lifted. Before that, f^2 and f^3 can be placed on and lifted from the surface as many times as desired, where the respective v_j , w_j always add to the final transformation. This way of transforming particularly complies with an important characteristic of the Big Bang Rubette's principle of gestural manipulation. During such a transformation a preview of the respective result is given both visually and audibly, but is not definite until f^1 is lifted. Furthermore, if a user wishes to undo a previous transformation, the entire transformation that happened between f^1 initially being placed on and finally lifted from the surface is undone. This way a user can experiment gesturally until a favored result is attained.

The general mode enables all conceivable affine transformations, but the composer might possibly want to work with more precision and only apply specific transformations. For this, we designed a smart mode, where again the user may place and drag one to three fingers. Depending on how many fingers are placed, the following decisions are taken:

one finger is interpreted as a drag gesture

two fingers lead to a dilation, if the variation in distance $|v_e| - |v_s|$ is larger

than the circular movement, $(\frac{|v_e| - |v_s|}{2}) \arccos(\hat{v}_e \cdot \hat{v}_s)$ or else to a rotation

three fingers are interpreted as a shearing, if the above defined shearing component is longer than the reflection component, and as a reflection otherwise

As long as f^1 is still down, whenever a finger is added or removed, the software switches to the corresponding transformation, forgetting previous actions. This mode lets the user perform only one of the five basic geometrical transformations at a time.

In an earlier version we implemented similar functionality to restrict the way certain transformations worked. The former shearing mode, for instance, standardly enabled general affine transformations. With the shift key down, however, shearing could be restricted to act either purely horizontally if the mouse was mostly moved horizontally, or vertically otherwise. The smart mode of this concept could be triggered in a similar way. In our implementation, if the shift key is down when a gesture starts the smart mode is used and else the general mode.

Apart from being more intuitive and direct, the advantages of using this novel multi-touch concept for the Big Bang Rubette instead of a mouse input are the following:

we can reduce the number of application modes for geometric transformation from five to one. There is now no need for on-screen mode buttons in this context anymore
we get new transformational possibilities such as the gradual reflection
several transformations can be applied simultaneously when operating the general mode. In the extreme case we can perform a transformation consisting of all five basic geometrical transformations with a single three-finger gesture

6. Implementation: The Big Bang Rubette Goes Mt4j

For the realization of these ideas the interface of the Big Bang Rubette was completely redesigned and rebuilt. Since Rubato Composer is written in Java, we did not hesitate for long before we decided to rely on the recently initiated project MT4j [14]. This project supplies an open-source java framework that establishes connections to many of the widespread multi-touch interfaces and provides templates for standard gestural control and a modern graphical user interface built on OpenGL.

To comply with MT4j, however, the old Java Swing application interface had

to be rewritten in their language. All the screen objects, such as the user interface buttons, the display axes, or the notes and the objects denoting their relationships, had to be expressed as MTComponents and MTShapes. The benefit of this adaptation is that for every such screen object we can now decide independently what actions can be performed with it. The transformations discussed above could now not only be applied to groups of notes as before, but easily to the screen objects themselves and thus directly change the appearance, shape, or position of a single note by touching it. Furthermore, through the use of OpenGL, all doors are open for us to take the Big Bang Rubette to three dimensions, where 3D-objects could represent notes in three of the musical dimensions *Onset, Pitch, Loudness, Duration* and *Voice* at a time.

The use of MT4j and its concepts brought even more simplification to the Big Bang Rubette's user interface. The gestures are interpreted differently depending on where a user touches the screen. We thus had the possibility to implement that when two-finger pinch or drag gestures are performed on the background area, the view zooms or scrolls. This way there is no need for a special *navigation mode* anymore, as there was in earlier versions of the Big Bang Rubette. Also the *selection mode* could be eliminated, since MT4j objects can traditionally be selected with one-finger lassoing gestures performed while touching the background area as depicted in Figure 6.

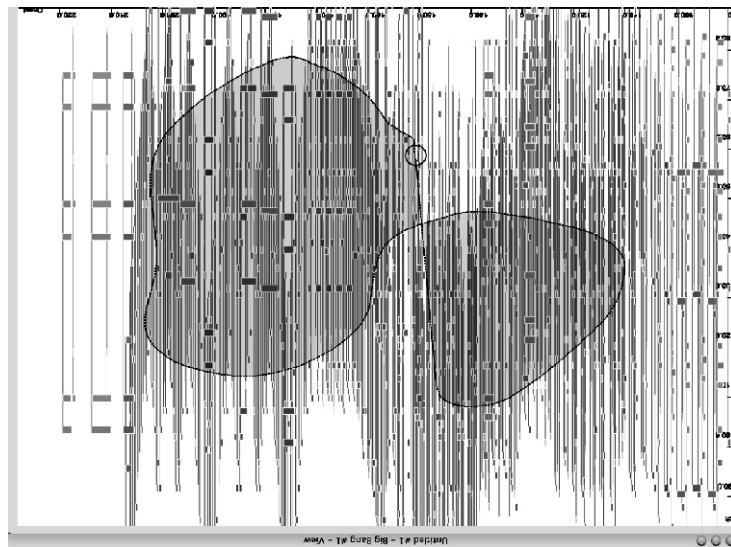


Figure 6. The new view of the Big Bang Rubette with notes being lassoed

The overall benefit of the user interface redesign based on standard multi-touch gestures and MT4j is therefore a reduction from a confusing total of eleven application modes to four modes, namely

navigation and transformation mode, where the view can be changed, and notes can be selected and transformed

drawing and shaping mode, in which new notes can be drawn with multiple fingers simultaneously. With the shift key down in this mode, the y-position of notes can be changed as described in [3]

wallpaper mode, where wallpapers can be created using the same transformation gestures as in the first mode (see [1])

alteration mode, where a composition can be altered gradually towards another one [1]. Local alteration degrees can be specified with one-finger gestures moving up or down

With all these simplifications, the process of working becomes significantly faster. It is probable that like this composers work with less accuracy and therefore make more mistakes. Our last suggestion is that since in the transformation mode we have only used gestures of at most three fingers up to now, we assign four-finger swipe left and swipe right gestures to the Big Bang Rubette's undo and redo operations. In so doing undo and redo become as quickly accessible as all other gestures and encourage composers to work in an even more spontaneous and experimental way.

7. Conclusion

In this article, we have extended the standard transformational multi-touch gestures to include all geometric as well as general affine transformations. Our first application of these gestures in the Big Bang Rubette shows how much such definitions can simplify the interaction with a software. This way musical operations become not only more intuitive and understandable but also more accessible, which means that they can be applied at much greater speed. The process of music composition here virtually reaches the spontaneity of improvisation, where decisions can be taken quickly and their implications immediately perceived. The technical distance here really seems to be reduced to a minimum and once a composer has gained some experience with such a software, he is ready to finally focus directly on inspiration and action.

8. References

1. F. Thalmann and G. Mazzola, "The bigbang rubette: Gestural music composition with rubato composer," in Proceedings of the International Computer Music Conference, (Belfast), International Computer Music Association, 2008.
2. G. Milmeister, The Rubato Composer Music Software: Component-Based

Implementation of a Functorial Concept Architecture. Berlin/Heidelberg: Springer, 2009.

3. F. Thalmann and G. Mazzola, "Gestural shaping and transformation in a universal space of structure and sound," in Proceedings of the International Computer Music Conference, (New York City), International Computer Music Association, 2010.
4. G. Mazzola, *La Verité du Beau dans la Musique*. Paris: Delatour/IRCAM, 2007.
5. G. Mazzola and F. Thalmann, "Grid diagrams for ornaments and morphing," in First International Conference of the Society for Mathematics and Computation in Music, (Berlin), MCM, 2007.
6. G. Mazzola, C. Losada, F. Thalmann, and S. Tsuda, "A boulezian creative analysis of boulez's structures," Somashekhar Naimpally, Giuseppe Di Maio (eds.): *Quaderni di Matematica, Theory And Applications of Proximity, Nearness and Uniformity*, vol. 23, 2008.
7. "Apple iphone." <http://www.apple.com/iphone/>.
8. "Perceptive pixel." <http://www.perceptivepixel.com/solutions.html>.
9. J. Y. Han, "Low-cost multi-touch sensing through frustrated total internal reflection," in Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology, (Seattle), pp. 115-118, ACM Press, 2005.
10. S. J. et al., "Touch screen device, method, and graphical user interface for determining commands by applying heuristics," January 2009. US Patent 7479949.
11. P. J. Ausbeck, "Method and apparatus for producing one-dimensional signals with a two-dimensional pointing device," November 2008. US Patent 7446754.
12. S. J. Bisset and B. Kasser, "Multiple fingers contact sensing method for emulating mouse buttons and mouse operations on a touch sensor pad," October 1998. US Patent 5825352.
13. J. Y. Han, "Multi-touch interaction wall," in ACM SIGGRAPH 2006 Emerging technologies, (New York), ACM Press, 2006.
14. U. Laufs, C. Ruff, and A. Weisbecker, "Mt4jan open source platform for mul-touch software development," *VIMation Journal*, 2010.
15. "Gestureworks. true multitouch for flash and flex." <http://gestureworks.com/about/open-source-gesture-library/>.

ABOUT THE CONTRIBUTORS

Guerrino Mazzola earned his Ph.D. in Mathematics from Zurich University, where he also qualified as a professor in algebraic geometry with Peter Gabriel and in computational science with Peter Stucki. Mazzola has profiled the European school of mathematical music theory since 1980 and has written six books on the subject, among them, *The Topos of Music*, published by Birkhäuser, and proposed by the American Mathematical Society as the mathematics book of the year, 2005. His research addresses the integration of (1) classical symbolic thinking, (2) transformational theory using category theory and (3) gestures in music. Such ideas have been operationalized in the composition program Presto (tamw.atari-users.net/presto.htm), as well as the Java program Rubato, for analysis, composition and performance. Complementing academia (he has written more than 100 scientific papers), Mazzola is an internationally acclaimed free jazz pianist. For his unifying efforts between science and culture, Mazzola was awarded the medal of the Mexican Mathematical Society in 2000. In 2006, he was elected member of the Philosophical Laboratory of Charles Alunni at the Ecole Normale Supérieure in Paris, and is a member of the research organization IRCAM in Paris since 1999. In 2007 he was elected Funding President of The Society for Mathematics and Computation in Music (SMCM). He is also the Chief Editor of the new Springer book series Computational Music Science.

Shantala Hegde is the Senior Scientific Officer in the Cognitive Psychology Unit, Centre for Cognition and Human Excellence, National Institute of Mental Health & Neurosciences (NIMHANS)- Bangalore. She is a post graduate in Psychology from Bangalore University, M.Phil in Clinical Psychology & Ph D in the area of Neuropsychology from (NIMHANS)-Bangalore. She is a specialist in the area of 'Neuromusicology/Music Cognition'. She has several research publications in both national and international peer reviewed journals. She is a Hindustani Classical Singer, undergoing training in the Agra Gharana and has immense interest in light classical vocal and ghazal singing. She is a student of the Tabla.

Jean Penny's career as flautist, researcher and educator stretches across a wide compendium of styles and arenas. Her work currently focuses on contemporary performance practices, specifically flute with electronics. Her professional background includes extensive work with many of Australia's major orchestras, recordings and concert appearances in Australia and Europe. Jean has performed many premieres of new music, and has had numerous works written for her. In 2009 Jean completed doctoral studies in music for flute and electronics at the Queensland Conservatorium. She has presented her research at Australian and international conferences and festivals, including the CMMR Re-New08 conference in Copenhagen, Denmark, the Australian Computer Music Conference - Sydney Conservatorium, The Musicological Society of Australia and at the Embodiment of Authority: Perspectives of Performance conference, Helsinki, Finland.

Ranjan Sengupta is a physicist, associated as a Consultant Scientist with Sir C V

Raman Centre for Physics and Music, Jadavpur University, Kolkata and Scientific Research Department of ITC Sangeet Research Academy, Kolkata, India. He is a PhD in Physics. He has many publications in the area of music acoustics, music cognition and music signal processing. He is a Life Fellow of the Acoustical Society of India.

Asoke Kumar Datta is an Artificial Intelligence expert, associated with the Scientific Research Department of ITC Sangeet Research Academy, Kolkata, India, as Consultant Scientist. He was the Professor and Head of the Electronics and Communication Science Unit of Indian Statistical Institute. He has many publications in the area of Speech and Music acoustics, music cognition and Speech and Music signal processing. He is a Life Fellow of the Acoustical Society of India.

Nityananda Dey is a postgraduate in Electronics Engineering, associated with the Scientific Research Department of ITC Sangeet Research Academy, Kolkata, India, as Research Scientist. He has many publications in the area of music acoustics and music signal processing. He is a Life Fellow of the Acoustical Society of India.

Soubhik Chakraborty, a PhD in Statistics, is currently working as a Reader in the Applied Mathematics Department of B. I. T., Mesra, Ranchi, India. He has published several papers in international journals in the areas of algorithmic complexity, statistical computing and music analysis and is an acknowledged reviewer. He is also a harmonium player (Hindustani classical) and the Prof-in-Charge of the SPIC-MACAY sub-chapter of B. I. T. Mesra, Ranchi.

Ms. Swarima Tewari, a B.E. in Electrical Engineering, is presently doing her M.Tech in Scientific Computing from the Department of Applied Mathematics, B. I. T. Mesra, Ranchi, India. She is doing a project in “Raga analysis using statistics and probability” as a part of her M. Tech. curriculum. She has a keen interest in music, having participated in several music related cultural events during her college days. She is an active member of SPIC-MACAY sub chapter of B.I.T Mesra, Ranchi.

Ms. Gargi Akhoury, an M. Sc. in Mathematics, is presently doing her M. Tech in Scientific Computing from the Department of Applied Mathematics, B. I. T. Mesra, Ranchi, India. As a part of her M. Tech. curriculum, she is engaged in a project related to Cyber Crime. She has a keen interest in music and has actively participated in several music related cultural events during her college days. She is an active member of SPIC-MACAY sub chapter of B.I.T Mesra, Ranchi.

Florian Thalmann has a Master's in Computer Science from the University of Bern, Switzerland. He has programmed and published on music composition procedures and mathematical music theory, especially in relation to the Rubato Composer software. He is currently attending the University of Minnesota with Guerino Mazzola to receive his PhD in Music Theory. His thesis focuses on gestural realtime methods for geometrical and topological musical composition.