NEW COGNITIVE THEORIES OF HARMONY APPLIED TO DIRECT MANIPULATION TOOLS FOR NOVICES

Simon Holland
Open University

IET, Open University, Milton Keynes, MK7 6AA UK
Telephone Milton Keynes (0908) 653296
Email address (UK & Europe) SHOL @alvey
ARPANET shol%uk.ac.rl.gm@cs.ucl.ac.uk

Abstract

Two recent cognitive theories of harmony can be exploited to design powerful direct manipulation tools to help novices sketch, analyse and experiment with harmony. Longuet-Higgins' (1962) and Balzano's (1980) theories are the focus of much current investigation by cognitive psychologists of music but they also offer considerable potential, so far virtually neglected, in music education and musician-machine interface design. Longuet-Higgins and Balzano start from quite different bases, (one from the overtone series and the other from mathematical group theory) but the two theories lead to closely related two-dimensional representations of harmonic structures and relationships. The paper discusses the design of direct manipulation tools based on a version of Longuet-Higgins' theory to allow novices to modify, sketch and analyse harmonic sequences simply and clearly by moving two-dimensional patterns representing notes, chords and key areas on a computer screen linked to a synthesizer. Such interfaces should enable novices to experiment intelligently with harmony in ways that might normally be barred to them because of lack of theoretical knowledge or instrumental skill.

1 Introduction

This research is part of a wider project to find ways of using artificial intelligence to encourage and facilitate beginners to compose music for enjoyment. The wider project is aimed at novices who may not have a formal musical education and at users outside as well as inside the formal education system. (For this reason we will use popular music and jazz illustrations, although the tools work with tonal harmony in any genre.) The research exploits two recent cognitive theories of harmony (Longuet-Higgins, 1962) and (Balzano, 1980) which give rise to principled and elegant representations for basic harmonic relationships and harmonic movement. Because of space limitations we will concentrate exclusively on interfaces based on a modified version of Longuet-Higgins' theory although closely related interfaces can be designed using adaptions of Balzano's theory.

2 Longuet-Higgins' theory

Longuet-Higgins' theory of the perception of harmony involves the investigation of an array of notes arranged in ascending perfect fifths on one axis and major thirds on the other axis[1] (fig. 1). (Readers not interested in the reasons for this could skip the rest of this section.) Longuet-Higgins' (62) theory asserts that the set of intervals that occur in Western tonal music are those between notes whose frequencies are in a ratio expressible as the product of the three prime factors 2, 3, and 5 and no others (Steedman 72). Given this premise, it follows that the set of three intervals consisting of the octave, the perfect fifth and the major third is the *only* non-redundant co-ordinate space for all intervals in musical use. We can represent this graphically by laying out notes in a three dimensional grid with notes ascending in octaves, perfect thirds and major fifths along the three axes. The octave dimension is discarded in most discussions on grounds of octave equivalence and of practical convenience for focussing on the other two dimensions (Fig 1). The theory is of great interest to cognitive psychologists of music (Sloboda 83), (Howell, Cross and West, 85)

[1] In Longuet-Higgins' presentations of the theory, and in all discussions of it in the psychological literature, the convention is that ascending perfect fifths appear on the x-axis and the ascending major thirds on the y-axis. In my discussions of educational applications I reverse this usage. The reversal originally happened accidentally, but I now maintain it in educational contexts on three grounds; firstly it allows students to switch more easily between the Balzano representation & the 12-note version of the Longuet-Higgins representation (which could both be available on a single interface for different tasks) - the x-axes coincide and the y-axes are related as if by a shear operation. Secondly it makes the dominant & subdominant areas coincide with Schoenberg's (54) dominant & subdominant regions (though of course the x-axes and the overall meaning of the respective diagrams differ). Thirdly the V-I movements that dominate Western tonal harmony at so many different levels become aligned with gravity in a metaphor useful to novices. To any readers who find this convention confusing, I apologise.

attempting to explain aspects of human musical intelligence, but we will focus here on using the theory for developing new educational tools.

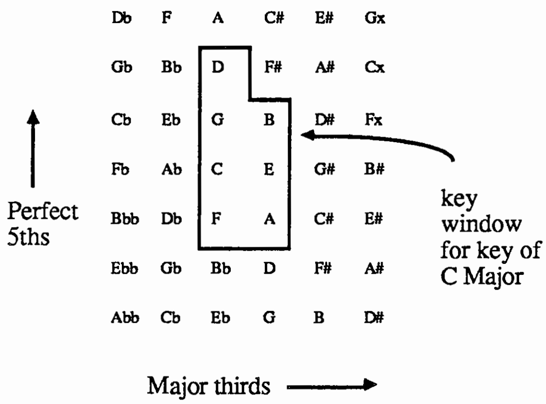


Fig 1

3.1 Representing the 'Statics' of harmony

3.1.1 Representing key areas and modulation

In diagrams such as Fig 1 all of the notes of the diatonic scale are "clumped" into a compact region. For example, all of the notes of C major, and no other notes are contained in the box or window in Fig 1. If we imagine the box or window as being free to slide around over the fixed grid of notes and delimit the set the notes it lies over at any one time we will see that moving the window vertically upwards or downwards, for example, corresponds to modulation to the dominant and subdominant keys respectively. Other keys can be found by sliding the window in other directions. Despite the repetition of note names, it is important to note that notes with the same name in different positions are not the same note, but notes with the same name in different keys (Steedman (72) calls these "homonyms"). This is an extension, motivated by Longuet-Higgins' theory, of the standard notational distinction that C double sharp , for example, is not the same note as D. (Steedman calls such pairs "homophones".)

However, for the purposes of educating novices in the elementary facts of tonal harmony we map Longuet-Higgins space onto the twelve note vocabulary of a fixed-tuning instrument resulting in what we might call '12-note two-dimensional Longuet-Higgins harmony space' or 2D harmony space for short. Consequently we lose the double sharps & double flats of fig 1, and the space now repeats exactly in all directions (fig 2). Notes with the same name really are the same note in this space. In fact a little thought will show that the space is in fact a torus, which we have unfolded and repeated like a wallpaper pattern.[2] One result of this is that instead of a single key window we have a repeating key window (fig 2).

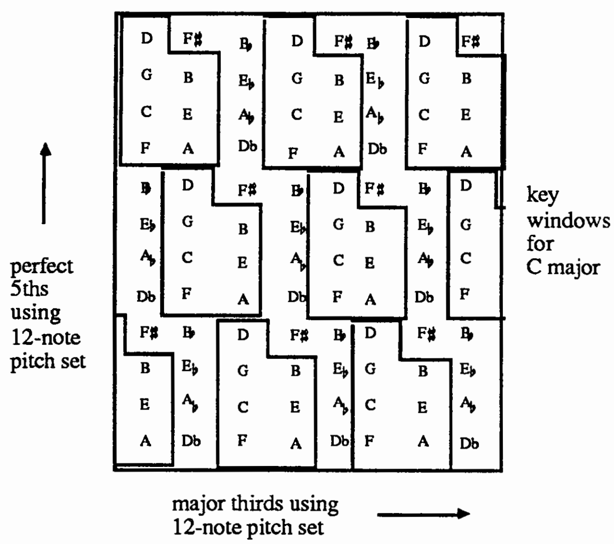


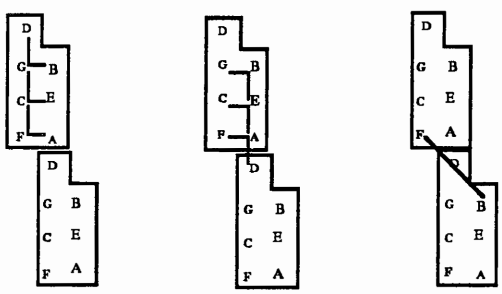
Fig 2

3.1.2 Representing chords and tonal centres

Let us now turn to look at the representation of triads and tonal centres. In 2D harmony space, major triads correspond to L-shapes (fig 3). A triad consists of three maximally close distinct notes in the space. The dominant and subdominant triads are maximally close to the tonic triad. We can instantly see from the diagram that the three primary triads contain all the notes in the diatonic scale.

[2]We have used arbitrary spellings in these diagrams (e.g.. F# instead of Gb etc.), but an environment could equally easily use neutral semitone numbers or any preferred convention. The convention could even be dynamically affected by changes in the position of the key window.

Notice also that we have a clear spatial metaphor for the centrality of the tonic - the tonic triad is literally the central one of the three major triads of any major key. We can make similar observations for the minor triads. Minor triads correspond to rotated L-shapes. Like major triads, they are maximally compact three-element objects in the space. The three secondary triads generate the natural minor (and major) scale (We could deal with harmonic and melodic minor scales by extending the key window, but we do not pursue this here). Also, the space gives a clear visual metaphor for the centrality of the relative minor triad among the secondary triads.[3] Completing the full set of scale tone triads for the major scale, the diminished triad is a sloping straight line. Some musical dialects, especially in areas of popular music consistently use seventh or ninth chords in place of triads. These chords similarly have memorable and consistent shapes in the space. See Fig 4 for the representation of scale tone sevenths.



Major triads in C major I, IV and V — Minor triads in C Major II, III and VI — Diminished chord VII

Fig 3 Triads in C major

3.1.3 Chords and keys in a direct manipulation environment

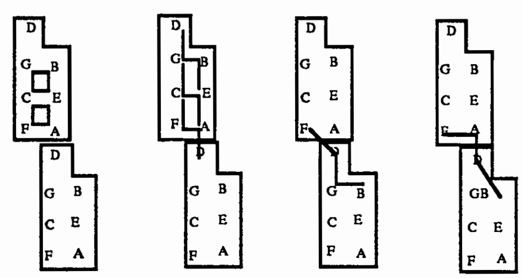
Let us pause to imagine a direct manipulation environment based on the theory as described so far. We have a grid of notes on the screen and two mice (or one mouse and a set of arrow keys). One mouse controls the location of a moving dot that sounds any notes it passes over, provided the mouse button is down at the time. Equally easily we can set the mouse to control the location of the root of a diad, triad, seventh or ninth chord. As we move the root around, the quality of the chord will change appropriately for the position of the root in the scale. There will be a clear visual metaphor for this constraint, because the shape of the chord will appear to change to fit the physical constraint of the key window. The other mouse can be assigned to move the key window. Moving this mouse corresponds to changing key. If, for example we modulate by moving the window while holding a chord root constant, the chord quality may change. Once again there will be a clear visual metaphor for what is happening since the shape of the chord will appear to be "squeezed" to fit the new position of the key window. Note that the proposed environment is linked to a synthesizer and that everything we have described can be heard.

[3]This argument is borrowed from Balzano (80). It cannot be applied in the full Longuet-Higgins space but works in the 12-note Longuet-Higgins version.

3.2 Representing the 'Dynamics' of harmony

3.2.1 Simple two and three chord movements

So far we have used 2D harmony space to look at the representation of key areas and chords. Let us now move on to look at harmonic succession and progression. A fundamental I V I progression can be seen visually as one that begins on the central major triad of the key, and then moves to a maximally close neighbour before returning home (fig 3). Similarly, progressions involving I, IV and V can be seen as oscillating either side of the tonal centre by the smallest possible step and then returning home.



Scale tone Major sevenths in C Major I and IV — Scale tone Minor sevenths in C Major II, III and VI — Dominant seventh chord V — Diminished chord VII

Fig 4 Scale tone 7ths in C major

3.2.2 Manipulating and representing Cycles of fifths

Moving onto wider chord vocabularies, progressions like II V I, VI II V I, III VI II V I etc. correspond to straight lines vertically downward in 2D harmony space with the tonal centre as their target (Fig 5). We refer to straight line motions in 2D harmony space to tonal goals as ***harmonic trajectories.*** [4] In particular, if we begin a trajectory vertically downwards in a cycle of fifths from I or IV, it turns out there are two classes of trajectory that we could make. In one case (tonal cycle of fifths) we use only notes in the key - physically this is a straight line that bends or jumps where necessary [5] to remain in the key window (Fig 5b and 5c excluding shaded points). In the other case ('real' cycle of fifths) the root moves unremittingly in a straight line down the perfect fifth axis, necessarily cutting across areas not in the key window[6] (Fig 5c shaded points). In the proposed direct manipulation environment we could sketch a tonal cycle of fifths by selecting an option that constrained the root to remain within the key window together with a second option causing the chord quality to be dynamically adjusted according to the position of the root relative to the window. Having done this, if we made a vertical straight line gesture with the mouse to play the cycle of fifths, the size of the root step could be seen and heard adjusting at one point to stay in key as the root came to the bottom edge of the key window (fig 5b or c), and the chord quality could be seen and heard flexing to fit within the key window (fig 3 and 4). This would work even if there were modulations (movements of the key window) mid chord sequence. To play a real cycle of fifths, we would simply switch off the options that constrained root position and chord quality. [7] The harmony of wide areas of Western tonal music is dominated by harmonic trajectories of chords and keys moving down the dominant axis. (See (Pratt 84) for a version of this thesis in more conventional language focussing on Bach, Schubert and Mozart). If we reverse the direction of the arrow and consider chord sequences moving vertically upwards, we have what might be called extended plagal sequences and cadences. This kind of chord sequence is occasionally used in popular dialects as in, for example "Hey Joe" (popular arr. Jimi Hendrix) as discussed by Steedman (83).

[4] in deliberate imitation of Levitt's (1985) use of the term trajectory to refer to the distinct but related phenomenon of melodic trajectory.

[5] e.g. if we are in the key of C, the root moves in a diminished fifth from F to B.

[6] For chords outside the key window, simple harmonic considerations cannot determine the default chord quality. Different pieces and different dialects use different solutions depending on the musical purpose involved .

[7] Other selectable options could include manually overriding the default chord quality at any time , holding it constant or putting it under program control.

3.2.3 Manipulating chromatic and scalic sequences

Scalic sequences (i.e. movement up and down the diatonic scale) can be represented as *diagonal* trajectories constrained to remain within the key windows (fig 6). So for example, the chord sequences I II III II I , IV III II I etc. can be represented as diagonal trajectories or diagonal oscillations. Scalic root movement occurs frequently in tonal music in short sequences and is often used in longer sequences in modal music. If the constraint is removed that the root must stay within the key window, scalic sequences become chromatic sequences (fig 7). Chromatic chord succession are widely used in some dialects of tonal music, particularly in jazz dialects. (Footnote 6 applies here too).

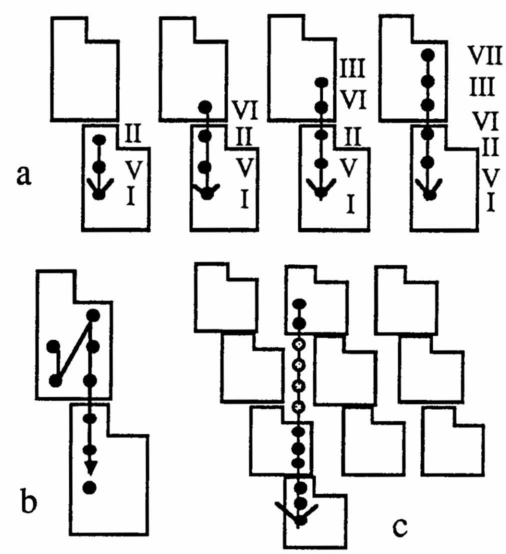


Fig 5 Cycles of fifths

4 Analysing real music in 2D harmony space

So far we have shown that a 12-note adaption of Longuet-Higgins space can provide economical descriptions of aspects of the statics and dynamics of harmony. Let us now turn to the harmonic successions of a real piece of music.[8] Fig 8 gives the chord sequence of the jazz

[8] This kind of analysis is different from and simpler than, for example Longuet-Higgins' analysis of Schubert

standard "All the things you are". Because of limitations of space we will only consider the first 8 bars, but fig 8 gives a harmony space trace of the whole chord sequence, and the analysis could easily be continued. For clarity, only the root of each chord is indicated, and chord alteration is indicated by annotation.

4.1 Example analysis

From an analytic point of view, the song breaks into a small number of recognisable harmonic plans (we only have space to deal with the first one). In the first eight bars, the sequence begins on a VI chord and makes a dominant-powered [9] trajectory towards the tonal centre (from what we know at this stage the goal being presumably the major I , though it could be the relative minor VI). But the sequence plunges on past I at the fourth bar onto IV at the fifth bar. The song at this point is in danger of breaking a standard convention for the dialect. In jazz "standards" there is a convention that we normally expect to reach a tonal goal when we hit the major metric boundary at the end of each line (normally eight bars). It is as though we were shooting for a tonal goal but overshot it. The solution used becomes in effect the harmonic motif of the whole piece - we move the goalpost. This is achieved by a timely transient modulation allowing the progression to reach the tonal goal in the nick of time[10]. In the direct manipulation environment, the 'moving goalpost' metaphor is demonstrated literally. The environment would physically show (fig 8) the 'goalpost' in the shape of the key window being moved sideways so that the tonal cycle of fifths drops into the goal or tonal centre at the audible metric boundary. The broad outlines of this analysis should be immediately comprehensible to a novice with access to the direct manipulation tools being discussed.

It is important to note that a visual formalism is not being proposed as a substitute for listening. It is being suggested that an animated implementation of the formalism linked to a sounding instrument may allow novices without instrumental skills to gain experience of controlling and analysing such sequences without knowledge of standard theory and terminology . But such an environment would be also be a good place to learn music theory if the novice desired.

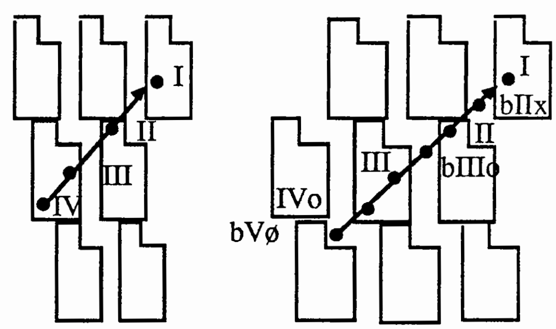
(62) and Steedman's analysis of Bach (72) using the full space.

[9] Pratt's (84) phrase.

[10] It is common in this dialect where all chords are routinely played as sevenths to emphasise arrival at the tonic (with restless chord quality major seventh) by repeating it as a more stable major sixth.

5 Educational use

The direct manipulation design considerations so far discussed could be the basis of a family of environments for analysis, modification, playback and sketching of harmonic sequences as well as an aid to learning theory. As a sketching device, multiple mice could control, for example, independent melody, bass and accompaniment voices. The system could permit dynamic association of varied rhythmic figures, Alberti patterns and arpeggiation with given voices. Analysis could be carried out at low or high level. Tonal centres and roots could have been already identified and the interface used to help illuminate the higher level harmonic structure of a piece as in Fig 8. Alternatively the student could use a version of the interface as an aid to help identify tonal centres, roots, modulations and perform the graphic equivalent of traditional harmonic analysis. Modification involves taking an existing piece and manually altering its annotated graphic trace to discover where small changes make big changes to the musical sense and vice versa. 2D harmony space tools could be used as valuable aids for studying practically any theoretical aspect of tonal harmony (e.g. the relationship between modal harmony and tonal harmony, aspects of the evolution of major/minor tonal centres etc.) The environment could automatically convert 2D harmony space displays into Common Music Notation and vice versa to assist this.

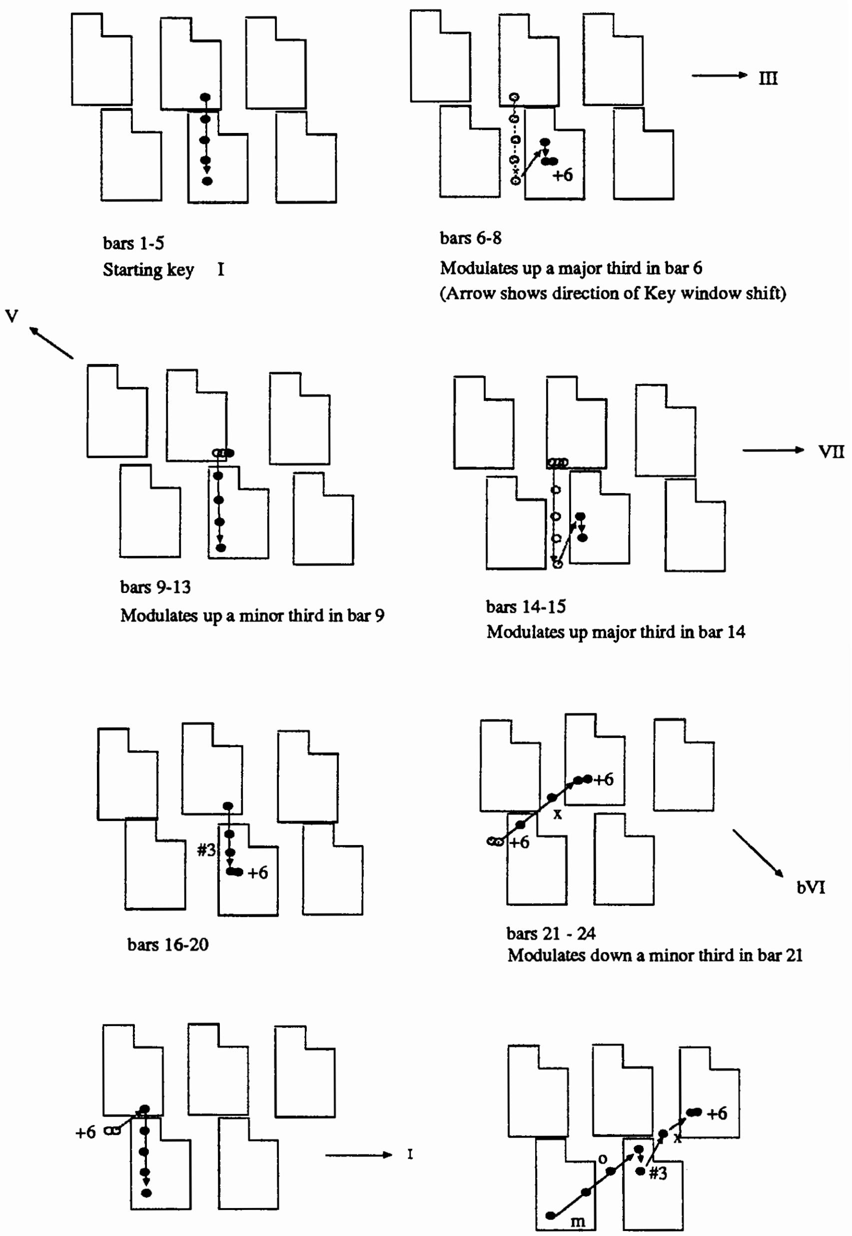


Figs 6 and 7 scalic and chromatic progressions

6 Problems and partial solutions

The ideas described so far give rise to a number of issues

"All the things you are" chord sequence in 12-tone 2D Harmony space



bars 1-5
Starting key I

bars 6-8
Modulates up a major third in bar 6
(Arrow shows direction of Key window shift)

bars 9-13
Modulates up a minor third in bar 9

bars 14-15
Modulates up major third in bar 14

bars 16-20

bars 21 - 24
Modulates down a minor third in bar 21

bars 24-28
Modulates up a major third in bar 24

bars 29-36

(Chord sequence reproduced from Mehegan (1959)). 'All the things you are' by Kern-Hammerstein © Chappell & Co.,Inc and T.B. Harms Co.

(Ab) VI / II / V / I / IV / (C) V / I / I+6 /
(Eb) VI / II / V / I / IV / (G) V / I / VI / II /
(G) V#3 / I / I+6 / (E) II / bIIx / I / I+6 /
(Ab) VI / II / V / I / IV / IVm / III / bIIIo / II /
(Ab) V#3 bIIx / I+6 / I+6 //

See appendix for explanation of chord symbol notation

Fig 8

which we can only outline here. These issues are all addressed in the research in progress of which this work forms part. The first problem is that to understand a chord sequence you usually need to know about its metrical context (as in the analysis above). A graphic notation has been developed to help address this problem. The second problem is that in a practical system, some means of controlling and displaying inversions and pitch register is needed. Some partial solutions have been devised. Thirdly, we have so far emphasised vertical aspects (in the traditional sense) at the expense of linear aspects of harmony. To a large extent this is an inherent limitation of 2D harmony space, but contrasting display strategies have been devised with partial success in special cases to selectively emphasise linear or chordal aspects of harmonic movement.

There is a potentially far more serious problem that fortunately has a satisfactory solution. The problem is that in a practical interface whose notes are not artificially constrained to fall within one octave, the diagonal (chromatic /scalic) axis will sound as a major seventh axis. Fortunately this problem can be addressed by using Balzano's representation in place of Longuet-Higgins' but bringing across the idea of a movable key window. Balzano's theory leads to a note array similar to the 12-note version of Longuet-Higgins but with axes of major third and minor third. One diagonal axis turns out to be the cycle of fifths and the other diagonal axis a true chromatic axis. It turns out that the Balzano representation works satisfactorily both theoretically and practically for the kind of interfaces we are discussing. Unfortunately space limitations do not permit us to discuss this class of interfaces and the interesting issues that emerge comparing the suitability of the two kinds of environment for different tasks.

7 New developments

2D harmony space is perhaps best viewed not as one tool but a family of tools. Other members of the family investigated include a rubber-band "MacDraw-like" version of the environment and a "Turtle Logo-like" programming language to control "harmonic turtles" in 2D harmony space. From an Artificial Intelligence and Education point of view, the 2D harmony space family of environments can be viewed as examples of what are known as discovery learning environments or microworlds. In order to tackle the problem common to such environments (Elsom-Cook, 84) of providing guidance tailored to individuals, 2D harmony space is being linked to an intelligent knowledge-based tutor for music composition under design as discussed in Holland (87). The linked system will then become a guided discovery learning environment for aspects of music composition in a variety of idioms.

8 Two further perspectives.

One useful perspective can be provided by an analogy between Turtle Logo which embodies the mathematically pervasive 20th century concept of the function and harmony space. A 2D harmony interface would embody a similarly powerful and appropriate representation for relationships in tonal harmony.

A second perspective is the Xerox Star/Macintosh human computer interface analogy. Design aims for the human machine interfaces of these machines included consistency, simplicity , reduction of short -term memory load and exploitation of existing knowledge. A 2D harmony space tool goes some way towards meeting these design aims for a harmony sketcher, since harmonic constraints and the relationships between keys are visually and physically externalised consistently to a higher degree than in the case of stringed instruments, keyboard instruments and common music notation.

9 Implementation

An early partial prototype of harmony space was implemented in Common Lisp using the interface design tool Dialog© on an Apollo Domain workstation controlling a Yamaha TX816 synthesizer via a Hinton MIDIC RS232 to MIDI converter in Jan 1987.

Conclusion

Longuet-Higgins' (1962) and Balzano's (1980) theories are the focus of much current investigation by cognitive psychologists of music, but they also offer considerable potential in music education and musician-machine interface design. We have discussed how a family of direct manipulation tools based on the theories can be designed to allow novices to modify, sketch and analyse harmonic sequences simply and clearly by moving two-dimensional patterns representing notes, chords and key areas on a computer screen linked to a synthesizer. Such interfaces should enable novices to sketch, analyse and experiment productively with harmony in ways that might normally be barred to them because of lack of theoretical knowledge or instrumental skill. The report from which this paper stems (Holland 86) appears to be the first discussion of educational use of Longuet-Higgins' theory, and its use for controlling as well as representing music.

Acknowledgements

Thanks to Mark Elsom-Cook for constant support, encouragement and guidance. This paper would not have existed without Mark Steedman's suggestion that Longuet-Higgins' theory was a good area to explore for educational applications. Thanks to Tim O'Shea for making it possible. Thanks to Mark Steedman, John Sloboda, Trevor Bray and Richard Middleton for comments on an earlier draft. Thanks to Christopher Longuet-Higgins, Ed Lisle and Mike Baker for valuable discussions. Thanks to Mark Elsom-Cook for help beyond the call of duty with specialised graphics, i/o and interface programming. Thanks to Caroline, Simon & Peter for three things more important than 2D harmony space. Not all criticisms have yet been adopted, so responsibility for all errors lies definitely with the author. The support of this work by the ESRC is gratefully acknowledged.

References

Elsom-Cook, M. (1984) Design Considerations of an Intelligent Tutoring System for Programming Languages. Phd, University of Warwick

Holland, S. (1986) Design Considerations for a human-computer interface using 12-tone three dimensional harmony space to aid novices to learn aspects of harmony and composition. CITE Report No. 7, Open University, Milton Keynes.

Holland, S. (1987) A knowledge-based tutor for music composition. CAL Research Group Technical Report No. 65, Open University, Milton Keynes.

Howell, P., Cross, I. and West, R. (1985) Musical Structure and Cognition. Academic Press, London.

Levitt, D. (1985) A representation for Musical Dialects. Unpublished Phd M.I.T.

Longuet-Higgins, H.C., (1962a) Letter to a Musical Friend. In Music Review, August 1962 pp. 244-248

Longuet-Higgins, H.C., (1962b) Second letter to a Musical Friend. In Music Review, Nov. 1962.

Mehegan, J. (1959) Jazz Improvisation I: Tonal and Rhythmic Principles. Watson-Guptil Publications, New York.

Pratt, G. (1984) The Dynamics of Harmony. Open University Press, Milton Keynes.

Schoenberg, A. (1954) Structural functions of harmony. Ernest Benn Ltd, London

Sloboda, J. (1985) The Musical Mind: The Cognitive Psychology of Music. Clarendon Press, Oxford.

Steedman, M. (1972) The formal description of musical perception . Unpublished Phd, University of Edinburgh.

Steedman, M. (1983). A generative grammar for jazz chord sequences. In Music Perception (2,1).

Appendix: Chord symbol conventions

Roman numerals representing scale tone triads or sevenths I II II IV V etc. are written in capitals, **irrespective** of major or minor quality. Roman numerals represent triads of the quality normally associated with the degree of the tonality (or modality) prevailing. We call this quality the "default" quality. In the jazz example, Roman numerals indicate scale-tone sevenths rather than triads. The following post-fix symbols are used to annotate Roman chord symbols to override the chord quality as follows ; x - dominant, o - diminished, ø - half diminished, m - minor, M - major. The following post-fix convention is used to alter indicated degrees of the scale; "#3" means default chord quality but with sharpened 3rd, "#7" means default chord quality but with sharpened 7th etc. The following post-fix convention is used to add notes to chords e.g. "+6" means default chord quality with added scale-tone sixth 6th. The prefixes # and b move **all** notes of the otherwise indicated chord a semitone up or down.