Journal of New Music Research

ISSN: 0929-8215 (Print) 1744-5027 (Online) Journal homepage: https://www.tandfonline.com/loi/nnmr20

A general pitch interval representation: Theory and applications

Emilios Cambouropoulos

To cite this article: Emilios Cambouropoulos (1996) A general pitch interval representation: Theory and applications, Journal of New Music Research, 25:3, 231-251, DOI: 10.1080/09298219608570706

To link to this article: https://doi.org/10.1080/09298219608570706

Published online: 03 Jun 2008.

Journal of New Music Research, Vol. 25 (1996), pp. 231–251 0928-8215/96/2503-231$12.00

A General Pitch Interval Representation: Theory and Applications

Emilios Cambouropoulos

ABSTRACT

Pitch and pitch-intervals are most often represented — in the western tradition — either by the traditional pitch naming system and the relating pitch-interval names, or as pitch-classes and pitch-class intervals. In this paper we discuss the properties, relationships and limits of these two representations and propose a *General Pitch Interval Representation* (GPIR) in which the above two constitute specific instances. GPIR can be effectively used in systems that attempt to represent pitch structures of a wide variety of musical styles (from traditional tonal to contemporary atonal) and can easily be extended to other microtonal environments. Special emphasis will be given to the categorisation of intervals according to their frequency of occurrence within a scale. Two applications of the GPIR will be presented: a) in a system that transcribes melodies from an absolute pitch number notation to the traditional staff notation, and b) in a pattern-matching process that attempts to discover repetitions within a melody.

INTRODUCTION

Many computer-assisted analytic/compositional systems represent pitch intervals as the number of semitones they contain. Some other systems that deal with the tonal system, use the traditional pitch-interval naming system. In this paper we will examine the possibility of devising a general representation that can be adapted to different scaling environments according to the musical task at hand.

A major difference between the traditional pitch-interval system and the semitone interval system relates to the degree by which each system allows explicit representation of different categories of intervals. On one hand, the traditional interval system allows multidimensional encoding of intervals in terms of scale degree distances (e.g., 2nd, 6th etc.), different sizes within the scale degree distances (e.g., major, minor, perfect, diminished, etc.) which further group together in the following way: perfect; major, minor; augmented, diminished. Thus, the traditional system allows explicit representation of different classes of intervals that relate to established hierarchies and functions. On the other hand, the semitone interval system abolishes any such possibility by representing all intervals unidimensionally and thus is adequate for the representation of 12-tone atonal pitch structures.

Various studies of music cognition (Deutsch 1982b, 1984; Bharuch 1984a,b; Sloboda 1985; Dowling & Harwood 1986; Krumhansl 1990; McAdams 1989) suggest that most musical systems establish different degrees of hierarchic taxonomies amongst the various musical elements that facilitate cognitive processing of a musical structure. In this paper we will examine one facet of such hierarchies, namely the hierarchic organisation of the pitches and pitch-intervals of a scale or set of scales over the full space of discrete pitch elements available in a given musical system.

Two enharmonic intervals in a tonal musical domain are very different although they consist of exactly the same number of semitones. The reason for this distinction lies in the structural properties that are assigned to each interval depending on the structural context in which they appear. For example, an isolated interval of three semitones can be heard in the tonal domain either as a minor 3rd or an augmented 2nd. If this same interval is preceded and followed by a semitone, it is recognised as an augmented 2nd interval, as this specific sequence is encountered only on the 5th degree of a harmonic minor scale. Our mind tries to match the heard sequence to the learned scale schemata of the major-minor system in an attempt to place the sequence in a higher level tonal framework. In the case of the above sequence, our mind makes a first selection, placing the sequence in the minor scale and considering the last note of the sequence as the tonic. As new intervals are encountered the first assumption is either reinforced or altered (if the new data give evidence that a better selection can be made).

The structural/functional properties of intervals within larger pitch schemata allow a finer classification than the one made if only their physical properties[1] are taken into account. This way, the 3 semitone interval can be further subdivided into the 3rd minor class and the "rare" and very characteristic augmented 2nd class allowing, thus, an explicit representation of intervallic properties that relate to more abstract tonal schemata.

Such structural properties may either be explicitly represented in a pitch representation of a specific musical system, or may be left to be implicitly inferred by other processes. Depending on the musical task at hand, a more refined representation may be more efficient (despite its seeming redundancy at the lowest pitch level) as it allows higher-level musical knowledge to be represented and manipulated in a more precise and parsimonious manner.

Brinkman (1990), in his discussion of encoding pitch and pitch intervals for computer applications, proposes a *binomial* system whereby he brings together the 12 pitch class set theory (Forte 1973; Rahn 1980) and the diatonic set theory (Regener 1964; Clough 1979, 1985). The latter suggests that the 12-tone pc-set formalism can be applied to the seven diatonic name classes; an integer from 0–6 stands for each letter-name (C → 0, D → 1, ... A → 6) and a modulo 7 mathematical formalism is developed. In the binomial system each pitch is represented by an integer couple the first of which is pitch-class and the second name-class (e.g., following the form [pc, nc] the note G♯ is [8,4] and A♭ is [8,5]). Pitch intervals are encoded in a similar manner (e.g., augmented 2nd is [3,1] and minor 3rd is [3,2]). This representation enables encoding of enharmonic pitches and pitch intervals. Brinkman proceeds in developing a set of mathematical operations that can be performed between the elements of the binomial system.

Following this direction of investigation, we will attempt to propose a General Pitch Interval Representation (GPIR) that can be applied to any M-tone scale set over an N-tone equal-tempered discrete pitch space ($M \leq N$). In the GPIR system the modality of a name-class interval is explicitly represented by the introduction of a separate symbol which is calculated from its frequency of occurrence – relating to Browne's theory (Browne 1981) on the importance of intervallic rarity. It will be shown that both the 12-tone and the traditional diatonic representations are conveniently accommodated within the GPIR and that this general-purpose representation efficiently expresses a wide range of other scale environments that may illustrate a varying degree of hierarchical organisation.

A GENERAL PITCH INTERVAL REPRESENTATION (GPIR)

In this study we will deal with equal-tempered scaling systems and more specifically with the 12-tone equal-temperament. The only equivalence assumed is octave equivalence under which any two pitches separated by a number of octaves are considered structurally equivalent (the octave equivalence assumption is an essential part of most musical systems). All other kinds of equivalence (e.g., inverse interval equivalence) are not embodied explicitly in the GPIR but can easily be inferred by the use of simple operations on the GPIR primitives.

Pitch representation

In the proposed system two pitch symbols are directly related to the structure of a scale. The first is taken from a set of integers that is used to represent the scale tones. The number of elements of this set is equal to the number of scale tones (i.e., 7 integers for 7-tone scales, 8 for 8-tone scales and so forth). Integer 0 is mapped onto note C of the diatonic system. This integer representation is a natural extension of the diatonic name-class representation discussed above. The second symbol is selected from a set of modifiers-accidentals. For these we use positive integers to stand for sharps, zero for natural and negative integers for flats (e.g., –2 → ♭♭, –1 → ♭, 0 → ♮, 1 → ♯, 2 → x). In Table 1 the traditional accidental symbols are used for matters of readability[2].

Table 1.

Traditional representation:													
7-tone diatonic scale	C	♯/♭	D	♯/♭	E	F	♯/♭	G	♯/♭	A	♯/♭	B	(C)
GPIR representation:													
7-tone diatonic scale	0	♯/♭	1	♯/♭	2	3	♯/♭	4	♯/♭	5	♯/♭	6	(0)
pentatonic scale	0	♯/♭	1	♯/♭	2	♯/♭♭	x/♭	3	♯/♭	4	♯/♭♭	x/♭	(0)
octatonic scale	0	♯/♭	1	2	♯/♭	3	4	♯/♭	5	6	♯/♭	7	(0)
12-tone scale	0	1	2	3	4	5	6	7	8	9	10	11	(0)

In the GPIR every pitch is represented by an array of the sort [nc, mdf, pc, oct] where *nc* (name-class) takes values from {0, 1, 2, ...M} for an M-tone scale, *mdf* (modifier) take values from {–u, ..., –1, 0, 1, ... u} and u is the number of pitch interval units in the largest scale-step interval, *pc* (pitch-class) takes values from {0, 1, 2, ...N} for an N-tone discrete equal-tempered pitch space and *oct* is octave range (middle C octave is 4). For example, in the diatonic system D4 is [1, 0, 2, 4], D♯4 is [1, 1, 3, 4], Eb5 is [2, –1, 3, 5] and Gb3 is [4, –1, 6, 3]. Enharmonic notes are represented with different arrays although enharmonic equivalence can be identified through the *pc* entry. In the 12-tone system D4 is [1, 0, 1, 4], D♯4 is [3, 0, 3, 4], Eb5 is [3, 0, 3, 5], Gb3 is [6, 0, 6, 3] and the first two entries become redundant as *nc* is identical to *pc*, and the modifier symbol is always 0.

This representation can easily be applied to any other equal-temperament scaling systems as, for example, the twelfth-tone Aristoxenian system (Aristoxenos 1989; Xenakis 1992. See Appendix for a brief presentation and discussion).

Before ending this section on pitch representation, we will briefly address some issues concerning the transcription of a piece of music from a traditional system of pitch notation (Western or otherwise) to the proposed GPIR, and the inverse. In general, the relation that allows conversion of a pitch structure from an M-tone to an N-tone representation (where M-tone is a subset of N-tone), is a mathematical *function*, i.e., for every element of the M-tone set there is one and only one element of the N-tone set that corresponds to it. In this case, transcription can be uniquely defined and realised (see Fig. 1).

When a pitch structure represented by an M-tone notation is converted to an N-tone notation and the M-tone is not a subset of the N-tone notation, the conversion relation is not a function and thus transcription is not a uniquely defined process (e.g., note 1 of the 12-tone scale can be either transcribed as C♯ or D♭ in the 7-tone diatonic scale). In this case, additional rules are necessary to allow selection of one possible transcription over another. This issue will be addressed in the section entitled Transcription of melodies based on GPIR..

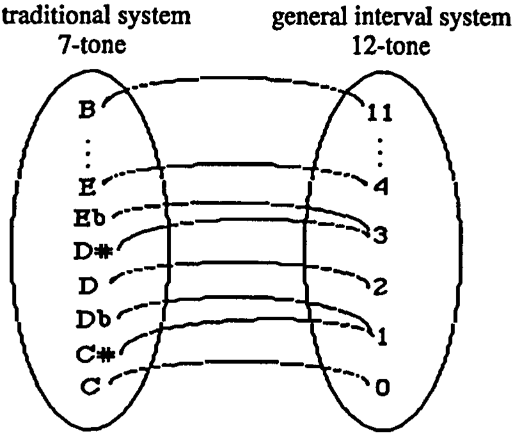


Fig. 1.

Pitch interval representation

The structure of a scaling system affects the functions and properties that may be assigned to other musical quantities, such as pitch intervals, that directly relate to it. In the GPIR two interval symbols are directly relating to inherent properties of a given scale:

1. *Name-class interval (nci):* this integer indicates the number of scale steps that an interval consists of and is calculated as the modulo M difference between the name-class integers (for an M-tone scale). Taneiev (1902/1962, pp.25–33) first introduced a similar way of naming intervals wherein the symbol 1st was used for the scale step interval — not 2nd as in the traditional interval system (this facilitates direct mathematical operations between intervals, such as addition and subtraction e.g., 1st + 4th = 5th). For a 7-tone scale the name-class intervals are depicted in Fig. 2.

2. *Modality*: the second interval symbol is determined by the *frequency of occurrence* of each member of the subset of intervals that are relating to the *nci*

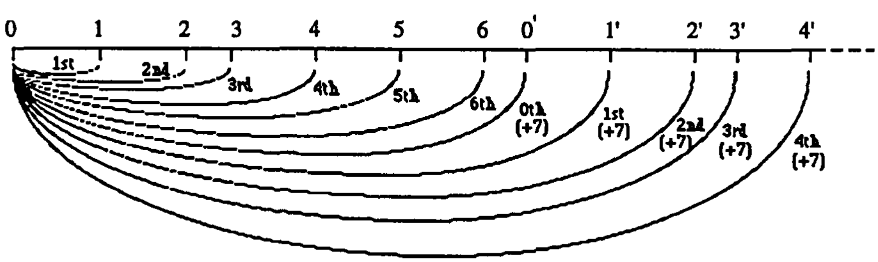


Fig. 2.

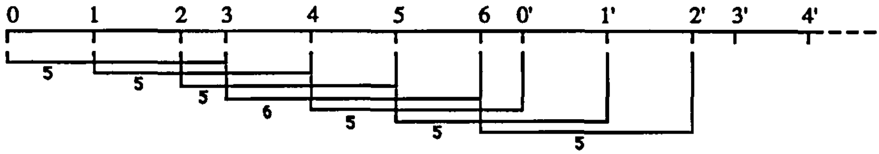


Fig. 3.

integer. If we calculate the number of times that all the different modalities of a specific name-class interval occur within a scale (taking as its lower note each degree of the scale), we can classify intervals depending on their frequency of occurrence[3]. For example, the interval of a fourth in the diatonic genre (Fig. 3) occurs 6 times at the size of 5 semitones (frequency of occurrence F = 6/7 = 0.86) and once at the size of 6 semitones (F = 1/7 = 0.14).

Table 2 illustrates the name-class intervals (as 1st, 2nd etc.), the frequency of their occurrences and the interval size in semitones (top row) for different kinds of genres of scales.

The naming process of the traditional interval system, wherein fourth is called perfect when it contains 5 semitones and augmented when it contains 6 semitones, seems to correspond to the above observation concerning the frequency of occurrence of intervals[4], i.e., perfect intervals occur most frequently between the degrees of the scale whereas augmented are rare.

The problem in defining the second symbol is the definition of the limits that will classify name-class intervals into different kinds. As a default we propose to have 3 classes (borrowed from the traditional system) defined by two symmetric limits (see Fig. 4 where x = 0.25 – this is an arbitrary selection of a limit that seems to work well for our purposes; further research may define a better value or range of values for limit x).

The frequency of occurrence of a scale interval of a specific size over the total number of scale degrees on which it can be based is F = n/N, where n = number of occurrences for that interval size and N is total number of scale degrees. For this limit (i.e., lower limit = 0.25 and upper limit = 0.75), class A contains at maximum one member (as each nci may occur only in one modality with a frequency over 75%), class B maximum four elements and class C maximum N elements. So, in general: class A = {A}, class B = $\{B_1, B_2, B_3, B_4\}$ and class C = $\{C_1, C_2, ..., C_N\}$. Intervals that do not appear between scale tones may be encountered between scale tones and nonscale tones or between nonscale tones. For these intervals, the modality symbol is selected from class D.

Table 3 depicts the resulting two-symbol names for all the intervals of the genres of scales presented in Table 2. Some comments on Table 3 are presented below:

Table 2.

Number of Semitones	1	2	3	4	5	6	7	8	9	10	11
	1st				3rd			5th			
Major Scale (t t s t t t s)	0.29	0.71			0.86	0.14		0.43	0.57		
			0.57	0.43		0.14	0.86			0.71	0.29
			2nd			4th				6th	
	1st				3rd			5th			
Asc. Mel. Minor Scale (t s t t t t s)	0.29	0.71		0.14	0.57	0.29		0.43	0.57		
			0.57	0.43		0.29	0.57	0.14		0.71	0.29
			2nd				4th			6th	
		1st			3rd			5th			
Harmonic Minor Scale (t s t t s tr s)	0.43	0.43	0.14	0.14	0.57	0.29		0.43	0.57		
			0.57	0.43		0.29	0.57	0.14	0.14	0.43	0.43
			2nd				4th			6th	
		1st					3rd				
Pentatonic Scale (t t tr t tr)		0.6	0.4				0.8	0.2			
				0.2	0.8				0.4	0.6	
				2nd					4th		
		1st				3rd				5th	
Blues scale (tr t s s tr t)	0.33	0.33	0.33	0.17	0.17	0.33	0.17	0.17	0.33	0.33	0.33
		0.17	0.17	0.17	0.5		0.5	0.17	0.17	0.17	
			2nd					4th			
	1st			3rd		4th			6th		
Octatonic Scale (t s t s t s t s)	0.5	0.5		0.5	0.5	1			1		
			1				0.5	0.5		0.5	0.5
			2nd				5th			7th	
		1st				3rd				5th	
Whole-tone scale (t t t t t t)		1				1				1	
				1				1			
				2nd				4th			
	1st		3rd		5th	6th		8th		10th	
12-tone Scale	1		1		1	1		1		1	
		1		1			1		1		1
		2nd		4th			7th		9th		11th

s: semitone, t: tone, tr: tri-semitone

frequencies	0	x		1-x	1
		(limit 1)		(limit 2)	
	$0 \leq$ class C $\leq x <$		class B	$< 1\text{-}x \leq$ class A ≤ 1	

Fig. 4.

a. In the octatonic scale there exist three class A intervals one of which is the tritone. There are no class C intervals ("rare" intervals).
b. The 12-tone scale[5] and the whole-tone scale consist only of class A intervals, and thus, the modality symbol becomes redundant and may be dropped altogether. For the chromatic scale the *nci* integer coincides with the *pci* (pitch-class interval) integer (e.g., the 4th interval is identical to the 4 pc-interval and consists of 4 semitones). One can see that the pitch-class interval representation is an instance of the proposed general system.
c. For the diatonic genre (including the major and natural minor scale) the traditional interval names emerge if the following "traditional" symbols are used: class A = {perf}, class B = {min, maj}, class C = {dim, aug}.
d. For the ascending melodic and the harmonic minor scales naming of intervals is somewhat different from the traditional system (e.g., 3rds and 4ths have a class B modality instead of class A). One may notice though that these scales hardly ever appear exclusively on their own. They are an integral part of a wider major-minor framework (even a piece of music that is composed solely on the harmonic minor mode cannot eliminate the significance obtained from the absent "opposite" major mode). If we weight[6] each kind of scale (e.g., 4 × major scale, 1 × natural minor, 1 × desc. melodic minor, 1 × asc. mel. minor and 2 × harmonic minor, add all occurrences for each interval and divide by 9) we arrive at the results depicted in Fig. 5a.
   From this weighted frequency of occurrence values we derive all the traditional interval names for the major-minor scales (Fig. 5b).
   It is obvious that the traditional interval representation is only an instance of the proposed general system.
e. "Blending" different scales together seems to be a useful method of obtaining a broader interval representation. The use of more than one genre of scales is commonly employed in some musical styles. Such scales usually exhibit a similar interval "character" i.e., they have a similar frequency of occurrence for all intervals or the most important ones. In Fig. 6 the similarity between the major-minor scale framework and the blues scale is illustrated (the blues scale appears usually in a major-minor context within jazz music). The same interval representation may also be used for the major scale and the pentatonic scale as the tones and intervals of the latter are a subset of the former.

Table 3.

Number of Semitones	1	2	3	4	5	6	7	8	9	10	11
	1st				3rd			5th			
Major Scale	B1	B2			A	C1		B1	B2		
(t t s t t t s)			B1	B2		C1	A			B1	B2
			2nd			4th				6th	
	1st				3rd			5th			
Asc. Mel. Minor Scale	B1	B2		C1	B1	B2		B1	B2		
(t s t t t t s)			B1	B2		B1	B2	C1		B1	B2
			2nd				4th			6th	
		1st			3rd			5th			
Harmonic Minor Scale	B1	B2	C1	C1	B1	B2		B1	B2		
(t s t t s tr s)			B1	B2		B1	B2	C1	C1	B1	B2
			2nd				4th			6th	
		1st					3rd				
Pentatonic Scale		B1	B2				A	C1			
(t t tr t tr)				C1	A				B1	B2	
				2nd					4th		
		1st				3rd				5th	
Blues scale	B1	B2	B3	C1	C2	B1	C3	C4	B1	B2	B3
(tr t s s tr t)		C1	C2	C3	B1		B1	C1	C2	C3	
			2nd					4th			
	1st			3rd		4th			6th		
Octatonic Scale	B1	B2		B1	B2	A			A		
(t s t s t s t s)			A				B1	B2		B1	B2
			2nd				5th			7th	
		1st				3rd				5th	
Whole-tone scale		A				A				A	
(t t t t t t)				A				A			
				2nd				4th			
	1st		3rd		5th	6th		8th		10th	
12-tone Scale	A		A		A	A		A		A	
		A		A			A		A		A
		2nd		4th			7th		9th		11th

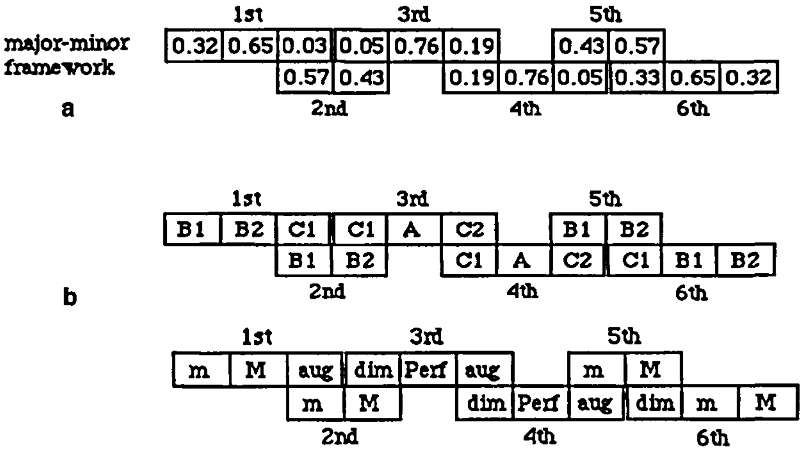


Fig. 5.

In the GPIR every pitch interval may be accurately represented by an array of the sort [dir, nci, mdl, pci, oct] where *dir* (direction) takes values from {–, =, +} depending on the direction of the interval, *nci* (name-class) takes values from {0, 1, 2, ... M} for an M-tone scale, *mdl* (modality)[7] takes values from classes {A, B, C or D}, *pci* (pitch-class interval) takes values from {0, 1, 2, ...N} for an N-tone discrete equal-tempered pitch space and *oct* is the number of octaves within compound intervals. For instance, in the traditional diatonic system an ascending augmented 2nd is [+, 1, C1, 3, 0], a descending minor 3rd is [–, 2, B1, 3, 0] and an ascending major 9th is [+, 1, B2, 2, 1] whereas the same intervals in the 12-tone system are [+, 3, A, 3, 0], [–, 3, A, 3, 0] and [+, 2, A, 2, 1]. In the latter case the nci and mdl entries become redundant.

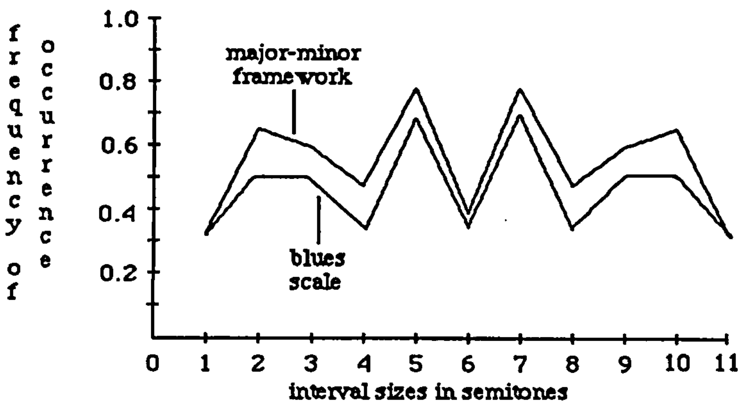


Fig. 6.

APPLICATION AND USES OF THE GPIR

The GPIR has been implemented in a PROLOG programming environment within a larger scale analytic-compositional project. The user presents to the system the interval array of a selected scale (or weighted set of scale interval arrays) and the system induces and stores the appropriate GPIR information (e.g., number of scale tones, number of discrete pitch elements, modality interval names, possible enharmonic spellings of notes and so forth). A set of operations has been developed that can be performed on the GPIR primitives in order to compute the interval between two pitches, the inverse of a given interval, the transposition of a pitch by a given interval and so on.

This representation increases the complexity of categorisation of intervals at the lowest level, but as it embodies structural properties that are inherent to the given scaling system, it facilitates reasoning and manipulation of the pitch material at higher levels of compositional and analytic processes. It has the advantage of encoding efficiently pitches and pitch intervals from a hierarchical tonal system down to a distributional 12-tone system.

Probably the most interesting aspects of this representation is the possibility to represent on computers other scaling systems in a way which is most relevant to them — e.g., pentatonic, octatonic, 9-tonic scales or even uncommon 7-tones genres (e.g., s s t t t t t). It may be the case that the lack of musical systems residing in the territory in-between the traditional highly hierarchical tonal system and the distributional atonal system is related to inefficiencies inherent in the traditional notation system. How can a composer notate, for instance, a functional 8-tone tonal piece on the traditional 7-tone stave notation? She/he either has to spend endless hours distinguishing the scale tones from the secondary nonscale tones (for instance, see Gillies 1993 on pitch notation and tonality in Bartók's music) or invent and learn a new notation system! The GPIR may enable computer-assisted compositional systems to compose music in hierarchical/functional systems other than the 7-tone diatonic system.

The GPIR could also be used creatively in analytic/compositional programmes by forcing an analysis (or composition) based on "wrong" scaling-interval representations (e.g., analyse 7-tone music with a 9-tone interval representation, etc.). One may impose the structural and functional properties of a given piece to different scale representations. This kind of experimentation could lead to interesting novel compositions.

This representation may easily be adapted or extended to meet the needs of musical systems (ethnic musics, experimental scaling environments etc.) other than the Western 12-tone equal-tempered system.

It is suggested that a flexible pitch-interval representation, such as the GPIR, may prove itself indispensable when devising a computer system that attempts to deal with a wide variety of musical styles. In the next paragraphs, two applications are presented that highlight the representational advantages of the GPIR in devising a) a transcription program; and b) a pattern-matching system.

Transcription of melodies based on GPIR

As stated earlier, the transcription of a piece of music from an M-tone system to an N-tone, where the M-tone system is not a subset of the N-tone, is not a function and thus, is not a straightforward process. We have implemented a system that converts melodies from a 12-tone notation (MIDI) to the traditional 7-tone notation based on the GPIR theory (an important similar system implemented from a cognitive perspective appears in Longuet-Higgins 1976/1987). The principle of classifying intervals according to their frequency of occurrence is strongly supported by this application. The transcription system applies two basic principles:

1) *Notational Parsimony* (i.e., "spell" notes making minimum use of accidentals[8]).
2) *Interval Modality Optimisation* (i.e., prefer intervals in the order of their frequency of occurrence — most preferable: class A — least preferable: class D).

A numerical grading of the different parameters that relate to these principles was devised:

Interval Notational Parsimony:	
nonenharmonic spelling of notes	0
enharmonic spelling of one note	2
both notes enharmonic	6
Interval Modality Optimisation:	
intervals of class A or B[9]	0
intervals of class C	1
intervals of class D	4

For any given sequence of MIDI pitch numbers all the alternative spellings of each pitch are found (see Fig. 7 for the beginning of the theme of Bach's *Musical Offering*). Then, the program calculates the total sum of the above values for each possible string of traditional pitch names and selects the ones with the minimum sum value.

As the system may find more than one string with the minimum value, we have added one additional rule:

60	63	67	68	59	67	66	65	64	63	62	61	60 ...
D♭♭		A♭♭		C♭	A♭♭		G♭♭	F♭		E♭♭		D♭♭
	E♭		A♭			G♭			E♭		D♭	
C		G		B	G		F	E		D		C
	D#		G#			F#			D#		C#	
B#		Fx		Ax	Fx		E#	Dx		Cx		B#

Fig. 7.

3) Prefer a sequence in which the *higher "quality" intervals appear last.* This rule accounts for the asymmetric temporally ordered aspects of musical perception (Deutsch 1984; Krumhansl 1990) according to which listeners, for example, tend to hear the last note of an interval as more prominent. When there are two alternative spellings of two intervals the system prefers the sequence in which the last interval belongs to a "better" modality class. This rule gives precedence e.g., to the sequence G — G♯ — A over the equivalent G — A♭ — A (they both have a total value of 4).

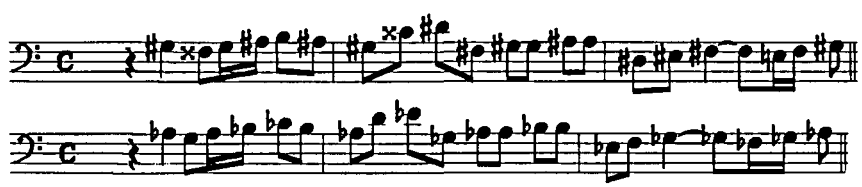
The system was tested over a set of diatonic melodies with unexpectedly good results for such a small and general set of rules (note that there is no higher level representation of musical knowledge such as keys, tonalities, modulations, tonics etc.).

The transcription programme was applied on the 24 fugue themes from J.S. Bach's *Das Wohltemperierte Klavier I.* All themes were accurately notated with only a few exceptions.

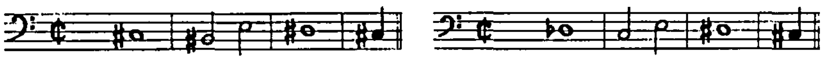
Fugue 14 in F♯ min (transcription) Identical with original.

Fugue 24 in B min (transcription). Identical with original. Note the use of enharmonic spelling of notes in bar 2 (E♯) and bar 3 (B♯).

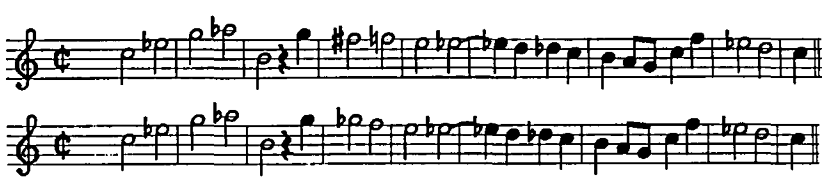
Fugue 18 in G♯ min (original and transcription). The system prefers the enharmonic key of A♭ minor. The same occurs in fugue 3 (C♯ maj) and Fugue 13 (F♯ maj).



Fugue 4 (original and transcription). This problem may be bypassed if additional rules are applied such as "avoid enharmonic spellings of a tone in a single passage", or if the optimisation method is additionally applied to intervals between noncontiguous notes, e.g., every other note.



The theme from *Musical Offering* by J.S. Bach (original and transcription). The selection of G♭ in the transcription is due to Rule 3. Both sequences have the same total value. Bach prefers F♯ for harmonic reasons.



The programme was applied to some melodies from later periods. For example: Opening from *Ballade* Op. 23 by F. Chopin (transcription). Identical with original.

The beginning from the English Horn solo from the third act of *Tristan Und Isolde* by R. Wagner (original and transcription). The incongruence in bar 2 is of the same nature as the one in fugue 4 (above).

AI methodology of the transcription programme

The total number (T) of all possible strings that can be derived from ν_1 pitches with 2 alternative spellings and ν_2 pitches with 3 alternative spellings is:

$$T = 2^{\nu 1} \cdot 3^{\nu 2}$$

This was significantly reduced by disallowing altogether a) two successive enharmonic notes and b) all class D intervals with the exception of chromatic semitones. T thus becomes approximately[10]:

$T = 2^{\nu}$ where $\nu = \nu_1 + \nu_2$ i.e., total number of notes.

The total number of possible paths given by this function is significantly reduced but still is an *exponential function* of ν leading, thus, to a combinatorial explosion and making it impossible to calculate the transcription sum values for larger sequences of pitches.

This problem was overcome by implementing an algorithm that transcribes the piece gradually by smaller sections. An overlapping technique was devised in such a way that only the middle part of the transcribed section is selected (marked by the bold segments of the lines in Fig. 8). This gives stability to the system and avoids misinterpretations of the interval qualities near the edges of the sections[11].

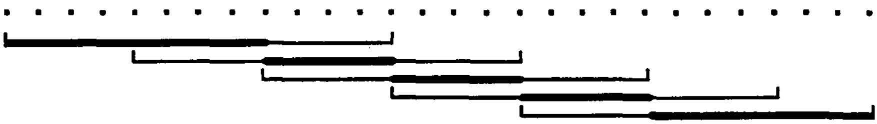


Fig. 8.

The above function now becomes:

$$T = c \cdot {}^{\nu}/_{\mu} \cdot 2^{\mu} = ({}^{c}/_{\mu} \cdot 2^{\mu}) \cdot \nu = k \cdot \nu$$

where μ = constant number of notes in transcription sections, ν = total number of notes and c = a constant that depends on overlapping. For the above example ν = 28, μ = 13 and c = 3 (each 5-element subsection is transcribed 3 times as beginning, middle and ending of the 13-element transcription sections).

This relation is a *linear function* and melodies of any length can be transcribed within reasonable computational times. Table 4 shows the values of the three functions for various values of ν.

How good are the transcription results obtained by this shifting overlapping technique compared to the results obtained by the method that transcribes a whole melody at once? Both methods were tested over a number of melodies generating

Table 4.

	n = 10	n = 20	n = 50	n = 100	n = 500
$T = 2^{n1} \cdot 3^{n2}$ (n1 = n2)	$8 \cdot 10^3$	$6 \cdot 10^7$	$3 \cdot 10^{19}$	$8 \cdot 10^{38}$	$3 \cdot 10^{194}$
$T = 2^n$	10^3	10^6	10^{15}	10^{30}	$3 \cdot 10^{150}$
$T = k \cdot n$ (k = 1890, c = 3, μ = 13)	$8 \cdot 10^3$ ($T = 2^n$, n = 13)	$4 \cdot 10^4$	$9 \cdot 10^4$	$2 \cdot 10^5$	$9 \cdot 10^5$

always identical results. The reason for this is that intervals of class C and D tend to appear isolated inbetween unambiguous stable sections of class A and B intervals.[12] The sections that may receive alternative spellings with a similar sum value are, in most cases, short — usually just a few notes. This localisation of the transcription process allows a shifting overlapping method to yield good results (although, in general, it is not necessarily true that the results obtained by the two techniques are always identical).

This technique of a step-by-step transcription by overlapping sections is also closer to the processes that take place while a listener is notating a little-by-little heard melody (melodic dictation). The listener hears and notates a few bars at a time making possible alterations to the immediately preceding notes if required by the new input.

This simple transcription system proves the importance of the hierarchical classification of intervals according to their frequency of occurrence within a scale and suggests that similar processes may be adopted by listeners when involved in notation tasks or by composers when notating their pieces.

Interval representation and pattern-matching

We will now briefly illustrate the importance of pitch-interval representation in the design of a pattern-matching process that detects repetition of pitch patterns. Our discussion will revolve around a matching process proposed by West, Howell & Cross (1992, p. 7) which they illustrate concisely in Fig. 9.

Although this process is very general and economic and will give successful results for the detection of repetitions in the majority of musical surfaces being presented to the system, we will argue that it has some inherent deficiencies relating to the way pitch-intervals are encoded. We will examine this system in two respects:

1. If the levels of representation of the pitch-intervals are considered to be strictly hierarchical, i.e., matchings that are detected first, starting from the lowest level

	a)	b)	c)	d)
chroma:	*, +2, +5, -2	*, +1, +5, -1	*, +2, +5, -2	*, +4, +4, -4
scale step:	*, +1, +3, -1	*, +1, +3, -1	*, +1, +3, -1	*, +2, +2, -2
contour:	*, +, +, -	*, +, +, -	*, +, +, -	*, +, +, -

Fig. 9. "A simple figure (a), requires at least three different methods of encoding pitch intervals for repetition to be detected by a matching process. Repetition with in-scale transposition (b) requires scale step encoding; repetition with simple transposition (c) requires chroma (pitch class) encoding; and repetition with contour preservation (d) requires contour encoding" (West et al. 1992, p. 7).

(chroma) upwards, are the ones to be selected (it is understood that this is suggested by the authors) then the system exhibits the following problems:

a. It disregards important differences[13] by matching (considering identical) enharmonic intervals in tonal surfaces. This shortcoming appears because the chroma level does not effectively represent a tonal surface. The system is not strictly hierarchical as it is possible to find situations, e.g., Fig. 10, where a higher (more abstract) level contradicts (does not match) a repetition detected at a lower level.

chroma:	*,	+3,	+1,	-7	*,	+3,	+1,	-7
scale step:	*,	+1,	+1,	-4	*,	+2,	+1,	-4
contour:	*,	+,	+,	-	*,	+,	+,	-

Fig. 10.

b. The scale step matching level is arbitrary in a distributional atonal environment (based on the 12-tone system). A quantification of the chroma level into equal numbers of semitones may be less arbitrary (e.g., two-semitone intervals, and so on).

c. Hierarchical tonal systems other than the 7-tone are not efficiently represented neither in the chroma level nor in the scale step level. The pitch and pitch-interval properties of such systems are not appropriately accounted for and thus the analyses obtained from this matching procedure are apt to diverge from the expected results.

2. If the levels of representation are considered to be complementary to each other (e.g., chroma and scale-step levels) then the problems discussed in 1a and 1b may be eliminated, as it is possible to infer implicitly the dissimilarity of enharmonic intervals in a 7-tone environment or to inactivate the scale-step level in a distributional 12-tone environment. This means that the system needs additional mechanisms that can induce these interlevel relations; but this way it loses on its simplicity and economic outlook. Even with the aid of an extra mechanism, problem 1c cannot be accounted for if the initial representations are not altered.

We suggest that the general pitch-interval representation proposed above may explicitly represent a wider range of pitch structures in a purely hierarchic fashion.[14] In Fig. 11 the first pitch pattern is matched[15] to each of the following patterns within: a) a 7-tone diatonic representation and b) a 12-tone representation.

This pattern-matching procedure gives rise to different analyses of a musical surface for different scaling systems. It is also possible to make use of more than one analyses in a multiple-viewpoint approach implementation.

	a)	b)	c)	d)	e)
For 7-tone diatonic representation:					
[dir,nci,mdf]	*,+1C1,+1B1,-4A	*,+1C1,+1B1,-4A	*, +2B1,+1B1,-4A	*,+1B2,+1B1,-4A	*,+2B2,+2B1,-2B2
[dir,nci]:	*, +1, +1, -4	*, +1, +1, -4	*, +2, +1, -4	*, +1, +1, -4	*, +2, +2, -2
[dir,*nci'*]:	*, +1, +1, -4	*, +1, +1, -4	*, +2, +1, -4	*, +1, +1, -4	*, +2, +2, -2
intermediate levels: . . .	. . .	. . .	. . .	. . .	. . .
[dir]:	*, +, +, -	*, +, +, -	*, +, +, -	*, +, +, -	*, +, +, -
For 12-tone representation ([dir,nc,mdf] level is redundant as mdf is always A):					
([dir,nci,mdf]	*,+3A, +1A,-7A	*, +3A, +1A, -7A	*, +3A, +1A, -7A	*, +2A, +1A, -7A	*, +4A, +3A, -4A)
[dir,nci]:	*, +3, +1, -7	*, +3, +1, -7	*, +3, +1, -7	*, +2, +1, -7	*, +4, +3, -4
[dir,*nci'*]:	*, +3, +1, -7	*, +3, +1, -7	*, +3, +1, -7	*, +2, +1, -7	*, +2, +2, -4
intermediate levels: . . .	. . .	. . .	. . .	. . .	. . .
[dir]:	*, +, +, -	*, +, +, -	*, +, +, -	*, +, +, -	*, +, +, -

Fig. 11.

CONCLUSION

It has been shown that the proposed general pitch-interval representation introduces a better way of encoding pitch intervals depending on the specific scale qualities of musical structures. It is maintained that the hierarchy of scale tones over a discrete pitch space makes possible — and even necessary — the more elaborate classification of pitches and pitch intervals according to their higher level structural properties. Special emphasis was given to the categorisation of intervals relating to their frequency of occurrence within a scale and it has been suggested that this classification method may be directly related to the traditional pitch interval naming system and to the way diatonic surfaces are notated. The GPIR enables a more accurate representation of interval properties according to the scaling framework in which they appear. Although this representation is more complicated at a low level it facilitates further reasoning about higher level properties of a musical structure. The GPIR can be applied not only to the 12-tone system but to any N-tone equal-temperament domain. The flexibility of this representation renders it an ideal candidate for (computer) systems that attempt to manipulate musical structures from diverse musical domains with a varying degree of hierarchic organisation.

ACKNOWLEDGEMENTS

Thanks are due to Peter Nelson, Alan Smaill and Raymond Monelle for useful comments and suggestions. This research was funded in part by a University of Edinburgh Postgraduate Research Award.

APPENDIX

In the Aristoxenian pitch system (Aristoxenos 1989; Xenakis 1992) the smallest pitch-interval unit is the twelfth-tone. The *tone* is defined as the difference between the perfect fifth (dia pente) and the perfect fourth (dia tessaron) and can be divided into two parts called *semitones* (6 twelfths), three parts called *chromatic dieseis* (4 twelfths) or four parts called *enharmonic dieseis* (3 twelfths). Three of these are combined to form *tetrachords* (total of 30 twelfths i.e., 2½ tones). There are three genres of tetrachords: a) *enharmonic* (3 + 3 + 24 = 30 segments); b) *chromatic* (soft: 4 + 4 + 22 = 30, hemiolon: 4.5 + 4.5 + 21 = 30 and toniaion: 6 + 6 + 18 = 30) and c) *diatonic* (soft: 6 + 9 + 15 = 30 and syntonon: 6 + 12 + 12). (If it is required that all intervals e.g., the ones in the chromatic hemiolon are expressed in integer numbers then the tone should be divided in 24 segments). Tetrachords and tones are further combined to form *systems*.

As an example, let us create a system which consists of two syntonon diatonic tetrachords (6 + 12 + 12 = 30) disjunct by a tone. If octave equivalence is further assumed, this *system* is the diatonic genre. This genre can be represented by 7 nc integers {0, 1, ... 6} for the 7-tone scale, 72 pc integers {0, 1, ...71} for the 72-tone discrete pitch space and 25 mdf integers {–12, –11, ...–1, 0, 1, ... 11, 12} since the largest possible scale step interval is the tone (12 units). For instance, between the scale tones [2, 0, 24, 4] and [3, 0, 30, 4] there exist 5 discrete pitches with two possible enharmonic spellings each e.g., for one of these: [2, 2, 26, 4] and [3, –4, 26, 4].

The Aristoxenian scaling system may accommodate a wide gamut of microtonal systems because of its fine resolution of intervals.

NOTES

1. Enharmonic intervals were originally physically different until the equal-temperament tuning forced them into identity, and even today, enharmonic intervals, when performed on nontempered instruments (e.g., voice, violin etc.), appear in different physical sizes (different intonation) depending on musical context (Schackford 1961, 1962).
2. Alternatively, integers may correspond to the symbols assigned to the elements of the discrete pitch space (columns in Table 5 consist of the same letter-symbols) facilitating thus pitch representations, especially in cases where within the same piece of music we have changes of scaling systems, as pitch names remain invariant within the overall pitch structure. Of course, in this representation, the modulo M (for M-tone scales) mathematical formalisms no longer apply.

Table 5.

7-tone diatonic scale	0	♯♭	2	♯♭	4	5	♯♭	7	♯♭	9	♯♭	11	(0)
pentatonic scale	0	♯♭	2	♯♭	4	♯♭♭	♯♭	7	♯♭	9	♯♭♭	♯♭	(0)
octatonic scale	0	♯♭	2	3	♯♭	5	6	♯♭	8	9	♯♭	11	(0)
12-tone scale	0	1	2	3	4	5	6	7	8	9	10	11	(0)

3. Every genre of scales will have exactly the same set of intervals and frequency of their occurrences i.e., it does not matter which tone is considered to be the tonic in a particular mode.
4. This view seems to relate to Krumhansl's observation (Krumhansl 1990, p. 273) that there is a link between the degree of consonance/dissonance of an interval to its frequency of occurrence, although we avoid to make any direct connection of modalities of intervals to degrees of consonance.
5. It may be preferable to analyse atonal music with an N-tone ($N < 12$) scale system as an atonal composition may microstructurally be based on N-tone scale fragments.

6. This weighting is not a result of cognitive, statistical or other studies; its aim is simply to represent all the different kinds of the major-minor scales in a balanced manner. It attempts to give half weight to the major scale and half to the minor scale (the natural minor scale actually reinforces both sides as it consists of intervals identical to those of the major scale — they both belong to the same genre of diatonic scales).
7. The modality symbol may be broken down into a two element list containing a modality symbol {a, b, c, d} or {1, 2, 3, 4} and an index number that is assigned to different members of the same modality class; the index number may indicate the number of units that an interval is greater or lesser than a reference size in that modality.
8. This actually means to avoid the enharmonic spelling of notes e.g., prefer C and avoid B♯ and D♭♭.
9. It is not possible to have for one name-class interval both modalities of class A and B, as this would give an overall frequency of occurrence greater than 100%.
10. For example, two notes with 3 alternative spellings may give $3^2 = 9$ combinations. Four of these are disallowed by the use of *constraint a* and usually one more by constraint *b* reducing thus the initial number of combinations to approximately $4 = 2^2$ (e.g., for the interval between MIDI notes 59–67 the spellings C♭-A♭♭, A♯-F♯, C♭-F♯, A♯-A♭♭ are disallowed by constraint *a* and B-A♭♭ by constraint *b*).
11. An instance of boundary problems caused by a nonoverlapping transcription technique can be demonstrated in Bach's fugue in B min. If the transcription section boundary is on 6th note of bar 2 then this note will be spelled E♯ as the last note of the preceding section and F as the first note of the following section!
12. This relates to the fact that "if X Y Z are three successive notes of a melody which, on paper, are separated by chromatic intervals XY and YZ, then there is always an alternative, simple interpretation of the middle note Y which transforms both intervals into diatonic ones". (Longuet-Higgins 1987, p.113)
13. For example, the minor 3rd and the "rare" augmented 2nd intervals are classified together as 3 semitone intervals. This way the important distinction between them is disregarded altogether. The opposite situation occurs when 12-tone music is analysed by a 7-tone scale-interval representation i.e., nonsignificant information is encoded as significant.
14. If hybrid musical systems are taken into consideration e.g., 12-tone music with 7-tone microstructural properties, then additional evaluation-selection mechanisms should be employed to combine different matching procedures.
15. Name-class intervals *(nci')* are matched if they are identical or differ by one unit.

REFERENCES

Aristoxenos, *Elementa Harmonica.* Translated in Barker, A. (1989). *Greek Musical Writings II — Harmonic and Acoustic Theory.* Cambridge (MA): Cambridge University Press.

Bharucha, J.J. (1984a). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology,* 16, 485–518.

Bharucha, J.J. (1984b). Event hierarchies, tonal hierarchies, and assimilation: A reply to Deutsch and Dowling.*Journal of Experimental Psychology: General,* 113, 421–425.

Brinkman, A.R. (1990). *Pascal Programming for Music Research.* Chicago: The University of Chicago Press.

Browne, R. (1981). Tonal implication of the diatonic set. *In Theory Only,* 5(6,7), 3–21.

Clough, J. (1979). Aspects of diatonic sets. *Journal of Music Theory,* 23(1), 45–60.

Clough, J. (1980). Diatonic interval sets and transformational structures. *Perspectives of New Music,* 18(2), 461–482.

Clough, J. & Myerson, G. (1985). Variety and multiplicity in diatonic systems. *Journal of Music Theory,* 29(2), 249–270.

Deutsch, D. (1982b). The processing of pitch combinations. In D. Deutsch (ed.), *The Psychology of Music.* New York: Academic Press.

Deutsch, D. (1984). Two issues concerning tonal hierarchies: Comment on Castellano, Bharucha and Krumhansl. *Journal of Experimental Psychology: General*, 113, 413–416.
Dowling, W.J. & Harwood D.L. (1986). *Music Cognition.* Orlando: Academic Press.
Forte, A. (1973). *The structure of atonal music.* New Haven: Yale University Press.
Gillies, M. (1993). Pitch notations and tonality: Bartók. In J. Dunsby (ed.), *Early Twentieth-Century Music.* Oxford: Basil Blackwell Ltd.
Krumhansl, C.L. (1990). *Cognitive Foundations of Musical Pitch.* Oxford: Oxford University Press.
Longuet-Higgins, H.C. (1976/1987). The perception of melodies. *Mental Processes — Studies in Cognitive Science.* Cambridge (MA): MIT Press.
McAdams, S. (1989). Psychological constraints on form-bearing dimensions in music. *Contemporary Music Review*, 4, 181—198.
Rahn, J. (1980). *Basic Atonal Theory.* New York: Schirmer Books.
Reneger, E. (1964). On Allen Forte's theory of chords. *Perspectives of New Music*, 8(1), 191—212.
Schackford, C. (1961). Some aspects of perception, Part I: Sizes of harmonic intervals in performance. *Journal of Music Theory*, 5, 162–202.
Schackford, C. (1962). Some aspects of perception, Part II: Interval sizes and tonal dynamics in performance. *Journal of Music Theory*, 6, 66–90.
Sloboda, J.A. (1985). *The Musical Mind: The Cognitive Psychology of Music.* Oxford: Oxford University Press.
Taneiev, S.I. (1902/1962). *Convertible Counterpoint in the Strict Style.* Boston: Bruce Humphries Publishers.
West, R., Howell, P. & Cross, I. (1991). Musical structure and knowledge representation. In R. Howell et al. (eds.), *Representing Musical Structure.* London: Academic Press.
Xenakis, I. (1992). *Formalised Music.* New York: Pendragon Press.

Emilios Cambouropoulos
Faculty of Music
University of Edinburgh
12 Nicolson Square
Edinburgh EH8 9DF
Tel. +44-131—6502427
E-mail: emilios@music.ed.ac.uk

Emilios Cambouropoulos has studied Physics (University of Athens, Greece) and Music (Conservatory of Attiki, Athens), and obtained an M.Sc. in Music Technology (University of York, U.K.). He is currently a research student at the University of Edinburgh reading for a Ph.D. degree in Music and Artificial Intelligence (title of thesis: "Towards a General Computational Theory of Musical Structure").