



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

THE FORMAL DESCRIPTION OF
MUSICAL PERCEPTION

MARK STEEDMAN



CONTENTS.

Part I : Introduction.	1
Part II ; Metre.	29
Part III: Key and Notation.	83
Appendices.	226

Abstract

This work concerns a problem in modelling people's understanding of music. The problem is cast in the terms of discovering formal rules for transcribing melodies into musical notation, as this might be done by a student in a harmony class, taking musical dictation from a 'deadpan' performance on the keyboard. The score that results reflects important aspects of the structure and interpretation of the piece, which are only implicit in the performance. In Part I it is argued that this paradigm raises questions of general relevance to the study of our perception of music.

In carrying out the task of notating a piece, two kinds of problem arise: what are the harmonic relations between the notes, and what are the metric units into which they are grouped? These two problems are considered in isolation from one another. In Part II, algorithms which embody two kinds of rule for the inference of metre are presented.

In Part III the harmonic problem is considered. It arises from the fact that the number of keyboard semitones between two notes does not, by itself, identify their harmonic relation, which is what the notation has to express. Among other considerations, the key of the piece is an important characterisation of this identification - but the key is not explicit in the performance, and must itself be inferred. An earlier theory of harmonic relations is further developed into algorithms for assigning key-signatures and notation to melodies.

By the definition of the problem, we are committed to the concern of music belonging to the tradition of Western tonal music, to which the/

the idea of key applies. Most of the musical examples discussed will be taken from the work of one of its outstanding exponents, J.S. Bach, and in particular we shall be dealing with the fugue subjects of his "Well-Tempered Clavier".

Some of the contents of Part II, Section § 2, and of Part III, Section § 3, have already appeared in a paper published in collaboration with H.C. Longuet-Higgins. Part III, Section § 2.1, describes his prior work in formulating the theory of harmonic relations, mentioned above, which forms the foundation of the work described in that section, and has been published elsewhere by him.

The work was carried out under the supervision and guidance of Professor H.C. Longuet-Higgins, whom the author thanks for his teaching.

Terminology

It is intended that the discussion to follow should be understandable by anyone, whether or not they have had much formal musical teaching. To that end, the use of abstruse musicological terms has been avoided where possible. Where they have been used, they are generally explained in the accompanying text, even at the risk of some tedium for the reader who is more familiar with them. However, certain terms are so basic as to be necessarily assumed in the text, and have been used there without explanation. Most people will be familiar with such terms as chord, scale, downbeat, tonic and dominant, and so on, but nevertheless these have been briefly defined or described in Appendix V.

All the programs were written in the language POP2 (Burstall et. al., (1971)).

The fugue subjects of the Well-Tempered Clavier were taken from the edition edited by Tovey and Samuel (1924). In transcribing them for the computer, each note was represented by a triple of integers, identifying the piano-key, or equally-tempered value of its pitch, its octave, and its duration as written, expressed as a multiple of the duration of demisemi-quaver, or thirty-second note.¹ Where the notes or rests of the score were written as tied, they were transcribed as single notes, or rests, in keeping with the idea of giving the programs only the information explicitly available to the ears of the student in the 'musical dictation' paradigm. However, trills, where the editors had written them as single notes, were input as such, rather than as the sequence of notes which would be used to play them. In one case, that of the very long initial trill of Fugue 13 of Book II, /

v.
13 of Book II, the notation of this edition is slightly confusing in this respect, (though not of course to the human performer, who knows the conventions of scores). A long note is marked as a trill, but is only in fact the beginning of the trill, whose ending is explicitly written out. This trill was therefore transcribed in full, beginning with an alternation of tonic and leading note.

Five windows light the cavern'd man: thro' one
 he breathes the air;
Thro' one hears music of the spheres; thro' one
 the Eternal Vine
Flourishes, that he may recieve the grapes; thro' one
 can look
And see small portions of the Eternal World
 that ever groweth;
Thro' one himself pass out what time he please,

Part I. Introduction.

1. Why Study Musical Perception ?	2
2. How To Study Musical Perception.	6
3. Two Problems.	25

PART I

Introduction

1. Why Study Musical Perception?

The musical faculty seems oddly useless, viewed in the context of the usual concerns of perceptual or cognitive psychology. Unlike such capabilities as those involved in vision and use of language, it can hardly be imagined to have come about as something of direct evolutionary value. Yet it has a power and attraction quite disproportionate to this non-utility. In these respects it is more like such arts as painting and poetry, which seem to be the 'by-products', or secondary developments, of the more directly useful corresponding faculties of vision and language, than like the independent perceptual mode that the title of this work might seem to imply.

Thorne (1970) characterized this secondary nature of art, as the offspring of more practical faculties, when he identified the relation of poetry to its parent, natural language, as one of creation of "a new language or dialect", to the end of saying things that "cannot be said in Standard English at all, - though they can be understood only by someone who understands Standard English". Gombrich (1960) offers a parallel definition of the relation of painting to vision.

Music is like this, inasfar as part of the activity of composition lies in the creation of a new 'language' or 'dialect', by modifying or discarding established or familiar rules. Given that the old

rules are understood by the hearer, he may learn the new variant.

It would be most unusual for a study of one of the faculties that have been described as 'basic', or 'primary', to be attempted by looking first at its 'secondary' artistic developments. You do not arrive at a theory of English by looking only at its artistic use by James Joyce, (though once even the rudiments of such a theory have been arrived at, such a study might direct attention to interesting problems.) Rather the reverse: the form of the current theory of the basic language will direct the study of the art, as Thorne(1970) implies when he says that 'there are special reasons for thinking that the introduction into linguistics of the notion of Generative Grammar will have an effect upon stylistics.' Meyer (1956), in the field of music, and Gombrich (1972), have explained aspects of their fields of study in terms of the psychological concept of expectancy. While one result has been the raising, in turn, of further interesting questions for psychologists, nevertheless Gombrich (1963) has suggested one reason why the reverse procedure - the direction of psychology by aesthetics - is unlikely, when he says that every image presented as art alters the frame of reference for itself and other art-works: the new dialect becomes a part of the established language. He compares the "strange precinct we call art" with a hall of mirrors, or a whispering-gallery: we might draw the inference that the psychologist who insists on finding the real figure, or the direction of the voice in that precinct, may get lost among "a thousand memories and after-images".

But music is unusual among the arts. If poetry is a process of 'playing about' with the rules of language and painting is play with the processes of seeing things, what, then, is the practical, down-to-earth, 'useful' faculty whose 'language' we play about with in music, the first shout in that particular whispering-gallery? There does not seem to be one, and in this respect we do not fully explain the strange nature of music by classifying it with other art.

Admittedly, music appears in some respects to involve a secondary development from language, but it does not stand in the same kind of directly secondary relationship as the other arts which constitute developments of language, such as poetry. It seems rather that, instead of being secondary to any one 'useful' faculty, music constitutes an abstraction of many. Perception of pitch and time are the by-products of perceptual apparatus evolved to quite separate practical ends. The structure of melody has often been associated with the idea of language and sentence structure, while the vocabulary of musical form, at a still higher level, abounds with the metaphors of discourse, as in the terms 'exposition' and 'answer'. But in none of these cases does music involve the exercise of the faculty itself, but rather of the pure capabilities that it involves, in isolation from their everyday subject matter.

Herein may lie the great attraction of the psychology of musical perception, in the promise of direct access to the very

bases of our perceptual processes.

2. How To Study Musical Perception.

When, for some area of study, there are several theories and few unexplained facts, then the way to pursue the study may be to find some new facts. There is a procedure that dictates which new facts should be sought: they should be those that discriminate most sharply between the various theories, in being predicted by some and contradicted by others. There is also a procedure that defines a relevant fact, and says how it is to be established^d, in the repeatable experiment. On the other hand, there may be large numbers of unexplained facts in the area of concern, and few theories, if any, to account for them, none of them differing as to their predictions. In this situation the experimental procedure is inappropriate. For one thing, it merely adds to what may be a very confusing plethora of data. More significantly, the procedure mentioned above for deciding which fact to look for does not apply. In this situation the appropriate procedure is rather to develop theories to account for the facts already to hand, and leave the acquisition of further data until it becomes necessary.

A suitable case study of the latter situation is offered in linguistics. It was not necessary to carry out an experiment to establish the fact that, of the now famous pair of sentences

John admires sincerity.

and

*Sincerity admires John.

only one is good English. In some sense, this fact was just lying around, not accounted for. It is by no means straightforward to

gleña) such facts, and it may require a considerable body of prior theory. (The fact that an active sentence and its passive are paraphrases can only be noticed and only makes sense in the context of a prior descriptive theory which says that the sentence is a meaningful unit which can sensibly be considered in isolation from extended discourse.) Nevertheless, the question of an experiment does not arise, and indeed it is hard to say how an experiment in which a hundred subjects were interrogated on this question could mean anything at all.

Music is another field where the theories, if there can be said to be any, are weak, and the data is just lying about. Experiments to establish facts, such as that people can, for a wide range of familiar and unfamiliar music, detect the occurrence of 'wrong notes', are out of place. It may be difficult to say exactly what we mean by such a statement, but that just makes it all the more difficult to even formulate a meaningful experiment. Some descriptive theory is necessary first, and the development of theory sufficiently to define just this kind of behaviour will occupy quite a few of the pages that follow. It is for these reasons that the extensive and venerable psychophysical literature concerned with the quantitative measurement of musical performance is not discussed here, with the exception of a brief consideration of Helmholtz's (1862) experimental account of consonance in a later section.

It may not be possible simply to distinguish theories in such fields as language and music on an experimental basis, but strong

requirements of formality and computability can be made of the theories. This is the case in this study. Accordingly, because of the vastly inferior position of musicology with respect to descriptive vocabulary at the simpler levels, as compared with the corresponding detailed vocabulary of constituent structure upon which formal linguistics was built, the concern will be with rather restricted problems. We shall accordingly have to leave on one side more ambitious theoretical works, which, whilst waiving the business of strict formality, are free to gain valuable insights into the highest reaches of musical comprehension. Despite this neglect, it is hoped that, in particular among works of musical analysis, this may claim some kinship, however remote, with Meyer's (1956) incorporation of the Gestalt psychologist's ideas of Expectancy into the study of music, and with Cooke's (1959) idea of the nature of the 'language' of music.

The analogy between music and language which is implied in the titled of these last two, and in the previous discussion, is common throughout the literature, particularly that part of it which derives from Computer Science and Artificial Intelligence laboratories. The analogy is an obvious one: music consists of sequences in time, of a highly 'structured' kind, with various levels of this structure identified in the traditional descriptive vocabulary - sonata form, theme or subject, phrase or cadence, arpeggio, scale, trill, and so on - and has a similar function of communication. Meyer and Cooke between them have lent considerable

more force and interest to the metaphor, by identifying ideas of paraphrase and ambiguity in music corresponding to those in language. Can we continue this process, and, by identifying further correspondences between music and language, subsume some of the problems of musical perception under existing theories of language? The particular temptation, of course, is to take over the powerful formal apparatus of generative grammar which Chomsky (1957, 1965) applied to English. This kind of linguistics concerns itself in particular with the representation of language in writing, so it is natural enough that several pieces of work, which implicitly or explicitly subscribe to this idea of investigating music in a way that language has been investigated, concentrate upon the written form of music, the traditional notation of the score.

Simon and Sumner (1968) proposed that programmes should be designed to read and write scores to and from a representation of music in terms of simple generative procedures. They gave examples of various aspects of harmonic progression and rhythmic structure that could be expressed in terms of finite-state and phrase-structure rules. Several projects can be discussed under these two headings of score-generators and readers.

Lindblom and Sundberg (1970) produced an outline of a score-generator which involved a rather close parallel to the methodology of linguistic generative grammar, in the sense that they outlined a generative procedure which was to stand or fall as a theory of

the piece concerned on the 'acceptability' of its products as music of that type. They attempted to produce a generative specification of a certain class of 19th Century Swedish nursery songs. These harmonised songs were characterised by a very stereotyped range of metrical frameworks, and an absence of syncopation, or 'non-congruence' between the rhythm and this metre. The most common metre was of an 8-bar verse, consisting of two phrases themselves divided into two groups of two bars each. Other patterns were of a similar nature, readily described by simple generative rules. Using rules similar to those used with surface-structural descriptions of English sentences by Chomsky and Halle (1968), they assigned 'stress contours' to such metrical structures as the above. This property of the metrical descriptions appeared to bear a close relation to certain harmonic and rhythmical aspects of the data, and formed the core of a set of rules for generating similar tunes.

The authors have hand-simulated a generative procedure of this kind, restricting themselves to the single eight-bar symmetrical form, (according to an algorithm which is not completely spelled out) and offer examples of its results. The main interest of this work is in the way various traditional aspects of music, such as metre, rhythm, harmony and melody, can be related in the framework of a generative algorithm. Actual generation of pieces by such algorithms is a very weak criterion for judging them: despite the well-known definition of a generative grammar as one which generates all and only the legal

sentences of a language, no linguist in his right mind sets a computer generating at random according to his grammar, and examines the results for their Englishness or otherwise. He has much more sophisticated techniques which lead him very directly to the crucial generations and derivations. Traditional linguistics, just as perhaps it suggested to him in the first place that the theory must express (say) the relation of the active and passive, will also specify the ways in which sentences of English are the same or different in their behaviour under such a relation. Whole infinite classes of sentences whose derivation would add nothing to the discussion are eliminated in this way, and the range of test sentences may be narrowed down to a handful. In the case of music, on the other hand, and such generative schemes as this, it is not clear where we find these strong criteria. It is not even clear that, even in the case of such a restricted class of melodies as this, we have a clear criterion of 'well-formedness' or of the domain of the rules. Even if we choose to find Lindblom and Sundberg's pieces "well-formed" nursery tunes, it is not clear how we are to be convinced that the grammar generates 'all and only' tunes of this type, in the absence of a clear definition of this set.

It seems as if the methodology of generative grammar, with its device of the 'informant' who tests the conclusions of the grammar against his knowledge of the language, is more appropriate to a study like linguistics which has a substantial descriptive theory already established, but no single formal theory, than to

music, whose descriptive theory is more fragmentary.

In the absence of such a helpful prior descriptive theory, we have to develop one, and can only do so by looking at tasks which human beings perform with music that are considerably better defined than the generation of whole pieces.

Smoliar (1971) extended the idea of a score-generator considerably, expressing a wide variety of musical forms in the medium of computational procedures. His main concern was with the computer as an aid to the composer, and to capture ideas of musical form in the traditional sense, so that, like the related studies of Bamberger (1972) in the field of teaching music, this work is of secondary concern to a theory of human interpretation of music, at least until we have specified the more trivial levels of musical utterance, which in these schemes are left to the human operator.

The other approach advocated by Simon and Sumner (1968) was the modelling of the opposite aspect of musical activity involving scores. This is the interpretation of the information in the score, as a fuller description of the piece, as a human performer does when he sight-reads a piece and recognizes its structure in such terms as scale and broken-chord, as consisting of repeated or varied occurrences of a motive, or on a larger scale, recognizes the recurrence of a theme, perhaps in some transformation, such as stretto or answer of a fugue-subject.

Forte (1967) had already written and programmed some early stages of a project to carry out this analytic reading of musical

scores. The scores concerned were those of atonal music, input in a form as closely related to orthodox scoring as the usual character-sets and input-devices of a computer would allow. The scores, which were for many instrumental parts, were 'parsed' into a description in terms of simple segments or melodic fragments of one instrumental part, and in terms of the temporal relations of such fragments with each other, either co-extensive, overlapping or in sequence, and also in terms of simultaneous groupings of notes. These rather abstract categories of description of the piece correspond in some sense to the categories associated with orthodox tonal music, of melody, counterpoint and harmony, but it would be unfair to interpret them as an attempt to model or explain these traditional facets of musical comprehension. Forte's intention was to develop new analogues of these traditional analytic categories for application to the new music, to which the old ones do not apply. Nevertheless, music, however new, is to be heard by people, and the new theories must be measured against what they hear, so it is fair to point out that the criterion for judging the analyses of such programmes as this is exceedingly weak, as the criterion for judging the well-formedness or otherwise of the generative programs was said to be. As an illustration of this, and of the extreme differences between tonal and atonal music, we can take Forte's own example of one of the analyses produced by a further program, carrying out higher analysis on the output of the 'parser' already described.

Forte (1967) uses as his example of the programme's operation

a piece for string quartet by Anton Webern, Op. 9/5. One of the features of a piece which the parser's description captures is the simultaneous occurrence of notes in the various parts - what in tonal music, at least, would be called a chord. Forte's further analysis program is concerned with detecting the repetition of such a coincidence of notes in the same interval relationship as the first, although possibly with different absolute pitch-values.

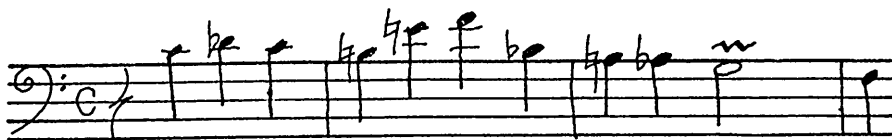
(This is perfectly familiar from tonal music, as when a chord G B D is recognized as the same configuration as a chord C E G at a transposition of a fifth.) The way in which his programs represent the music is such that the two groups of three simultaneous notes ringed in the fragment below are said to be related to each other by "a transposition ... at the interval of seven half-steps" or semitones, that is, "a fifth down."

It is not for someone who really does not understand this music to comment upon the analyses of such an eminent authority as Forte, so

to point out that if this were a passage in tonal music this would not be a correct 'reading' of this score is merely to point out the difference between tonal and atonal music. Although on a piano keyboard these groups of notes are related by a transposition of a fifth, since the keyboard equates, (in this case), $G\flat$ with $F\sharp$, it is only by making this equation that we can make $G\flat$ C B ~~from~~ into B F E by a transposition of a fifth. (It is a consequence of the fact that the program 'reads' notes as their keyboard equivalence classes that makes it produce this analysis.) We shall see in great detail later what such distinctions as that we have stated between $G\flat$ and $F\sharp$ actually signify, and say that the drawing of them is absolutely crucial to the business of comprehending tonal music, that both in reading scores and listening to music, far from ignoring such distinctions, we draw some which are not even apparent in the notation. For the moment, however, we shall just note that there is a distinction and that one of the few really solid pieces of information we have about the underlying structure of a piece lies in its score, where at least the composer has indicated his interpretation in his choice of key and time signatures, and notation. If Webern had meant that transposition, surely he would have written it, and ^d used an $F\sharp$, a notation he has no objection to using elsewhere in the very same piece. On the other hand, if the analysis does describe what we hear, and Webern has written it down wrongly, or for some other reason, then we lose the only really solid criterion we have on which to judge the

theory, and the main point, for our purposes of building psychological theories, of concerning ourselves with scores in the first place.

This necessity of, on the one hand, using the structural information in the score for all it is worth, and on the other, of checking theories against it, is a very good reason for confining our attentions to tonal music. Whatever we may think about the Webern, there is no doubt at all about the notation of the first two bars of the first F minor fugue subject of Bach's 'Well-tempered Keyboard':



The fourth note, for example, is clearly the leading-note (major seventh) of the dominant, for such reasons as that the following note is the leading-note, (followed by the tonic, as if there were any doubt) and the interval that the fourth note makes with it must be a perfect fifth, or that (for stylistic reasons) it must be related to its predecessor by a diatonic semitone. These are among the facts that a musician would read from the score, and the kind of statement that we might require from a 'score-reader' of tonal music.

Winograd (1968) programmed such a tonal score-reader, again with an explicitly 'linguistic' approach. His concern was with one aspect of musical analysis, that of tonal harmony in polyphonic

music. When a performer reads a passage of such music, one of the things he does is describe certain groups of notes which occur simultaneously as chords, to assign certain functional descriptions to them, in terms of their relation to some tonal centre, and to describe sequences of such chords, in the rich structural-descriptive vocabulary of harmony theory. If the piece were Beethoven's 4th Piano Concerto, for example (see below), he might say that the first long drawn-out chord was a major chord, identify its root or centre as G, its inversion or arrangement of the notes relative to the root as 'root position'. He would observe the key-signature to be that of G major, and therefore describe the function of the chord in the larger context as that of tonic. He would also say that this chord is only a part of a sequence of G major chords which can be considered as a whole as a statement of that tonic chord. The next group of chords are also major chords, with root D and root position, and are related to the previous chord and the key as their dominant, or chord of the fifth. This is followed by another sequence of tonic G major chords, and form a higher harmonic unit with their predecessor, as a cadence, from dominant to tonic, and if this cadence is taken together with the initial statement of the tonic they make a still higher unit.

Handwritten musical notation for a piano concerto, showing a sequence of chords in G major. The notation is written on a grand staff (treble and bass clefs) with a key signature of one sharp (F#). The chords are labeled below the staff: G, D, G, Am, D, D7, G^{sus} 4, D.

Winograd's program was designed to carry out this kind of analysis. There are other things than chords in such music, and analysis is concerned with other things than their relations. Bars four and five of the above fragment contain notes which do not occur with others, and must be described as scale movements. Such aspects of the music were not his concern, and so the input to his program was provided by first converting the pieces concerned into pure lists of chord groups of notes, either by omitting non-chord notes, or including them in the surrounding chords. Naturally this limited the range of the subject matter to those pieces for which this could be done without doing intolerable violence. Nevertheless, an attempt to do this in as straightforward and 'programme-like' manner as possible, and the work should still be classed as a 'score-reader', since the result is input to the program with two kinds of information that are given in scores, and not in music that is simply heard. These are the key signature, and the notation of the notes on the staff. This means that, as in the description above of the Beethoven, the first chord can be identified as tonic, and for example, that the F# of the second chord is identified as such, rather than being possibly Gb, a possibility that someone who only heard the music would have to

consider, no matter how briefly. [The identification as F# makes the identification of the chord simple.] The effort of the program is in deciding that the G and D chords are tonic and dominant of a key or tonality, as opposed to, say, subdominant and tonic respectively of G, or even as a non-harmonic passing chord onto the major triad of the minor third of A minor. A musician's reading of a score involves decisions between such alternatives, and this is what Winograd's program did. It parsed the list of chords in terms of the categories (cadence, modulation, etc.) of textbook harmony, systematized within a grammar-like framework. As we have seen, these descriptions of a given piece are ambiguous, and the program compares the various descriptions on grounds of complexity, and yields the simplest reading. As a theory of tonal harmony, or of perception of tonal harmony, the actual program only adds to traditional theory a proof that it can be made formal,⁷ as well as an illuminating systematization. The computer affords an implementation of existing harmony theory, rather than helping in its formulation. The main interest of the work is in the extension of the definition of the parallel between language and music by the identification of concepts corresponding to linguistic syntax and semantics in music. For Winograd, a chord is a syntactic category, (trivial in his system), and its function, as tonic of one tonality, or dominant of another, say, is a semantic category, or what it means. His implementation, which foreshadows his later work in natural language in its interactive use of rules

in the two domains in the form of systemic grammar, is not a direct concern here, but the basic distinction, so intuitively clear in the example of chord function, will be developed further in the pages to follow: like Winograd, we shall consider such specifications of relation as statements of the meaning of the piece.

But perhaps this analogy between the written forms of music and language is misleading. A written sentence simply consists of a segmentation of the utterance, with no explicit indication of structure. A written piece of music contains much more information than a mere segmentation of the piece into notes, as we have already seen. It also has all the structural information that the above examples of 'score-reading' programs used, as people use it, in 'interpreting' the events of the melody - the key-signature, time-signature, and other notational information. But, while a performer needs this structural information, say to sight-read at a reasonable speed, the listener does not have it: the key signature is not played, so he has to infer from the notes of the piece which of them is the tonic, and whether the mode is major or minor. He must do this, if he is to understand what the piece means in the sense defined above, and it is clear that he does. It is also clear that the listener infers other structural information of the kind that is indicated in the score, for example the metric structure. Thus, while written language represents a rather 'primitive' level in the analysis of spoken language, the score represents a more sophisticated analysis of the musical utterance.

(A more direct parallel to written language for music would be a Pianola-roll. This is a paper-tape which controls an automatic piano. Each tape specifies the notes of a piece, as to their piano-note pitch, their duration, and the time or order in which they occur. (On real pianola rolls there are a few more parameters controlled by the tape, but they will not concern us here.) Thus the pianola-roll corresponds to the same kind of basic segmentation of the 'utterance' as written language.)

We can take as a paradigm of this kind of interpretative activity in music the example of a class of students taking musical dictation. These students are played a piece on the piano and required to write it down in the standard notation. For the purpose in hand we assume that the piece is played 'dead pan', that is, every note is played with the same loudness, and with phrasing, and note-durations as exactly as written in the original score. If we examined the results we should find complete unanimity upon certain aspects of the structural descriptions of the piece as revealed by each student's transcription. They might well have written the piece in many different keys, but all would agree as to which degree of their key-scale the melody began. Similarly, all would agree as to the positions of the bar-lines, the groupings of the notes within the bars, and whether the first note of the piece fell on a downbeat or an upbeat of this metre. If there were accidents, then they will agree in each case as to whether it is a sharpened or flattened degree of the scale they have used.

In short, at this level of analysis there is a clearly defined right answer. The following two scores are both 'interpretations' of the same 'pianola-roll' utterance - the National Anthem. The first is clearly wrong, and the second is clearly right, and reflects the way in which the melody is perceived.



The well-defined character of the task of writing down a piece in standard musical notation, and the central role played in musical perception (as revealed by the other projects discussed above) by the description of metre and key, provides the motivation for concentrating on this aspect of musical perception. The inference by the listener of these two categories of description in particular, metre on the one hand, and relative key-signature and notation on the other, have been singled out as tasks for which rules have been written. The remainder of the work concerns these two problems.

But it might seem that we have just been talked out of the

central arena of psychology concerned with faculties shared by nearly all people, into a very restricted area, where the ability being modelled is one that only a minute, highly skilled, proportion of the population possess. Most people never write down or read the music that they deal with, and would be completely unable to function in the situation of the musical dictation class, so at first glance this paradigm does not seem to come under the heading of a fact that is 'just lying about', or obvious. It seems rather to be a very artificial and contrived situation, of doubtful relevance to normal behaviour. However, this is not the case, for having once produced this concrete example of the abilities that the models concern, in sufficient detail to show that an (unnecessary) experiment could be done to demonstrate it, we can readily generalise it to more 'everyday' abilities. For example, it was said a while ago that an 'obvious' fact about everyone's dealings with music was that they could recognize the occurrence of 'wrong notes' even in unfamiliar pieces. But to say more about this ability, it is necessary to have a definition of a 'wrong note'. Some alterations we might make to melodies are hardly 'wrong' at all, while others are quite unarguably incorrect. But two ways in which notes may be said to be 'wrong' whatever other ways there may be, are when they are inconsistent with the key or metre of their surroundings. So when these two parameters of music are investigated in the context of the 'dictation-class' paradigm, we are dealing with two aspects of musical structure that

are not confined to written music, but are fundamental to all musical interpretation.

3. Two Problems.

The business of this thesis has now been defined as the writing of rules for the inference of metre, and of key and associated notation. These rules are to be written in the form of computer programs that will carry out the tasks. The input to the programs and the criterion upon which their performance will be judged are taken from the paradigm of the musical dictation class. The input is a representation of a 'deadpan' performance of the pieces as something rather analogous to a pianola-roll, a simple segmentation of the piece into notes, which are specified as to their pitch, identified as equally-tempered, or keyboard, value from 0 (for C, B $\sharp$ etc.) to 11 (for B, C $\flat$ etc.) together with the octave in which they occur, and as to their duration expressed as a multiple of the unit duration of a demisemiquaver, or thirty-second note. (Rests are specified like true notes in the input, as a triple, but their pitch is naturally represented by a special "undefined" value.) The criterion of success for these programs is the same as that for the students in the dictation class: they should reach as their conclusions in their various domains the descriptions of the piece which the composer indicated on the original score, in his time signature, bar-lines, and note-grouping, and in his key signature and notation. This criterion restricts us in our choice of data to written music for which we can be reasonably sure the composer has written it down correctly, and where distinctions he makes are really reflected in the music. This leaves us with a vast range of possible subject matter, but in the event the music of J.S. Bach has been chosen, as being rather above suspicion in this respect. In particular, since it was decided to examine these

problems in the context of unaccompanied melody, it was decided to use the subjects of the forty-eight fugues of 'The Well-Tempered Keyboard' as the data for both problems. The reasons for confining the work to the field of melody, rather than harmonised or polyphonic music will have to await the detailed formulations of each problem. But since this is the data which will be used in both, it is appropriate to describe it at this introductory stage.

A fugue (Higgs) is a musical composition of at least two, and in general more, instrumental parts, which may be played on several instruments or on one. The point of the fugue is in the relationships of harmony and counterpoint that hold between these different melodic parts, or voices. Very little of this will be concerned in the work in hand.

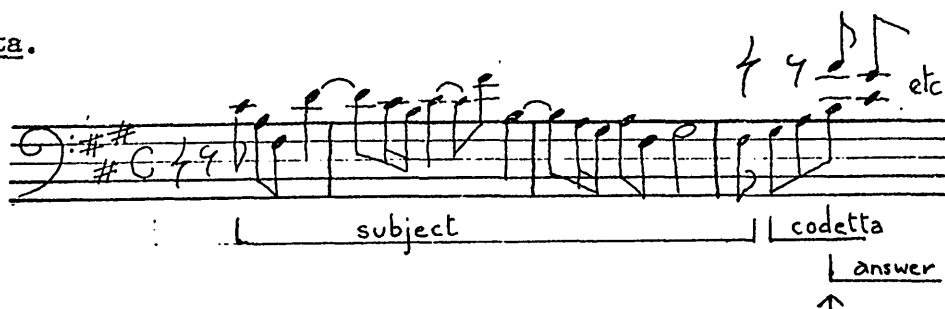
A fugue in general begins with a theme or subject stated by one voice as an unaccompanied melody. This melody or melodic fragment will form one of the bases for the development of the fugue, and will recur in the various voices throughout the fugue. When it has been completely stated, the first voice continues with further melodic material. At some point the second voice joins in with a transformation of the subject called the answer. It is a particularly simple transformation, being basically a simple transposition of the subject into the region of the dominant by a fifth upwards or a fourth downwards. (There are important exceptions to this rule, but they are of no concern here.) These are the only parts of fugue that are referred to

in what follows, and the answers will only be referred to incidentally. The main concern ^{is} and the subjects which are used as the unaccompanied melodies for input to the programs.

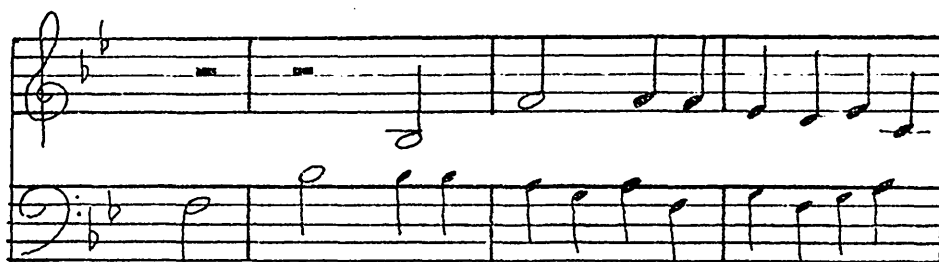
The answer in many cases begins on or during the last note of the subject.



However, the true subject, the material which appears again and again in the development, may end before the entry of the answer, the interval being filled by a quite separate fragment, called a codetta.



The subject may also extend beyond the point at which the answer enters.



In general none of these strict definitions of the extent of these melodies matter to the programs, and are not likely to until they deal with much more sophisticated aspects of melody than they do now. Rather than consider such subt^alties, the subjects of the forty-eight were just taken to be the melody of the first voice up until the note which includes the onset of the first note of the answer.

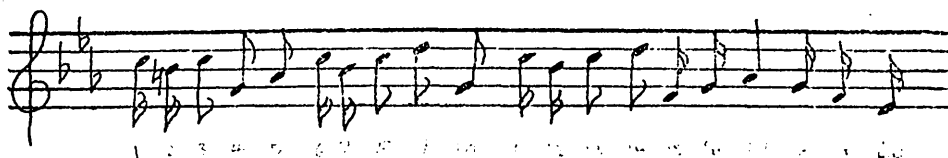
The two problems were initially considered in isolation from each other. That is, it was assumed that key and notation could be decided without reference to rhythm, metre, or note-durations, and that metre could be decided without reference to such properties of notes as their status as tonic etc. For this reason the two are dealt with in separate parts, although towards the end of each some qualifications will be made to this extreme position. (In particular, the later metrical rules make explicit reference to the key of the piece.) Nevertheless, these two parts can be read almost entirely independently of one another. The treatment of the problem of metrical inference comes first, because it is the simpler.

Part II. Metre.

1.	Introduction.	30
2.	Rules For Inferring Metre Based Solely On Relative Durations.	36
3.	Further Rules For The Inference Of Metre Based On Repetitions.	48
3.1	Introductory.	48
3.2	Figures And Repetitions.	49
3.21	Real Repetition.	49
3.22	Virtual Repetition.	52
3.23	Relation of Repetitions To Previously Established Metre.	56
3.3	Some Special Cases.	58
3.4	An Algorithm.	66
3.41	Re-formulation Of Rules.	66
3.42	Repetition Rules Restated.	68
3.421	Initial Conditions.	68
3.422	Virtual Repetition.	68
3.423	Real Repetition.	69
3.424	Remarks.	69
3.43	Other Rules Restated.	70
3.44	The Algorithm Itself.	70
3.441	Applying The Repetition Rules.	70
3.442	Applying The Alternation And Scale Rules.	74
3.5	Results.	77

PART IIMetre1. Introduction

The most 'deadpan' performance in the world of the following melody will lead us inexorably to certain strong ideas about its metre. Certain notes are felt to be 'accented', hence to fall on the stressed first beat of metric groups.



semiquavers seem in the context of their neighbours to be comparatively unaccented. The suggestion seems to be that they fall on an offbeat of the higher metric unit, and that it is a binary, which establishes the next higher level of the metre. Finally, the similarity of the group of notes 10 to 14, both to the group 5 to 9 and to the first four notes when the rest established in the last stage is included with them, seem to establish two higher metric units, both of them binary and with the stresses on the beginning of the 'repeating' figures, establishing the half-bar and the whole-bar respectively. This is sufficient to specify to the time signature exactly as Bach wrote it:



In fact it also specifies a fact that Bach expressed in the 'bracketting' of the quavers and semiquavers, namely that the four crochets, or quarter-notes, of his time signature are each made up of pairs of quavers, and these in turn are made up of pairs of semiquavers.

As in this example, we shall often be concerned with statements of ^f/ metre or time signature which are either more or less detailed than the conventional time signature, expressed as the ratio of two numbers. For example, for a time signature such as 6/8, six quavers to the bar, it is always understood that the

grouping within the bar is that of two groups of three, rather than the reverse, which would receive the time signature $3/4$. For the purpose in hand a slightly more explicit notation will be useful. We shall express these two versions of a metre with six quavers to the bar as $2.3/8$ and $3.2/8$ respectively. A time signature of say, $6/8$ as a result of one of the programs will indicate that it has failed to detect the detailed inner structure of such a metre. Finally, if such a program yields a wrong analysis, either by grouping the wrong number of subunits, or by getting the phase, or position of the stress, of the group wrong, then the level of grouping at which it went wrong will be marked with an asterisk. Thus the result $3*.2/8$ means that the program grouped the quavers, or eighth-notes, correctly in pairs, but was wrong in grouping the pairs into threes. In the case of the C minor fugue examined above, we would say that the time-signature was $2.2.2.2/16$, or $2^4/16$ for short.

In that discussion, it was implied that a note which is long with respect to its predecessors sets up a metric unit, at least when it is as simple a multiple of the previous unit as two or three times. However, in the same melody, at note number 17, a long note occurs, in the shape of a crotchet which does not occupy a crotchet metric unit, but rather straddles two of them. We could make this absolutely explicit by writing the crotchet as two quavers 'tied', i.e. to be sounded continuously, across the metric boundary.



But any program which drew any inference from this long note would be wrong, and wrong in a way which people are not. We easily recognise that crotchet as indicating an accent which is at odds with the metric structure - a so-called syncopation. While the intuition, which ^{we} will eventually make into a formal rule, correctly reflects the fact that we perceive the syncopated note as a rhythmic accent, we need a further statement to say when a rhythmic accent can be taken as implying a coincident accent of the metre, and when, as in this case it is, it may not be coincident with the metre. It turns out that the appropriate place for such a statement is as an aspect of an algorithm, or procedure, rather than the above kind of rule.

This principle was implicit in the earlier discussion, for we talked about each piece of evidence under the assumption that the metre established by earlier evidence could be assumed to be established. Moreover, we did not go on to talk about the evidence that the crotchet at note 17 might offer, simply because we had already established a crotchet metre against which it could be seen to be a syncopation. We tacitly assumed that the earlier information had precedence over the later. The earlier could be assumed to be a rhythm consistent with, or (to use Cooper and

Meyer's (1960) term) congruent to, the metre, even though this caused later information to be non-congruent, or syncopated. But it would be absurd to assume the reverse, where later information established a metre against which earlier rhythm was seen as non-congruent.

This principle will be referred to as 'The Principle of Congruence', and is stated in a general form, for simple as it is to state in the context of metre, it is the principle under which all of the algorithms and corresponding programs are designed, in both the domains of metre and key-identification. Both metre and key are regarded as structural 'frameworks', against which the rhythmic and harmonic events of the piece are to be respectively viewed. Rhythmic accents or pitches may be congruent, or non-congruent to these frameworks, but the principle of congruence states that

"No non-congruent event will occur until such a point in the piece that the event can be perceived unambiguously to be non-congruent, by the prior establishment of sufficient framework".

In other words, in a 'well-formed' melody, the early rhythmic (and harmonic) events can be assumed to be congruent, for a melody that did not establish these frames before departing from them could not be understood. It is a corollary of the Principle, that once a frame or part of a frame is established, it cannot be overthrown, (except by a completely fresh start, as when a piece has a double

bar, or a modulation.)

The rest of this section will be concerned with capturing intuitions as to the features used to infer metre in unaccompanied melody, such as were offered above, using as subject matter the fugue subjects of the Forty-eight as defined in Part I. In the discussion of the C minor no appeal was made to features of the notes such as their harmonic function in the key, for example, to the fact that the tenth note of the subject is the dominant of the key, and as such might be argued to be accented. In particular, certain of those remarks were phrased entirely in terms of the relative durations of the notes, completely ignoring their pitch, when it was said that the third note was long with respect to its predecessors, and the sixth and seventh were short with respect to their neighbours. These and other purely durational cues for the inference of metre will be discussed first, in the section that follows. The kind of repetition of similar fragments of melody that, it was suggested, was a cue for setting up such high metric units as the whole bar, are considered separately, in a later section.

2. Rules For Inferring Metre Based Solely Upon Relative Durations.

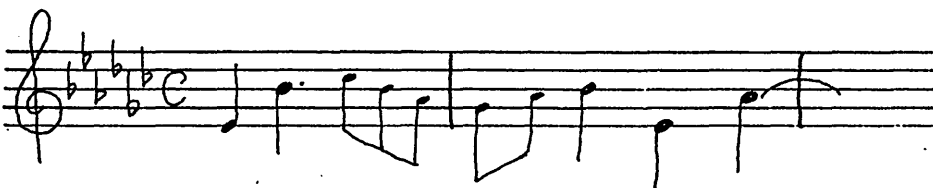
All statements so far have been in terms of building new metric units out of old. This process must have a starting point, in a basic metric unit. We define the basic unit for a melody by a general rule, which is a special case of the Principle of Congruence. This is that, whatever its length, the first note of a subject (or the first two, if the second is shorter than the first and third (see below)) may always be taken to define a metrical unit at some level in the hierarchy - though usually at rather a low level. This may be seen as a manifestation of the rule of congruence, and would, for example, exclude the unlikely possibility that the first note of the subject occurred at the end of a bar and was tied over the bar line. [Once a metrical unit has been adopted it is never abandoned in favour of a shorter one, or another one which cuts across it.¹ The only way, then, in which the metrical hierarchy can be built up is by the progressive grouping of metrical units into higher units. Any rules for doing this must clearly be framed with great circumspection, since any mistake, once made, will vitiate all that follows.

What, then, is good evidence for enlarging a metrical unit which has already been established? A natural suggestion, offered earlier, would be that if a note, falling at the beginning of a metrical unit, lasts for two or three such units, then the metrical unit should be doubled or trebled accordingly. This suggestion would account very neatly for our appreciation of the metre of *Henrich*

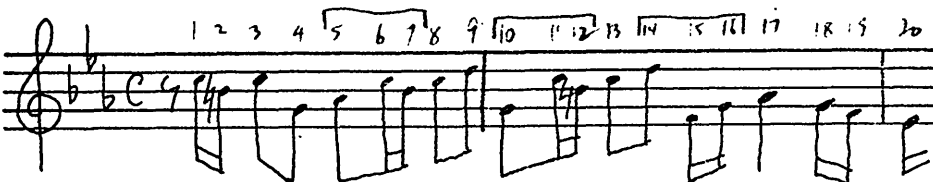
Fugue of Book I,



which begins with a quaver, establishing a quaver metre, and continues with a crotchet, which lasts two quavers and therefore becomes the new metric unit. And stated thus cautiously, the suggestion does not, ⁿ Fugue 8 of Book I



call upon us to abandon the crotchet metre established by the first note in favour of a dotted crotchet metre, for which the unit would be only $1\frac{1}{2}$ times as long. Nor do we allow ourselves to set up units of five, seven, etc., times the previous unit: Bach never used such metres, and we shall see later that we must disallow them if the algorithm is not to make errors. But it raises a serious problem in the familiar Fugue 2 of Book I.



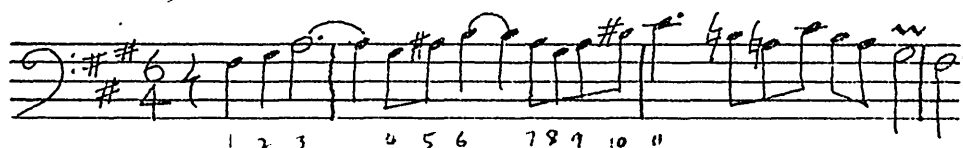
Here the semiquaver metre gives place to a quaver metre when the

third note is reached, and no problem arises for a while. But eventually we come upon that crotchet, and the above suggestion would lead us to adopt this as establishing a crotchet metre which actually cuts across the bar lines in the score. In fact no musician would be so deceived; but how does he avoid deception, and come to realise that the crotchet is in fact a syncopation, allowable under the principle of congruence?

The answer seems to lie in the early occurrence, in this subject, of a rhythmic figure which seems to play a central role not only in Bach's music but in the music of every succeeding generation. We name this figure the "dactyl" after its counterpart in the metre of poetry.¹ (This usage differs from that of Cooper and Meyer (1960)). It consists - in its simplest manifestation - of a note followed by two equal notes of half the length, these being followed by a longer note. With no exceptions in the Forty-Eight we find that if the first note of such a dactyl occupies one unit of an already established metre, then the metrical unit must be doubled so as to accommodate the dactyl. Thus in Fugue 2 of Book I notes 5, 6 and 7 form a dactyl, establishing a crotchet metre into which the dactyl 10, 11, 12 fits neatly; there is another dactyl 14, 15, 16 soon afterwards, which cuts across the crotchet metre, but is powerless to overthrow it. (There are a few organ fugue subjects by Bach in which the dactyl figure occurs in other metres, but these seem to be melodies which confuse us, as well as the rules.)

There are, however, other sorts of dactyl to which we undoubtedly pay attention in discerning metre. In the first fugue of Book I, for example, we find a pair of demi-semiquavers preceded by a dotted quaver and followed by a quaver. It is difficult to resist the view that the dotted quaver and the two demi-semiquavers provide the same metrical information as would a quaver and two semiquavers, so we therefore count this figure as a dactyl too. As a result, the initially established quaver metre is doubled so as to accommodate the dactyl, and gives way to a crotchet metre.

We may define a dactyl in general terms as 'the first three notes in a sequence of four, such that the second and third are equal in length and shorter than the first or the fourth.' But we must then be very careful how we state the metrical implications of its occurrence. Consider, for example, Fugue 14 of Book I.



The first two notes establish a crotchet metre, and are followed by a dactyl consisting of a semibreve and two quavers (followed by a minim). The dactyl is five crotchets long, but we are not allowed to group metric units into fives. At this point we could of course draw in our horns and say that only the figures described earlier count as dactyls, but it is preferable to restrict the range of the dactyl rule by a clause which says that when the general

dactyl under consideration occupies an "unreasonable" number of current metric unit, no higher grouping is to be attempted. We shall return below to the question of what can be inferred from the occurrence of an 'unreasonable dactyl.'

Earlier it was suggested that 'if a note, falling at the beginning of a metrical unit, lasted for two or three such units, then the metrical unit should be doubled or trebled accordingly.' The natural generalisation of this would be to combine the metrical units into sets of 2, 3, 4, 6, 8, 9, 12, 16, etc. if any note lasts for such a number of units. But then we run into immediate trouble with Fugue 5 in Book I.



After a flourish of eight demi-semiquavers we come upon a dotted quaver, of six times the length. If we take the dotted quaver as the new metrical unit, we make a mistake which is apparent as soon as the following semiquaver is succeeded by another dotted quaver.

The root of the trouble, clearly, is that we did not wait long enough to notice that the semiquaver following the dotted quaver was an isolated short note. For this particular fugue subject we should obtain a correct result if we treated the semiquaver and the preceding dotted quaver as a dactyl, occupying eight demi-semiquaver units, and thereby establishing a new crotchet metre.

But this expedient would lead to trouble in other connections - for example in Fugue 8 of Book I.



Perhaps erring on the side of caution, we therefore rule that when a note falling at the beginning of a metrical unit is followed by a single shorter note (which is followed by a longer note) the metrical unit is to be doubled, trebled, etc., only if the length of the first note minus that of the second exceeds the current metrical unit by a factor of 2, 3, etc.⁷ Then in Fugue 5 of Book I the demi-semiquaver unit is multiplied by 4 rather than by 6 on the evidence of the dotted quaver-semiquaver pair.

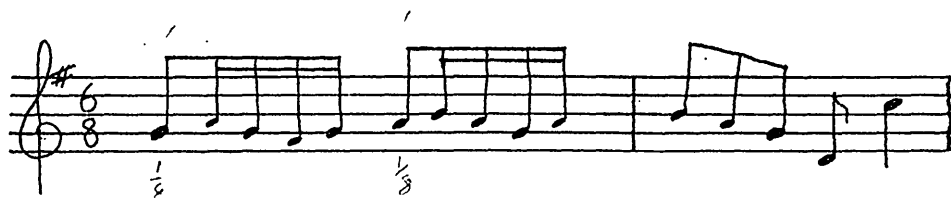
At this point it is appropriate to return to the question of what conclusions can be drawn from the occurrence of an "unreasonable" dactyl. The supplementary rule is proposed as a generalisation of the above idea, that when a dactyl occupies an "unreasonable" number of current metrical units, then the metrical unit is to be doubled, trebled, etc., only if the length of the first note minus the combined length of the two short notes exceeds the current metrical unit by a factor of 2, 3, etc. If we re-examine Fugue 14 of Book I we find that this rule leads to a correct inference - the metric signature that it finds is $3/4$ consistent with the actual time signature $2.3/4$.

We are now in a position to state more precisely the conditions under which a new metre can be inferred from the lengths of the incoming notes and rests. If a note is the first note of the subject, or falls at the beginning of a unit of the most recently established metre, then various possibilities must be explored.

(i) The note may be the first note of a dactyl - a fact which can only be established by examining the durations of the following three notes. If the total length of the dactyl is 2 or 3 [or ...] times the current metric unit, then this must be adopted as the new metric unit. If not, but the length of the first note minus the combined length of the shorter notes satisfies this condition, then this difference in length is taken as the new metric unit (as in Fugue 14 of Book I). Otherwise (but there are no such cases in the Forty-Eight) things are left as they are. (ii) The note may be followed by a single shorter note (followed in turn by a longer note) - a fact which can only be established by examining the following two notes. If so, we subtract the length of the short note from the end of its predecessor. If the result is 2 or 3 [or ...] times the current metric unit, it is adopted as the new metric unit; otherwise (but again this does not occur) the previous metre is maintained. (iii) The note may be of neither of the above types, but may endure for 2 or more current metric units. If so, these units are combined together in groups of n , where n is the largest [round number] (2, 3, 4, 6, etc.) of units which are filled by the note. (iv) If the note is not of types

(i), (ii) or (iii), then the current metre is retained.

The above procedure enables us to take a step upwards in the metrical hierarchy whenever we encounter an accented note or dactyl of sufficient length, ⁷an accented note being one which falls at the beginning of a unit of the current metre.⁷ But when applied to a subject such as that of Fugue 15, Book I,



it fails to indicate the triple grouping of the quavers in the first bar - though this is quite obvious to the hearer - so that on reaching the crotchet in the second bar it incorrectly suggests a crotchet grouping of the quavers. To deal with such cases we need to extend the concept of an accent to metrical units as well as to individual notes. ⁷A metrical unit is "marked for accent" if a note or dactyl begins at the beginning of it and lasts throughout it.⁷ We can now state a new rule for establishing metre at a higher level, as follows: ⁷if, in the current metre, a unit which is marked for accent is followed by a number of unmarked units, and then by another marked unit, which in turn is followed by an unmarked unit, then the two marked units are taken to establish a higher metre, in which they occur on successive accents. In Fugue 15 of Book I the first two quavers of the subject are, according to this definition, both marked for accent,

and the second is flanked by unmarked quaver units, as required for the application of the rule, so the rule establishes a higher metre, of dotted crotchets, in which the first and the fourth quaver units of the subject occur as the first two accents.⁷ But there is an important qualification to the rule, namely that if the higher metrical unit so defined is an unreasonable multiple of the lower one, it must not be adopted. With this qualification the new rule, which we call the "isolated accent rule", can be applied without absurdity.

We may illustrate the application of these rules in relation to Fugue 2 of *f* Book I.



The first note establishes a semi-quaver metre, to which the second note conforms. The third note converts this into a quaver metre, fitting the fourth note. The fifth note is the first of a dactyl, lasting a crotchet, establishing a crotchet metre in which the dactyl is marked for accent. Notes 8 and 9 constitute an unmarked crotchet unit; notes 10, 11 and 12 constitute another dactyl, which is also marked for accent. Notes 13 and 14 make up an unmarked crotchet unit, so the isolated accent rule can be applied to the two marked crotchet units, revealing a minim metre, which is in fact the metre of Bach's half bars. The fact that 14 is the first

note of a dactyl does not disturb this analysis, since it is unaccented in the crotchet metre which has been established by the time it is sounded.

This completes the set of rules based on relative durations of notes in the melody.⁷ They were incorporated in a program (Longuet-Higgins and Steedman (1971))(See Appendix I, section § 1) for assigning metrical structures to the 48 fugue subjects; the results are given in Table 1. By and large, the program avoids mistakes, at the cost of a rather incomplete analysis; for example, it is powerless to deal with any subject in which all the notes are of the same length, and where the musician would obviously rely on harmonic clues or skillful phrasing for discerning the metre. But there are some interesting mistakes. (See Figures below) In Fugues 8, 14 and 19 in the second Book the program makes just the sort of mistake which a musician would probably make in performing the same task. In all three Fugues the rhythm departs from the underlying metre before the latter has been made explicit; by thus violating the rule of congruence Bach plainly intends to throw the listener off the track.⁷ There are three other subjects for which the program gives a wrong result. In Fugue 6 of Book II it makes a mistake which would probably not be made by a musician who paid attention to the pitches of the notes. In Fugue 22 of Book II the program places accents on the second and fifth minim unit rather than the first and fourth, and a musician might well make the same mistake. And in Fugue 11 of Book II it

is misled by the semiquaver rests⁷ which to a musician are merely indications of phrasing and are equivalent to staccato markings of dotted quavers. In three Fugues (4 and 21 of Book I and 13 of Book II) the program actually carries its analysis beyond the bar; the resulting bar groupings seem quite acceptable, though they do not carry Bach's imprimatur. But one of these analyses (of Fugue 21, Book I) must be regarded as a fluke, because all the program does is to group quavers in twelves, without specifying whether the metrical structure is 2.2.3/8 or 2.3.2/8 or 3.2.2/8.

It is perhaps surprising that a program operating on so little information, and embodying such simple rules, can reveal so much metrical structure.

II 8

II 14

II 19

II 6

II 22

II 11

TABLE 1

Metrical Analysis

	<u>Book I</u>			<u>Book II</u>		
	Correct	Found	Comment	Correct	Found	Comment
1	$2^3/8$	$2/8$		$2^3/16$ $\frac{2}{4}$	$2^2/16$	
2	$2^4/16$	$2^3/16$	IAR	$2^3/8$ $\frac{2}{4}$	$2/8$	
3	$2^3/8$	$2/8$		$2^3/8$	$1/8$	=
4	$1/1$	$2/1$	>	$2^2.3/16$	$1/16$	=
5	$2^5/32$	$2^3/32$	IAR	$2^3/8$	$2/8$	
6	$3.2/8$	$2/8$		$2^3.3/16$	$2*.3/16$	*
7	$2^4/16$	$2^2/16$		$1/1$	$1/1$	
8	$2^2/4$	$1/4$		$2^3/8$	$2*.2/8$	*
9	$2^3/8$	$2/8$		$2/1$	$1/1$	
10	$3.2^2/16$	$2/16$		$2^2.3/8$	$2.3/8$	IAR
11	$3/8$	$3/8$	IAR	$2.3/16$	$3.2/16$	*,IAR
12	$2^2/4$	$2/4$		$2^2/8$	$2/8$	
13	$2^3/8$	$2/4$		$2^4/16$	$2^5/16$	>,IAR
14	$2.3/4$	$3/4$		$2^3/8$	$2*.2^2/8$	*,IAR
15	$2.3/8$	$3/8$	IAR	$3.2/16$ $\frac{2}{8}$	$1/16$	=
16	$2^3/8$	$2/8$		$3/4$	$1/4$	
17	$2^3/8$	$2/8$		$2^3/8$	$2/8$	
18	$2^2/4$	$1/4$		$2.3/8$	$1/8$	=
19	$3^2/8$	$1/8$	=	$2^4/16$	$2*.2/16$	*,IAR
20	$2^4/16$	$2^2/16$		$2^2/4$	$1/4$	
21	$3.2/8$	$12/8$	>,IAR	$3.2/8$	$1/8$	=
22	$2/2$	$1/2$		$3/2$	$3*/2$	*,IAR
23	$2^3/8$	$2^2/8$		$2/2$	$1/2$	
24	$2^3/8$	$2^2/8$		$3/8$	$1/8$	

> means that whole bars are grouped together.

= means that notes are all equal in length.

* indicates an erroneous result.

IAR means that that rule applied.

3. Further Rules For The Inference Of Metre

3.1 Introduction

The algorithm just described involved only rules concerned with purely durational information, ignoring pitch or sequence of pitch. The isolated accent rule is a crude identification of what a musician would call a repetition of a "figure" in a melody. It is crude because it only recognises a "figure" as a group of units the first of which is marked for accent, and only counts any other "figure" of this kind which follows immediately as a repetition. In this it is like a more general and less powerful version of the dactyl rule. (It is more general because it does not demand any particular configuration apart from the marking of the first group, and less powerful because it demands a repetition.) But the musician's idea of a figure and repetition is more powerful. Consider Fugue 6 of Book II



The old algorithm finds that the semis are grouped in triplets when it gets to the dotted quaver at note 13. But the groups of six notes 1 to 6 and 7 to 12 also form a repeating series of intervals and interval durations, in that they are each made up of three ascending scale steps and two descending scale steps,

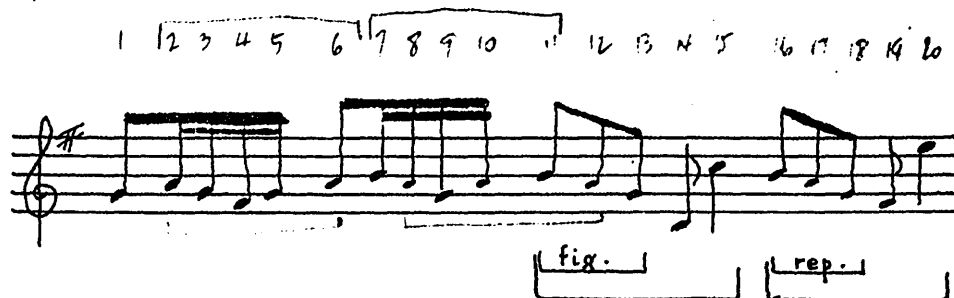
(although a major tone may be repeated by a minor tone or by a semitone.) It seems clear that the human "interpreter" recognizes this fact and infers the binary metric grouping of the triplets. It is this idea of repetition of melodic figures and the metrical inferences which may be drawn from them which the following discussion tries to capture.

We note at this point that when in the following discussion we refer to the "equivalence" of two intervals we mean that they are equivalent under the same relation that makes two notes on adjacent lines of a staff always some kind of third apart (be it minor, major, imperfect, etc.). We shall see in a later section concerning notation, that we can map the pitches of a fugue subject unambiguously onto the "functional" names (second, third, etc.) of these equivalence classes.

3.2 Figures and Repetitions

3.21 Real Repetition

In the above example of melodic repetition we have tacitly assumed a rule which says that a new level of metric group can be set up having a period (or extent in time) of the duration between the start of the figure and of its repetition, and what we can call phase (position of the downbeat) such that the accent falls on their beginnings. This definition extends to non-immediate repetition of a figure, as in the following example of no. 15 of Book I.



As in the old algorithm, we ignore such evidence if it proposes "unreasonable" groups with prime factors other than 1, 2 or 3. In order to apply this first approximation to a rule for deriving metrical structure it is clear that we must not allow any of the smaller repetitions among the first twelve notes to confuse the issue. Clearly notes 2 to 6 and 7 to 11 constitute a repetition, but the fact that earlier notes can be included means that they must be included. We may call this the "principle of earliest figures", and shall often invoke it in describing the later ramifications of our rules. In effect, it is an aspect of the familiar Principle of Congruence.

Now if we look a little closer at the earlier example of no. 6 of Book II we see that in fact the figure can be extended by one more note, running from note 1 to 7, and is repeated by notes 7 to 13, if we allow the last note of a figure and repetition to differ in duration. We do allow this, and for a very commonsense reason. The definition of a figure and its repetition is in terms of a sequence of intervals and not of notes. [The duration of the last note does not affect the last interval, and so does not come into the comparison between the figure and the repetition.]

The rule outlined above for assigning period and phase of a

metre to a melody applies to the overlapping figure and repetition in exactly the same way. The period is the lapse of time between the respective onsets of the pair (although it may, as here, be less than the length of the figure); the phase is such that the downbeat falls on both these onsets.

With one more reservation we have the nucleus of a rule. This reservation is that the figure must include at least two intervals (i.e. three notes). It is obvious that the mere recurrence of a single interval is very weak evidence for structure. It is in the context of this restriction that the overlapping condition comes into importance. In the case of a long figure and a (relatively) short overlap - as in the example no. 6 of Book II - the fact that we are allowed to make figure and repetition overlap does not afford^d us any more information, as far as metre goes, than if we had prohibited this. However in the following artificial example, in which all notes are of the same length, (and hence the old algorithm does nothing) the use of the overlapping condition allows us to establish a duplet. Otherwise we would miss this metre because of the above restriction on the length of a figure.



However the importance of the overlap condition is much greater than this, although this importance is not our immediate



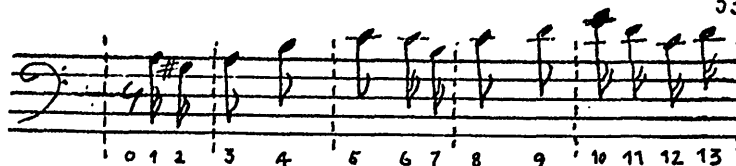
concern here. The above example if encountered in a melody would be recognized as parallel or conjunct scale movement by our human 'interpreter'. In fact all occurrences of the overlapping repeat condition constitute such conjunct motion of several 'voices', and most commonly a conjunct scale movement. The importance of the idea of overlap extends beyond the immediate problem to this neat definition of this kind of constituent of melodies.

It is in the light of the significance of the overlap condition that we make the rather obvious point that a figure is defined as extending for as long as possible, and if possible into its repetition, rather than just the minimum of notes (3) which constitute a legal figure and allow us a metrical inference. This trivial point has relevance to an algorithm for the rules. At risk of increasing pomposity we might include it in the principle of earliest figures, and rename it the "principle of earliest and longest figures".

If we apply this nuclear rule we soon run up against problems.

3.22 Virtual Repetition

In Fugue 20 of Book I, the old metrical algorithm produces the metre indicated below by dotted bar lines:



The above rule alone yields:



and 'virtualisation' yields:



The old algorithm finds a metric unit of crotchet duration (from the dactyl 5, 6, 7) and thus finds the initial quaver rest 0.⁷

The repetition rule as it stands finds the repetition of 1 to 5 in 6 to 10. The period of two crotchets is right but the phase is absurd. Not only does it not reflect what we hear: we cannot set up this phase which contradicts a lower metric unit under the principle of congruence. Clearly in some sense note 5 belongs in the repetition, but it has no "real" counterpart in the figure. But there is a sense in which the rest, established at the lower level corresponds to note 5. We say that there is a figure made up of the notes (more strictly the intervals between them) 0 (the rest) to 5 which is repeated over the notes 5 to 10 (overlapping) with a minim period and the phase shown. We call such a repetition ^{et} virtual repetition and distinguish the "literal" repetition of the early definitions by the name real repetitions. The initial variable section, where the intervals of the figure are not exactly repeated, is called the virtual part, and the literally repetitious final section is called the real part. The establishment of a virtual repetition demands the prior establishment of a real repetition, subject to all the previously stated restrictions.

The process of virtualisation does not alter the period suggested

by the real repetition, but only the phase, since the onsets of both figure and repetition move back by the same amount.

We must state some restrictions on this process of creating virtual figures from established real ones, or we shall lose many of the successes of the early formulation.

Naturally the virtual part must have the same durational composition in figure and repeat.

In the previous example of no. 20 Book I we saw that the virtual part of the figure was an initial rest. Clearly, though, we cannot unrestrictedly virtualise back over more and more initial rests, for we need never stop. That very example suggests the answer: the initial rest was independently established by the old algorithm, and hence may take part in virtualisation of the real repetition. Where no initial rest has been previously established, none may be constructed.

The intervals which form the virtual part are restricted in another way. Each pair of corresponding intervals in the virtual parts of figure and repeat must have the same sign (ascending or descending). (Scale steps (seconds) are not distinguished from other intervals for this purpose.) Thus in no. 17 of Book II.

new rules: | | | | | | | | | | | |

virtualisation to include note 3 in the figure and 8 in the repeat cannot occur, since the intervals between notes 3 and 4, and 8 and 9 are of opposite signs.

There are a few special cases.

- 1) A "tied fragment" of a long note which lasts over a metric boundary may act as a note in its own right for these purposes. Fugue 10 of Book II points out the necessity for this: without this clause the repetition of bar 3 by bar 4 cannot be observed.



The interval between a note, (or a rest, because of Fugue 19, Book II) and its tied fragment is, naturally enough, defined to be unison.

As an aside, it should be mentioned that the clause just stated leads to an unexpected bonus. The initial notes of Fugue 5 of Book I are demi-semiquavers, or thirty-second-notes, so the dotted quavers of notes 9 and 11 are described in terms of this small metric unit to be [] []. This description in terms of a note and two tied fragments, (in other words, two intervals) is enough to establish a repetition when it reoccurs later, even though the 'figure' only really consists of one note. This is a little odd, but is not obviously

going to cause any errors, since by and large, long notes will preempt its occurrence by establishing a higher metre, and not getting broken up in this way.



- 2) The pitch of a rest is undefined. We have seen that a rest can behave in virtualisation like any other note but have not yet said how the interval between a rest and a note is defined. In virtualisation the value of this interval, which is initially undefined, is taken to be equal to the sign of its correspondent in the repetition. The size of an interval to or from a rest is undefined. So a rest is always allowed to take part in the virtualisation. If both signs in the figure and repeat are undefined then they may be really equal if their sizes are also the same.
- 3) The sign (ascending or descending) of a unison is undefined like that of a rest, but its size of interval is defined (as 0). In real repetition only a unison may correspond with a unison, but in virtual repetition any interval may.

3.23 Relation Of Repetitions To Previously Established Metre

So far we have said nothing about where a figure must fall with respect to some established metrical grouping of the melody. Any repeating figure, whether virtual or real, must, to make sense,

start on the downbeat of the established metre. This fact has very strong implications for the algorithm which will apply the rules we are discussing. Most importantly it means that the process of virtualisation must take us back to a previous downbeat, or it must be abandoned. If the latter part only of a metric group is involved in the beginning of a repetition then it must not be considered as the true start of the figure, and must be discarded, whether its correspondence is virtual or real. For example in no. 1 of Book II crotchet units have been established by the old algorithm:

The diagram illustrates a musical sequence of 24 crotchet units, numbered 0 to 23. The sequence is divided into three main sections: units 0-8, 9-16, and 17-23. The first section (0-8) is a single group. The second section (9-16) is a repeating series of six units, which is repeated at units 17-23. The analysis below the notation shows three different ways to bracket the sequence:

- old alg:** A single bracket spanning the entire sequence from unit 0 to unit 23.
- repeating series:** Two brackets, one from unit 9 to 16 and another from unit 17 to 23, indicating the repeating series.
- figure & rep.** A bracket from unit 9 to 16 labeled 'fig.' and a bracket from unit 17 to 23 labeled 'rep.', indicating the figure and the repeat.

the six notes 9 to 14 are repeated (really) at the six notes 17 to 22: (dotted brackets above). However the first of these form the last half of a crotchet group. Since virtualisation cannot take us back to the downbeat (opposite signs) note 9 is not part of the figure and note 17 is not part of the repeat: (Solid brackets). We here receive a solid hint about the form the algorithm will take:

we are going to pay particular attention to the beginnings of established groups. If we don't find at least a virtual repetition between the beginnings of two groups then they cannot be the starts of a figure and its repetition, and we need not examine that pair any further.

Within the above restrictions the rule is that, once a real repetition has been established, we must virtualise as far back in the melody as possible, under the previously stated principle of earliest figures. This will also be evident in the algorithm.⁷

3.3 Some Special Cases

Before discussing that algorithm we must look at some special cases. These will require special considerations beyond the frame of the simple definitions of real and virtual repetitions and the principle of earliest and longest figures.

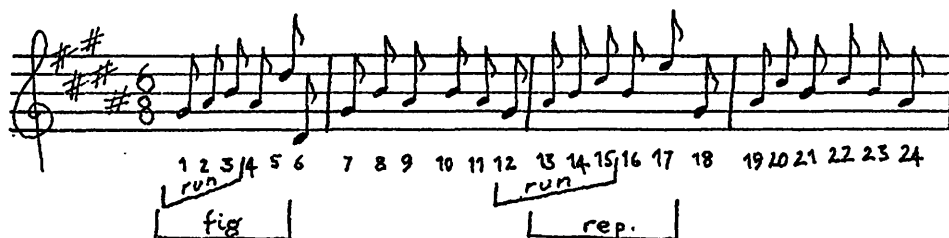
Firstly a special statement must be made about scale sequences or runs. Under the present scheme such a sequence if it contains more than three established metric units will always give an overlapping repetition, of period one and the obvious phase



This trivial information is a waste of time, nor must we automatically

assume a repetition of phase 2. We therefore make the rule that when the note of the downbeat of a group is continued by a scale movement all subsequent groups up to but not including the last to include any part of that scale movement are passed over in the examination for figures and repeats. This basically has the following consequences: a) either the beginning of the run is in a figure or repeat or none of it is; b) the above situation of a figure and repeat both within a scale is avoided.

However there is a slight complication to be made to this rule in respect of a situation illustrated by Fugue 18 of Book II.

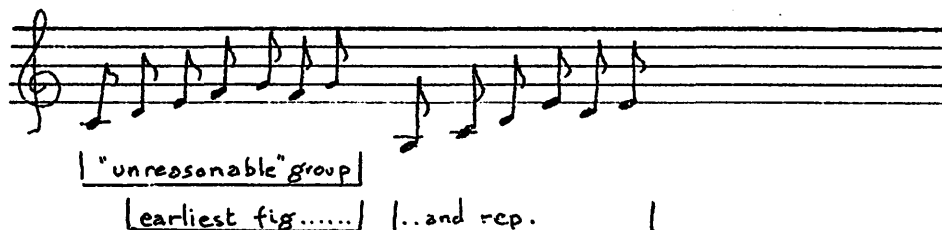


The prohibition on starting a repetition in mid-scale correctly prevents the establishment of notes 11, 12, 13 as a repetition of a figure made up of notes 3, 9, 10, which would give a wrong phase for the triplet half-bar. The repetition we want to find is a later one, where the third bar of Bach's score repeats the first. However, the repeat, in notes 13 to 16, of the first six notes, which we ourselves hear quite clearly, begins in the midst of the scale comprising notes 12 to 15!

The difference between these two situations is that, in the latter case where we do want to find the repeat, it includes a

substantial part besides the notes of the run, whereas in the latter case, where our rule is correct in ruling out the repeat, it only contains one repeating interval other than those inside the run. In other words if we allow a repetition to begin at any point in a run (instead of only at its beginning or after its end) but do not count the intervals within the run towards the real part of the repetition, (unless the repetition begins on the first part of the run), but count them as constituting the virtual part of a virtual repetition, then we achieve the desired result. The sequence of five notes at the beginning of the subject is repeated virtually at note 13, but the sequence from 8, while at first it looks as if it will be virtually repeated, does not in the end have a large enough real part to establish the (wrong) repetition. Intuitively this seems right. The repetition of 1 et.seq. by 13 et.seq. does seem to be established only in retrospect, when the real part arrives.

While we are about it we may as well generalise this new rule to include the possibility that a (virtual) figure should begin in mid run. We do this simply on grounds of intuition and symmetry, since the situation does not appear to arise in the 48 subjects.



However we can devise examples such as the above in which that statement of the new rule allows of an ambiguity: there are several possible phases for the virtual repeat (though of course only one period). Which do we want? The answer is intuitively obvious and is already established in the principle of earliest and longest figures used in the old algorithm and embodied in the old program. This principle implies a feature of the new rule, namely that in seeking to establish one of these new virtual repetitions you examine all possible figures (in order from earliest to latest) before considering the possibility that the repeat starts later in the run. [With this statement the rule is complete.]

Fugue 15 of Book II suggests another similar kind of constituent, or configuration that must receive special treatment, en bloc.



The old algorithm yields no metric structure above the note level.

The first repetition is 123/567. This gives a period of 4 semis,

and that is wrong. It does not seem to be an error a person would make: the repeated accent on the same note sounds very unlikely. This is not the only place in which this happens, so we propose a rule that when the same note alternates with others, (all being of equal duration) and there are more than three such same notes (cf. Book II, Fugue 4) a binary metre with those notes falling on the offbeat is set up. This rule must apply before the repetition 123/567 is found.

Fugue 4 of Book II, as well as the G# minor fugue that was discussed in the context of runs, show that the above definition of 'alternation sequences' must be refined.



desired result for the E minor, Fugue 10 of Book II, of establishing the correct binary metre, but it would exclude Fugue 15 of Book I, and allow the consequent error described before to arise.

Instead, we formulate the rule in the form of a hybrid of the two-sided and one-sided forms of the alternation sequence's definition. This is as follows:

[An alternation sequence is a sequence of notes, all of the same duration, in which every other note is of the same pitch, and which includes a sequence of at least three such same pitches such that the non-same pitches which intervene are either all above or all below the repeated pitch.]

In other words the new definition amounts to saying that an alternation sequence may or may not be one sided, but must contain within it a one sided alternation.

The "interpretative" rules for alternations remain exactly as before. Under the new definition the opening sequence of Fugue 15, Book II, is an alternation because notes 5 to 11 are a "one sided" sequence, and the correct binary metre is established. The opening sequence of Fugue 4 and the notes 1 to 13 of Fugue 18 of the same book, are not alternations and the incorrect binary metres are not established.

Having defined the alternation sequence a rule that uses it further is suggested by that same no. 15 of Book II. Establishing the binary metre by the above rule (which establishes the initial rest 0) does not stop the application of the repetition rule, which

would construct a higher binary virtual repetition 0123/4567.

This metre does not agree with Bach's key signature 3.2/16. It cannot be considered a failure of the rules, because $\sqrt{\text{this}}$ rhythmic structure is clearly there, and is clearly heard to be there: it is in fact repeated in the next four semis and in the four after that. Two conclusions might be drawn. On the one hand we might prohibit the repetition rule within the extent of alternation sequence, as we do during a scale: hence the repetition rule would not apply within the span of notes 0 to 11. This means that the program for these rules would reach no conclusion before it reaches the real overlapping repetition of notes 12 to 23 by notes 18 to 29, giving a metre of 3.2/16 which is exactly right.

The alternative is to allow the repetition within the alternating sequence to establish its 2.2/16 metre and to have new kinds of rules which allow the over-riding of previous conclusions by later ones. Such rules are of a totally new kind and of an obscure nature, so the first alternative will be followed here. We may note however that there is other material which strongly suggests that rules of this kind will eventually be necessary, and that if such rules produced the following structure for 15,II they would be reflecting real facts about the melody.

note numbers: (0,1)(2,3)		 (8,9)(10,11)(12,13)		... (16,17)(18...23)	
2.2/16	2.2/16	2.2/16	3.2/16	3.2/16	
3.2.2/16			2.3.2/16		

However we do not know how to set about this yet.

In the light of choosing the first alternative we must further note that we may not virtualise the overlapping repeat 12.., 23/18.. .29 back into the alternating sequence (as we may with a run) as this would give a wrong phase.

It is open to question whether in the light of the unity of the alternation sequence in 15,II, we should allow inferences to be drawn from the virtual repetition of the beginning of the sequence in the notes 12 to 15. There is no evidence from the other fugue - Number 10 of Book I - in which alternation occurs, so we may allow it. It makes perfect good sense, although it yields a bar of twice Bach's time signature.

While we are looking at special cases of the repetition rules' applicability we should look at no. 16 of Book II.



The old algorithm does nothing but establish crotchet groups. The repetition rule yields 1,2,3,4,5,6,7,8/5,6,7,8,9,10,11,12 - a triplet grouping of crotchets, with the wrong phase. This is a revealing mistake because the human interpreter will probably make the same mistake - but realize his error when he reaches the repeating C which exactly fills the metric unit of the score, and seems to establish it against all odds of repetition.

This characteristic of our own interpretative behaviour with this subject is a very clear indication that our own rules do involve the possibility of over-riding early decisions by later ones, and called the principle of congruence into open question, at least in this field of metre. In an earlier discussion of this point it was said that we would sidestep this issue for the present. We might do this here by adding a "glueing rule", (which would say that successive groups entirely occupied by one pitch tend to be "glued" together into one higher group,) onto the old algorithm. We must however realise that this is only putting off this problem. We cannot seriously believe that our present division of the metrical interpretative process into two "passes" - one for mainly durational rules of the old algorithm and one for the new melodic repetition rules - is a true reflection of performance of the task by humans.

Strength is lent to this conjecture by the fact that an extension of the "glueing rule" to trills as well as repeating pitches in a rule which could override earlier conclusions would save the error made by the old algorithm in no. 14 of Book II, if it were included in it. This would again reflect our intuitions about our own performance for this subject. But the rule has never been programmed, for the above reasons.

3.4 An Algorithm

3.41 Re-formulation Of Rules

Some rules have been stated - the various forms of repetitions, the conclusions that can be drawn from them and the two kinds of 'constituents', the run and alternation sequence. Some properties of the algorithm have been suggested in the principles of congruence and of earliest and longest figures. Before discussing the algorithm we will restate the rules in a more appropriate way. Since the most appealing kind of algorithm is one which works from left to right in the simplest possible way, that is how we restate the rules. In the following statements it is assumed that the input to this algorithm is the output of the old algorithm - that is a) a list of metric groups which may themselves be lists of metric groups or at the lowest level lists of notes and b) a notation of the pitches of those notes as sharpened, flattened or natural degrees of a key scale. When we talk about a group we mean one at the highest level of metric grouping. Groups may be identified by subscripts ordered from left to right. The groups yielded by the old metric algorithm are (nested) lists of notes. For the purposes of this algorithm we shall consider ~~them~~ as (nested) lists of intervals. An interval has a size, a sign and a duration. Its size is an integer representing unison (0), second (1) third (2) etc.: its sign is 1 if ascending, 0 if descending; its duration is that of the first note of the pair of notes between which the interval passes. (Certain special cases of size and sign concerning rests and tied notes were described before, in section 3.22) Intervals are derived from notes in the obvious way.

3.42 Repetition Rules Restated

3.421 Initial Conditions

A pair of groups $[G_i, G_j]$ ($i < j$) may respectively fall on the downbeats of a figure and its repeat if and only if:

- 1) The time interval $j-i$ which will be the ~~period~~ of the proposed repetition (expressed as a multiple of group length) has only the prime factors 1, 2, and/or 3. (As in the previous algorithm)
- 2) The first interval I_1 of each are the same in sign and duration. ?

If both these initial conditions are fulfilled, the analysis can proceed. There are three possible continuations: G_i and G_j may stand at the beginnings of a virtual or of a real repetition, or neither.

3.422 Virtual Repetition

If the initial intervals I_1 of G_i and G_j are not equal as to their third property of size then G_i and G_j may begin a virtual repetition pair.

The next intervals must also be equal at least in sign and duration in order to establish a repetition. If they are equal in these two respects, but not in respect of size then this step iterates with successive intervals. The iteration proceeds, and may pass on into successive pairs of groups $[G_{i+1}, G_{j+1}]$, etc; until it terminates in one of two ways. Either a pair of intervals are found which are not even virtually equivalent i.e. there is no repetition - this eventuality has been described; or a pair of

intervals are found which are really equivalent in size, as well as sign and duration. In this case the possibility that we are dealing with a virtual repetition is still open and depends on the extent of the real part which we have now entered.

We look at successive intervals in an iteration exactly like the one we did for the virtual part, checking this time that successive pairs of corresponding intervals in the supposed figure and repeat are equal in all three attributes. This iteration terminates if such a pair of intervals are not equal in this way. and this event indicates that we have found the end of the figure. A final check is then made that the real part contains at least two intervals.

3.423 Real Repetition

If the first intervals of G_i and G_j are equal in all three attributes then the possibility that these two groups are the beginning of a real repetition is open. To find the extent of the repetition we proceed exactly as for the real part of a virtual repetition, making the same check at the end.

4.434 Remarks

In order to check the length of the real part of either kind of repetition we must have a counter which is incremented at each stage of the relevant iteration.

Once we have established that there is a 'legal' repetition the

subscripts of G_i and G_j provide all the information which is needed to deduce the metre. The rule is that a new metric group is set up, having period $j-i$ times the old metric group, and phase such that old metric groups G_i and G_j both fall on the accent. However if we want more information - such as whether the repeat in question is real or virtual, overlapping or separate, or how long the real or virtual part is - we can insert more counters into the above procedure. Such counters would mainly count groups rather than intervals, and the way to introduce them into the algorithm is obvious.

3.43 Other Rules Restated

The repetition rules are naturally at the heart of the algorithm. It remains to describe the "special case" rules for runs and alternations.

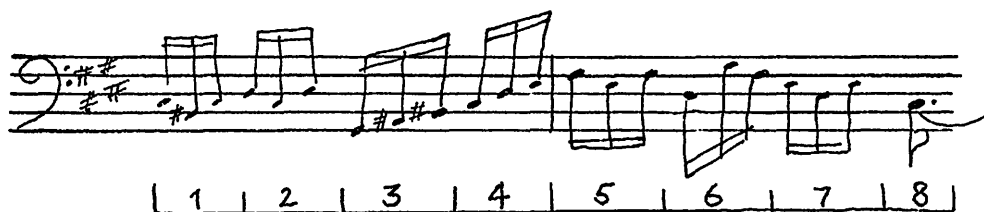
To further define the alternation and running rules we must first show a little of how this algorithm works. We have already stated the condition for a pair G_i, G_j of groups to be the start of a repetition: the algorithm is just a sensible way of ordering a search for such pairs.

3.44 The Algorithm Itself

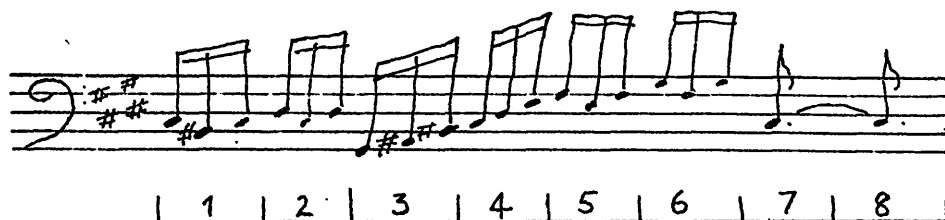
3.441 Applying The Repetition Rules

The algorithm passes along the melody from left to right in

"bites" of the current metric unit (CMU), or highest level of metric grouping achieved to date. It is equipped with a store, which is effectively a list, of previously-encountered units, onto the end of which it sticks each unit as it deals with it. Each newly-encountered unit is compared with each previously-encountered one on the list in order; that is the earliest first and the latest last. If this order of search seems less likely than the reverse order, where the most recent figure would be looked for, rather than the least, Fugue 4 of Book II shows that this is nevertheless the right order.



The algorithm just suggested - searching from earliest to latest - correctly associates triplets 1 and 5 as figure and repeat giving a metre of $4.3/16$, whereas the alternative - later before earlier - associates 2 and 5, and therefore wrongly produces the metre $3.3/16$. Theoretical considerations also show this to be the logical algorithm. We can illustrate the argument by supposing that this subject had been

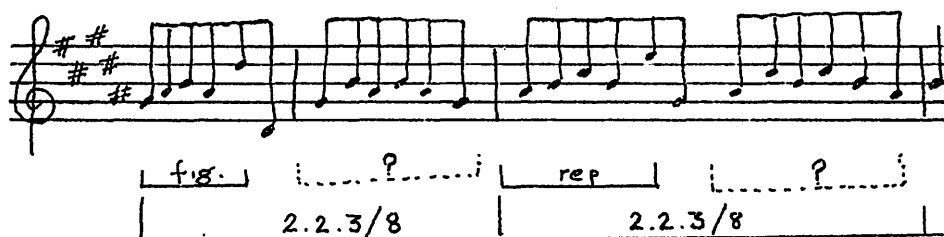


By associating 5 with 2 rather than 5 + 6 with 1 + 2, as repeat and figure respectively, we should clearly have violated the principle of earliest and longest figures.

Each previously encountered unit in the list may be the beginning of a figure of whose repeat the unit in hand may be the beginning. If this is so (§3.421) the rest of that unit and the one in hand are further examined (§3.422). This examination may continue to subsequent units (§3.422). If either the initial or the further examination fail to reveal a legal repetition, then the next group on the list is examined. If no repetitions are found over the whole of the melody to date, then the unit in hand is tagged on the end of the list, and the process repeats. If a repetition is found then the stored representation of units at the CMU level are grouped into higher metric units according to the rule of §3.4242. The analysis proceeds as before, with the new CMU.

This procedure can result in a very large metric unit - say six times the CMU. In this case we say that there is a "missing level" in the time signature, for it is uncertain whether the metre is 3.2 or 2.3. We should like our algorithm to give us the missing information if it becomes available later. (It cannot have been available earlier, or we should have found it.) When we get to the results of the algorithm we shall see that this is a fairly common occurrence. When we get a multiple with only one prime factor 4, 8, 9 etc. we could just automatically interpolate the missing levels. Where there is ambiguity (6, 12, etc.) we might

take each such group as it arrives and look for internal repetitions. This clearly helps in the case of no. 15 of Book II, but less obviously in the case of 21, II (because, although if we were given the correct six quaver bars it would give the right answer, we need to know the duplets in order to avoid the wrong answer this algorithm yields.) It does not help at all in II, 18 which we quote here without any suggestion of how to deal with it, since it is unique, and therefore inconclusive.



The obviously binary structure should be derivable. No rules for filling in "missing levels" of metric analysis have been included in the program. A convenient way of visualising the algorithm outlines above, and of looking at the slightly tricky operation of the run and alternation rules, is to think of the melody as a series of blocks (representing CMUs) which the algorithm takes one by one and places on the top of a 'tower' of previously dealt with blocks.

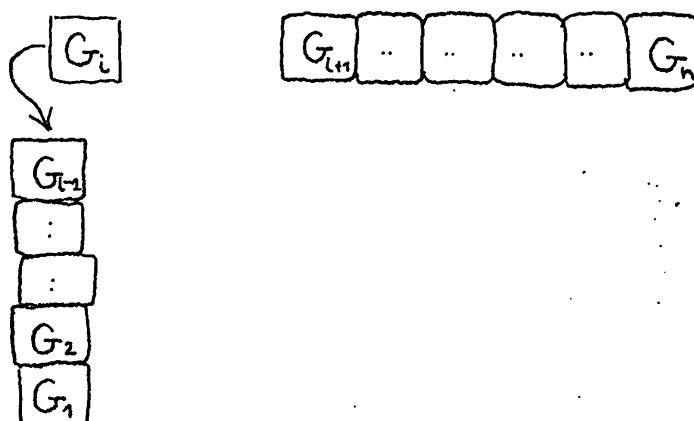


Fig. 1

Every so often the program will take the tower to bits, glue the blocks into duplet, triplet or higher blocks, then rebuild the tower with these and proceed as before. From now on, however, it will take new blocks in twos, threes or more, or in multiples thereof, according to the new CMU.

In the light of this analogy we can describe the special cases of alternation and scale sequences (cf § 3.3)

3.442 Applying The Alternation And Scale Rules

The way the algorithm realizes that it can glue blocks into higher blocks is by searching up the tower from earliest to latest and applying the test of § 3.42 to a block in the tower and the block in hand. However in our process up the store we should not consider blocks which are in a scale or an alternation sequence, except the earliest such (cf. § 3.3). In this respect scale and alternation sequences behave just like single blocks - as in the rules of § 3.42 if we don't find the possibility of a repetition

in the first part of the object, we may ignore the later part, with the exception that was seen in the case of Fugue 18 of Book II: the later part of a run must be examined for the virtual part of a repetition. However, such a sequence is not like a block in the sense that it does not represent a metric unit. For the purpose of detecting a repetition such a sequence behaves like a CMU block (though of course its anomalous length must be taken into account); but in the process of building new metric units, or gluing blocks together, it is no more than its constituent ordinary CMU blocks.

In our building block diagram we could represent an alternation series or a scale as a horizontal sequence in the tower, to represent the fact that in the sequence of blocks it has the nature of a single step

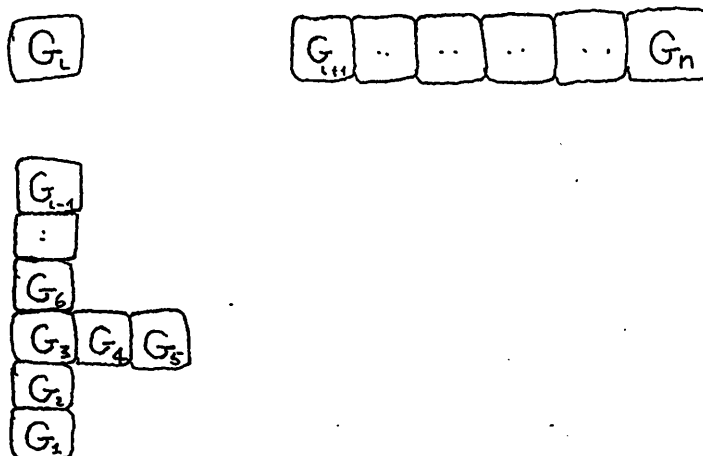


Fig. 2

Note that one of these special sequences includes only group or blocks which are made up only of notes in the sequence. If some but not all notes of the following group could be part of the scale or alternation, nevertheless it is not included.

But how is an alternation or scale sequence detected in the first place? When the program takes a new block or a number of new blocks, before it examines the tower for repetition, it examines the new unit (and the succeeding blocks if necessary) for such sequences. If it finds one such as in Fig. 3. then in the case of an alternation, it updates the tower with a binary metre, according to the alternation rule. It then, in either case, examines the beginning of the sequence for repetition of earlier material and acts accordingly. If the sequence is a run, the remaining subunits are examined for the virtual part of a virtual repetition. Finally, in both cases, the tower is updated by putting the whole sequence on top, as in (b) below.

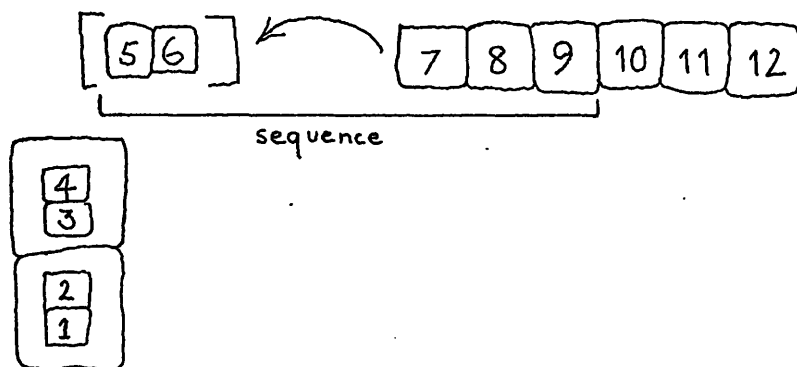


Fig 3 (a)

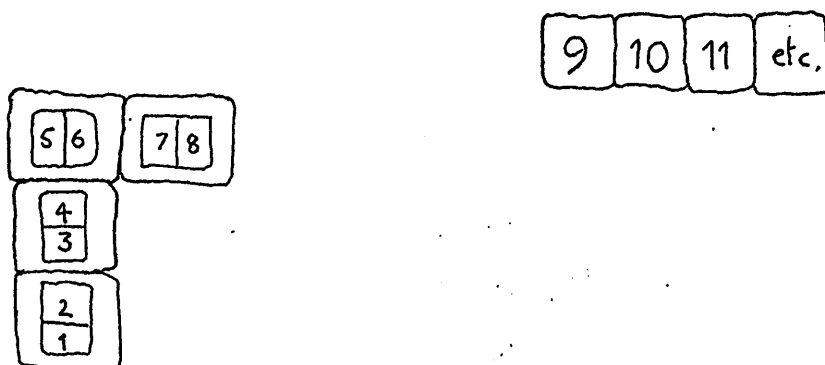


Fig. 3(b)

Note that in this case the sequence ended halfway through a CMU pair (9,10). The unit placed on the tower does not include 9, and the next group examined for repetition is 9, not 10. If an alternation sequence is found this indicates a binary metre, as described before, and this evidence has precedence over any evidence there may be in the group in hand for repetition.

3.5 Results

This algorithm was programmed and run on the fugue subjects of the Forty-eight, [using the earlier program as a 'first pass,'] as well as the first of two notation algorithms to be described in the next section (See Appendix II for details of program and examples.) The results can be summarised briefly as follows. (Subjects which contained no repetitions, hence no further metrical information omitted.)

Subjects for which this algorithm produced new, correct information:
 24
 Book I, Nos. 2, 5, 15, 20, Book II, Nos. 1, 4, 7, 10, 12, 15, 17, 18, 24.
 Total: 14
 Subjects for which this algorithm produced new incorrect information:

not adequate to have identified the extent of the figure and its repetition. Even if they were, there would still remain the problem of identifying the virtual part - the initial rest in the figure, with note 6 as its repeat - since the old algorithm does not establish that rest before the Isolated Accent Rule applies.

All of the subjects for which new information was forthcoming are listed in Table 2. The conventions are the same as those for Table 1 of this section. Again, a star against a particular grouping in one of these 'time signatures' indicates that the program made a wrong analysis, either as to period or phase at that level, and the sign 'greater than' in the margin indicates a metre which groups Bach's own bars into an even larger metric unit.

Of the four errors, one, (No. 16 of Book II,) has already been fully described. The three remaining require further comment. They are Fugues 3 and 19 of Book I, and 21 of Book II. (The successes that have not been mentioned before, Book I, Nos. 2, 3, 5, 19 and 24, and Book II, Nos. 7, 10, 12, 21, and 24, are quite straightforward, and are better described by the Table and inspection than in words.)

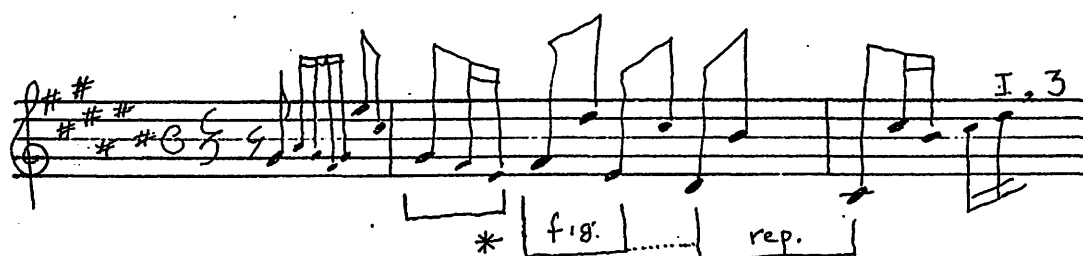
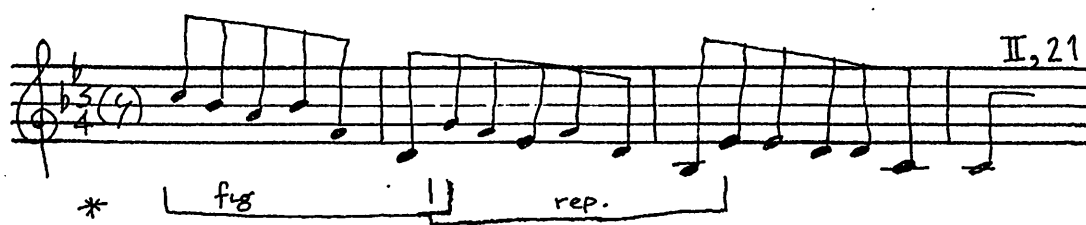
No 21 of Book II reflects a real inadequacy in the rules: the repetition is set up out of phase because the lower-level grouping of the quavers into pairs has not established the initial rest, to allow the correct virtual repetition of the first bar in the second. Yet there is something clearly wrong with the grouping that the algorithm sets up. It is particularly absurd in the third bar, but the rules completely fail to capture this fact.

Fugue	Correct	Previous	New	Remarks
Book I,				
2	$2^4/16$	$2^3/16$	$2^4/16$	
3	$2^3/8$	$2/8$	$2^*2/8$	*
5	$2^5/32$	$2^3/32$	$2^4/32$	
15	$2.3/8$	$3/8$	$3.2.3/8$	>
19	$3^2/8$	$1/8$	$2^*/8$	*
20	$2^4/16$	$2^2/16$	$2^3/16$	
24	$2^3/8$	$2^2/8$	$2^4/8$	>
Book II,				
1	$2^3/16$	$2^2/16$	$2^4/16$	>
4	$2^2.3/16$	$1/16$	$2^2.3/16$	
7	$1/1$	$1/1$	$2/1$	>
10	$2^2.3/8$	$2.3/8$	$2^2.3/8$	
12	$2^2/8$	$2/8$	$2^2/8$	
15	$3.2/16$	$1/16$	$6.2/16$	>
16	$3/4$	$1/4$	$2.3^*/4$	*, >
17	$2^3/8$	$2/8$	$2^2/8$	
18	$2.3/8$	$1/8$	$12/8$	>
21	$3.2/8$	$1/8$	$6^*/8$	*
24	$3/8$	$1/8$	$3/8$	

Table 2.

In Fugue 3, (see Table 2) the first pass sets up a metre of crotchets (quarter notes) by the 'dactyl' at the beginning of bar 2. This preempts the need for the overlapping sequence of the rest of the bar to establish the same thing. However the overlapping sequence is just long enough to allow a repeat at the next level, of the half-bar. Unfortunately this is out of phase. The overlap condition was really designed with the lower level of such 'conjunct movement' as we see here in mind and not for this sort of thing. It looks as though we should treat such sequences as this rather as we did runs, and prohibit further establishment of repetitions, once the basic metre has been established, for the duration of the conjunct sequence. This would have the attraction of reflecting the feeling that the sequence in Fugue 3 is a kind of run or scale movement.

In Fugue 19 of the same book, the error does not so much reflect a lack in the rules, as a deliberate attempt by Bach to mislead. In this subject, the conjunct ascending scale is not in a binary, but in a ternary metre. Bach violates the rule of congruence, and people show their adherence to that principle by being completely misled, and mightily surprised when the answer comes in and sets the record straight. We should not be surprised, for it is the business of the artist to take liberties with the rules, whether they are programmes', or own own.



Part III. Key and Notation

1.	The Concept Of Key.	85
2.	A Theory Of Harmonic Relations.	95
2.1	The Three Dimensions Of Harmony.	95
2.2	Earlier Theories.	111
2.3	Music In Three Dimensions.	122
3.	Two Wrong Algorithms For Identifying Key.	133
3.1	Clustering.	135
3.2	The 'M16' Algorithm.	140
4.	The Triangle Game.	156
4.1	Introduction.	156
4.11	Introductory Proposal.	156
4.12	Three Interesting Errors.	161
4.121	Initial Ascending Tone.	161
4.122	Tonic Dominant Hypothesis.	164
4.2	Playing Games With Intervals.	165
4.21	The Concept Of Key And Triad.	165
4.22	Implication Of Triads.	166
4.221	Introductory Examples.	167
4.222	Interpretative Rules.	173
	4.2221 Transitions As Chord-Diads.	
	4.2222 Further Interpretive Rule for "Run" Transitions.	
4.3	"Syntactic" Rules.	181
4.31	Introductory.	181

4.32	Further Syntax.	183
4.321	Methodological.	183
4.322	Further Definitions.	186
4.4	Notation.	192
4.41	Equal Temperament As A Separable Problem.	192
4.42	Notation In A Key.	194
4.421	Key Frames Again Plus Melodic Convention.	194
4.422	The New Semitone Rule.	195
4.4221	The Old Rule.	
4.4222	The New Rule.	
4.423	Accidentals.	202
4.43	Notation Algorithm.	205
4.5	Interpretation And Identifying Key.	205
4.51	Introductory.	205
4.511	Recap Parsing, Notation.	206
4.512	Recap Interpretation.	206
4.52	Interpreting A Parse In Some Key.	207
4.521	Implied Intervals And "Virtual" Notes.	207
4.522	Ambiguity Of 2nds And 7ths.	211
4.6	Determining Key By Comparing Interpretations.	212

PART III

Key and Notation

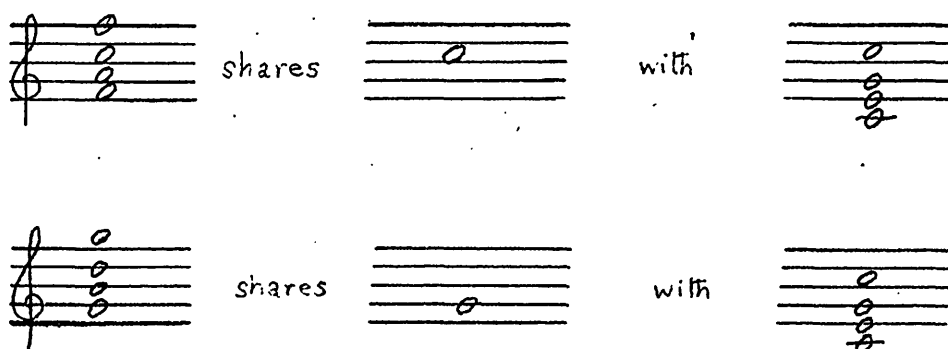
1. The Concept Of Key

Two problems were posed at the end of Part I as being fundamental in people's interpretation of music. The second of these was that of identifying the key of a novel melody input in equal temperament, together with the associated task of resolving the ambiguities engendered by that temperament in the notes as to their harmonic function, by writing them down in traditional notation. In that earlier discussion the problem of accounting for such everyday musical capabilities as the detection of 'wrong notes' was reformulated in terms of this traditional musicological term, via the paradigm of the musical dictation class. The next task is to reformulate this traditional musicological term with another one - harmony.

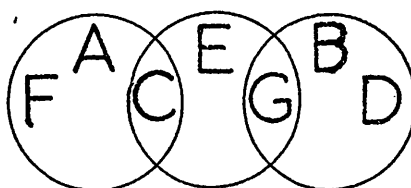
There is more to the idea of key than just a set of note-values 'expected' or 'permitted' to occur in the melody: a key is associated with three triad chords which specify a certain kind of relationship over these 'expected' notes. This relation is a part of the idea of a key.

These three chords of a key contain between them all the seven notes of the scale of that key. They overlap in this respect: some notes or degrees of the scale are included in more than one

of these three chords. In particular, the chord of the subdominant (that is F in the key of C) and the chord of the dominant (G in the key of C) both have one of their notes in common with the chord of the tonic or keynote (in C, obviously, the chord of C). For this key of C, we can illustrate this property of key, either in standard musical notation -



- or simply by representing the chords as overlapping sets of notes.



This latter representation shows that being in the key of C means more than a simple statement that we expect such notes as D, and do not expect ^{notes} such as D#. It also means that a D will, so long as we are in the key of C, make us think of a certain harmony - that of the chord of G, the dominant, in which it occurs.

This commonplace of music has very far reaching implications

for the definition of the term key which we are looking for. The figure above does not only suggest that every note of the scale can be identified with one or more of the common chords or harmonies of the key. It further suggests that there are different kinds or degrees of relationship between the various notes of the scale.

The relationship between G and D in the key of C is a "close" one - they ^{are} both involve ^d in the same harmony, the dominant chord of G major. On the other hand, the relationship of the A and the D in the same key seems to be of a totally different character.

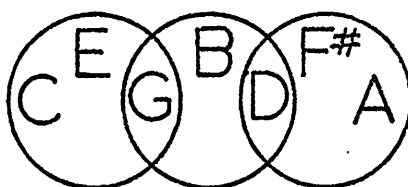
Not only are they in different chords - the chords that include them do not share any notes in common, and to get from one to the other we must pass through all three of the triad "sets" that make up the scale. In fact the relationship of D to A in this key looks a very "distant" one, compared to that of the same D to its neighbour, G.

This is a musical fact, not a theory or a consequence of the representation in the above figure. That figure just illustrates what musicians mean by the term "harmonic relationship" as applied to pairs of notes in a key. The example given of the difference in the relation of G to D from that of D to A in the key of C simply states in a convenient way a fact of music that any musician would

confirm, even to the extent of using some term such as "distance" to refer to the difference. He would say that the former was a "perfect fifth" while the latter was an "imperfect fifth", a quite different sort of interval, and a more "remote" one.

But what does all this have to do with the problem in hand, that of defining the concept of key sufficiently clearly for a discussion of how people identify it in a melody to get off the ground? The point here is to associate the idea of a set of these harmonic relations upon the set of notes of a scale, with the idea of a key. In the next example we see that a choice of key is inextricably involved with this choice of harmonic relations.

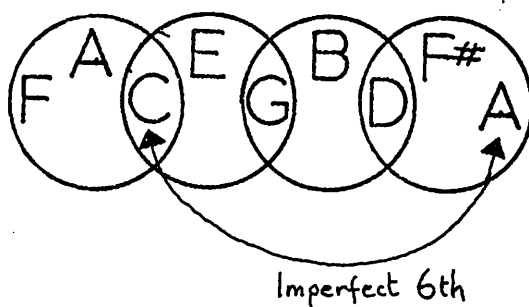
The key (as opposed to the chord) of the dominant of C (that is, G) also contains three notes called G, D and A, as we see by drawing the same kind of picture of its three chords of C major, G major and D major.



When the same comparison as before is made between the relationships of G to D, and of D to A, there is a different result. This time the relation between the D and the A is a close one, rather than the remote "imperfect fifth" relationship that they held in the key of C. In fact, their relation, being that between tonic and

dominant of the same chord of D major, is the same as that of G and D in both of these keys - the "perfect fifth".

Since the two keys are closely related, with two of their three chords in common, the two diagrams can be conflated by superimposing the chord sets that they share.



Two points corresponding to the note A appear in this diagram.

Once again, the diagram merely describes a musical fact, and this fact can be seen in two ways.

We might say that the two positions labelled "A" in the diagram correspond to different notes, that despite being labelled the same, the "A" in the key of C is a different note from that in the closely related key of G. This indeed is one of the musical facts of life. If two violin players, one thinking that he was in the key of C while the other believed the key was G, were asked to play a descending scale from E to A, they would play all of the notes in tune with one another, except the final A. (Helmholtz (1862) reports a similar experiment to show that string players do observe such distinctions.) As the diagram indicates, all the

harmonic relations defined by the two keys are identical, until the A is reached. At this point the player in C must reason "A is related to my C by the fact that both are in the chord of F. They are therefore separated by a major sixth, minus an octave. Therefore I must play a note $5/6$ ths of the frequency of our C." But the player who thinks of the key as G must reason that "A is in the chord of D major, and is therefore a perfect fifth away from the D that I just played. But D is related to our C by two perfect fifths minus an octave. Therefore I must play a note $(3/2)^3 \times (1/2)^2$ of the frequency of the C, which is $27/32$ nds of the frequency of our C." These frequencies are not identical: they differ in the ratio 80:81, so the two performers will play this note perceptibly out of tune.

But there is another way of looking at the fact that the diagram points out by having two positions for the note called A, one for the key of C and another for that of G. The piano keyboard only has one note called A in each octave, so perhaps it does not quite make sense in this context to talk about different notes with the same name. On the piano they are all replaced by a single note whose pitch represents a compromise between the several different versions involved in different keys. However, the fact that the different versions all sound alike, whichever key is involved, does not affect the fact that the harmony, or chord with which the note is associated will depend upon the key in the way described. This in turn implies that such distinctions

as that which the diagram draws between the "A" notes in the keys of C and G will still be drawn by different keys. The only difference is that in the situation above, where a performance in equal temperament obscures the distinctions of frequency that are associated with different keys, the different sets of harmonic relations associated with each are to be viewed as different sets of harmonic interpretations which will be placed on a given equally tempered interval.

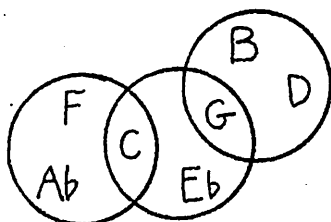
Thus, looking again at the previous case of the fifth between D and A in the keys of C and G, we may view the diagram as illustrating the fact that the former key is associated with the interpretation of that fifth as the remote "imperfect" fifth, while the latter is associated with the interpretation of that same equally-tempered fifth as the perfect fifth.

It is just such a question of the interpretation of that very interval, from second to major sixth, that gives the fifth note of the following familiar melody its peculiar quality.

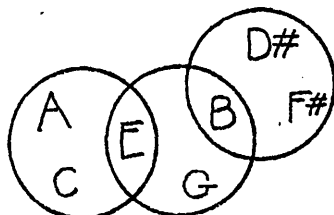


The long duration of the major sixth note, number 6, before the rather anticlimatic resolution of this remote excursion back to the right, in the dominant region where it started, and the consequent difficulty of singing it correctly, make the special imperfect nature of this interval embarrassingly clear.

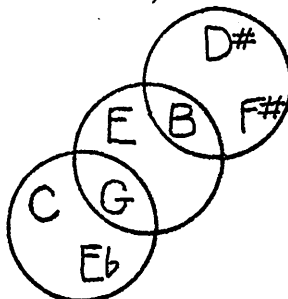
We might call these notes which are given the same name and sound the same in equal temperament "homonyms", by analogy with language. There is a similar class, which by the same analogy we should call "homophones", of notes and intervals which are also not distinguished in equal temperament. Again they sound alike, though this class of intervals are "spelt" differently - they have different names and are written differently in musical notation, unlike pairs of homonyms. An example of such a pair is offered by $D\sharp$ and $E\flat$. If we represent the key of C minor by drawing its three overlapping chords -



and do the same for the related key of E minor-



and then superimpose the notes that they share in common, again as before - (leaving some out, for the sake of clarity.)



we see just as in the case of the "homonym" pair of As, the two

functions of the note three semitones above C are each associated with one of the keys. In the key of C the note is in the same chord as the C, in the key of E minor it is in a separate chord, and stands in a quite different harmonic relationship to the same C - and in this case traditional notation reflects the distinction by giving the two notes different names.

For the present purposes of modelling the student taking musical dictation, the term 'notation' will be taken in the fuller sense to include the distinction of homonym pairs, as well as the distinction of homophones made by standard 'notation'. This fuller sense of the word is to be understood throughout, when reference is made to algorithms for achieving the 'notation' of a piece.

In short, while the ambiguous equal temperament obscures these distinctions of harmonic function, as for example between the perfect and imperfect fifths, nevertheless, a given key is associated with a single harmonic interpretation for each interval between all the degrees of its scale. To identify the key of a melody ^S~~i~~ to identify these harmonic relationships. To attain the goal of specifying rules by which people might carry out this essential part of musical interpretation, we must therefore begin by accounting formally for the nature of harmonic relation between notes of music, and explain the data which were used descriptively above in setting up the current definition of key.

In view of the vast amount of musical literature concerning

harmony, involving so many centuries and such great traditions, it might be reasonably supposed that at this point we could automatically round off the definition of key in terms of harmony, by appeal to that body of theory. This is not the case, and it will be interesting to ask why not, in the light of a recent theory of harmony which is adequate for the purpose in hand. The theory in question was originally formulated by Professor Christopher Longuet-Higgins, the originator of the project of which this thesis concerns a part, and it is the main concern of the next section.

2. A Theory Of Harmonic Relations.

2.1 The Three Dimensions Of Harmony

In three papers published in the early sixties, Longuet-Higgins (1962(a), 1962(b), 1965) presented various aspects of a theory of the fundamental bases of tonal harmony. The work related in this part of the thesis rests entirely on this theoretical foundation, so it will be necessary to describe that theory in some detail. It will be easier to see why some traditional theories are inadequate for our purposes in the light of that description, than to develop it in terms of that tradition, since the chief virtue of Longuet-Higgins' theory is its extreme simplicity.

It is not the case that any pair of frequencies, in any ratio, constitute a musical interval, as will be shown below. A minimum requirement for a theory of relations between notes of music is that it should define and denumerate the set of such relations - the set of intervals of music from among the larger general set of frequency ratios. We can arrive at such a definition by looking at a sample of some of the simpler frequency ratios, and observing a simple rule that distributes them correctly as between the two sets, musical versus non-musical intervals.

The simplest frequency ratios are to be found in those made between a fundamental and its individual overtones. These are simple integer ratios, set out in order below. If we look at each in turn we find that they are of two kinds.

<u>freq./fundamental</u>	:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<u>notename rel. to C:</u>	C	C	C	C	E	G	!	C	D	E	!	G	!	!	E	C	

If people are asked to sing the ^xintervals, or to tune two strings of an instrument to them, they find some of them very easy. For example, the first six in the series above are like this: as the note-names of the overtones of C that are written below them show, they are some of the simplest and most common of musical intervals, the tonic or fundamental, and its fifth and third, transposed by various numbers of octaves. On the other hand, some, such as the seventh in the series, are qualitatively different. People find it difficult or impossible to sing these intervals, or tune an instrument to them. In place of the note-name for the overtone relative to C the symbol ! appears - to indicate that this ratio, which has no place at all in music, cannot be named musically. (In fact its frequency relative to C is that of a rather flat B $\flat$ - but as we shall see, the genuine minor seventh has absolutely nothing to do with the seventh overtone.) Such intervals do not occur in Western tonal music, except as a joke.

If we continue to examine the series in order after this fashion we find more intervals of each of these two quite different types. The eighth, ninth and tenth harmonics are all of the 'musical' variety: they can all be easily tuned, and are the fundamental, the note, a major tone above it, and another major

third, again at various octaves which we can ignore (for now). The eleventh harmonic is another 'non-musical' noise - a flat F#, if the fundamental is C. The twelfth is another version of the easily-tunable fifth. The thirteenth and fourteenth are untuneable noises, while the fifteenth and sixteenth are the leading note and fundamental respectively - again readily tuned. We don't need to look any further in order to frame a definition which turns out to be completely general, for the class of intervals which occur in tonal music. It is 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 three prime factors, namely two, three and five, and no others. Thus in the limited set of simple ratios that we examined in the harmonic series, the 'musical' intervals were those whose frequencies were expressible as a multiple of the fundamental by the formula $\text{frequency} = 2^x 3^y 5^z$, where x y and z are whole numbers. Ratios involving higher primes, such as the seventh, eleventh and fourteenth harmonics, showed the qualitative difference of being extremely hard to tune, and were labelled 'non-musical'.

Although the rule was induced on a very small sample of frequency ratios, in fact on those with positive indices in the above equation. it is quite general. Assigning negative indices corresponds to carrying out the above experiment, where people are asked to tune an interval, between two harmonics in the series, rather than between a harmonic and the fundamental: it is clear

that if each of the harmonics is of the 'musical' variety, and they can be tuned with the fundamental, then they can be tuned with each other, if only by going via that fundamental. (We quietly took this for granted in the previous discussion, when we ignored the extra octaves labelling the ratios of the harmonics to a fundamental C ^{above} Implicitly we were using negative powers of 2.) Armed with this more general statement^{ment}, we can draw up a table of some of the more common musical intervals, identifying the values of the indices x, y and z for each.

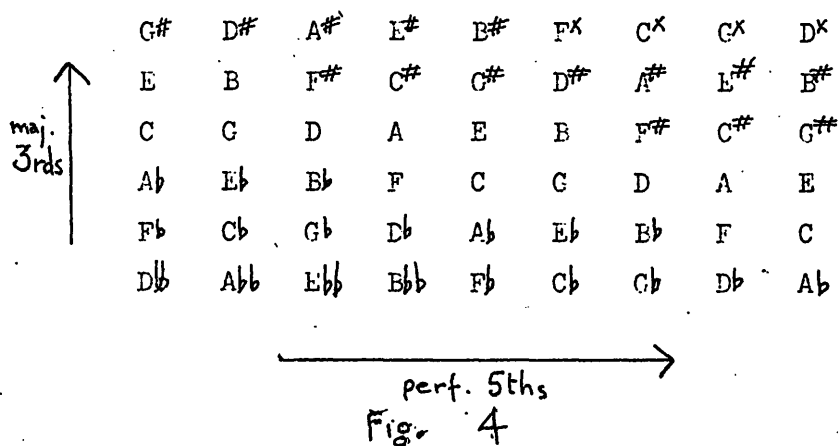
Interval	Ratio		x	y	z
Octave	2:1	2.0	1	0	0
Perfect 5th	3:2	1.5	-1	1	0
Major 3rd	5:4	1.25	-2	0	1
Perfect 4th	4:3	1.333	2	-1	0
Major 6th	5:3	1.666	0	-1	1
Minor 6th	8:5	1.6	3	0	-1
Minor 3rd	6:5	1.2	1	1	-1

Any reputable textbook of Harmony, such as Hindemith's (1937), let alone our "experiment" on people's ability to tune intervals, will confirm that these ratios are correct.

But the above statement does more than simply denumerate the set of musical intervals. It further implies that there are ^{three} degrees of freedom⁷, corresponding to the three parameters x, y and

z in the equation above, to the relationship of two pitches that form a musical interval. This is to say that the intervals may be represented as vectors in a three-dimensional space whose axes correspond to the notes reached by those interval-vectors from the origin.

The unit ratios along the three axes of this space can be identified with three unit intervals, the factor of two with the octave, the factor of three with a fifth plus an octave and the factor five with a third plus two octaves. If this space is projected along the "times two" axis, then the result is the following two-dimensional space, where the points are labelled with note-names relative to an original C, but not distinguished as to octave.



This two-dimensional projection is the version that we shall use in the rest of the discussion. We console ourselves for the loss of one of the three-dimensions by observing, in Figure 4, that the missing one corresponds to intervals which are octaves or multiples of octaves. What we have done in suppressing this dimension is in

effect to say that all such intervals are equivalent, to the unison, and that an interval plus an octave, (e.g. an interval of a fifth and of a twelfth) are harmonically speaking, equivalent. From a musical point of view, this makes perfectly good sense. Having identified the x dimension with the octave, and agreed to disregard it, then the remaining y and z axes can be associated with steps of a perfect fifth and a major third respectively. (It is just bad luck that the interval called a "fifth" is associated with the factor three, and the interval called a "third" is perversely associated with the factor of five. To avoid this confusion in future we shall refer to the horizontal axis either as such, or as the "fifth axis", and the vertical one, similarly, as the "third axis".)

Having drawn the "harmonic space" in this way, we are in a position to make some observations about its properties which will be relevant to the problems in hand.

Firstly, it is indefinitely extending in all directions.

Secondly, any of the points can be considered as the origin. The neighbours of any point stand in the same relation to it as those of any other, and although note-names were attached to the points in terms of an origin of C, any other origin would have resulted in the same diagram.

It is a consequence of the Fundamental Theorem of Arithmetic, which states that every integer can be expressed in one and only one way as a product of prime factors, that every integer ratio can

also be expressed uniquely as a product of prime factors. In particular, the particular ratios that we are concerned with in music, which as we have seen involve only the three prime factors identified above, also are expressible as the product of those factors in one and only one way. Thus, since every point or vector in the space corresponds to a different product of the three factors, and there is a one-to-one correspondence between such expressions and musical intervals, it follows that every point in the space corresponds to the result of passing through a different musical interval, and each is musically distinct from every other.⁷

Looking back at Figure 4, where we labelled the points of the space with their note-names in normal musical notation, this result appears to be contradicted: several positions with the same name appear, in fact there appears to be a periodic repetition of the whole space in a direction down and to the right. Nevertheless the foregoing result is true, and this occurrence reflects an ambiguity in our traditional notation, rather than one of the space. The Fundamental Theorem assures us of this. What then is the force of this distinction that is made in the space but not in our standard notation?

It is exactly the kind of distinction that was associated earlier with a distinction of key. We saw that the definition of a given key in terms of the three basic chords forced us to realize that, although the two keys of C and G major are closely related, and both contain an A note and a D note, nevertheless the harmonic

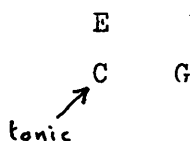
relation between the D which they have in common and A is quite different in the two keys. In fact in the "experiment" with two violin players we saw that, notwithstanding that in musical terminology the A in C has the same name as the A in its dominant G, in just intonation they are in fact different notes.

Looking for convenience at a rather smaller area of the harmonic space, we can see that that argument and those facts are reflected clearly and concisely.

F#	C#	G#	D#	A#	E#
D	A	E	B	F#	C#
Bb	F	C	G	D	A
Gb	Db	Ab	Eb	Bb	F

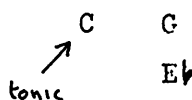
Fig. 5

There is a close similarity to the set diagrams that were used earlier to convey those simple facts. The representation used there of a chord as a set of notes is replaced by the representation as a cluster of notes in the space. For example, the chord of C major, containing the notes C, E and G, appears as the L-shaped cluster,



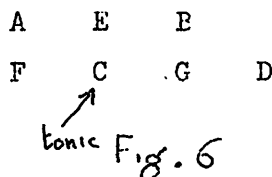
and in fact it is an obvious consequence of the way the space was generated, and of the observation that any point in it can be considered as the origin, that any major triad will appear as an

L-shaped cluster, with the keynote at the corner. (It is an equally obvious consequence that the minor triad will appear as a rotated L, with the dominant at the corner, as for example the C minor triad C, E $\flat$, G)



Whereas the full harmonic relations between all of the three notes were not spelled out in the previous presentation, they are made fully explicit here. The fact that the scale of a key is the union of three triads is again captured, and moreover the relations between all the notes are expressed in their simplest way.

Each major and each minor scale also form a consistent shape or cluster in the space. The major scale is quite straightforward—it is the intersection of three major triads, each separated from the last by one step along the fifth axis.



The minor (fig. 7) scale is also formed by the intersection of three triads. In different contexts different notes are taken to belong to the minor scale, as we shall see later, but for the body of

work that we are chiefly concerned with here, the basic minor scale, of which others are but temporary variations, seems to be the one made up of the minor triad a fourth to the left of the tonic, the minor triad of the tonic itself, and the major triad of the note a fifth to the right. For the familiar key of C, this gives the following "frame" of notes for the minor scale.

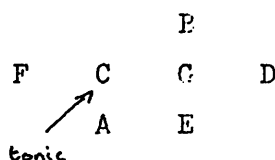
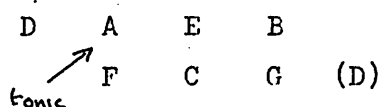
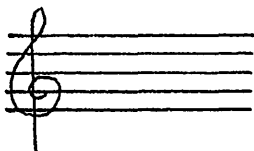


Fig. 7

This is of course not quite the same scale as is identified in the traditional key signature. A minor key is always written as the union of three minor triads. The A minor key-signature identifies the frame as



So the key signature is the same as that of C major, since both this scale and C major have seven notes with the same names. However, we have already seen how misleading this can be; only six of these are really the same, while the "D"s disguise a "homonym" pair.

It is exactly this kind of musical distinction that the harmonic space makes clear. Following the nomenclature of Longuet-Higgins (1962 (b)), all of these distinctions which are not so wildly remote as to be of little interest can be shown by

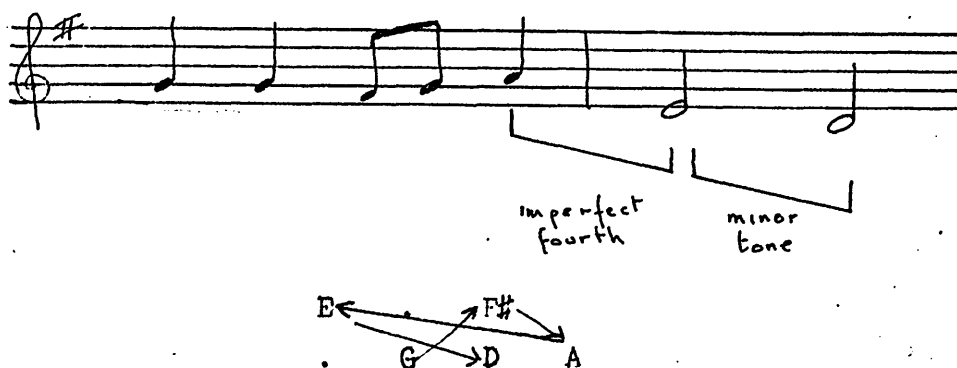
re-drawing the space of Figure 4, with the addition of the names of the intervals made by each position with the central origin, here taken to be C. (Since any point can be considered as the origin, the same arrangement of interval-names would result for any other origin. Only the note-names, here included, would differ.)

Augmented Seventh B	Augmented Fourth F#	Small Half-Tone C#	Augmented Fifth G#	Augmented Second D#	Augmented Sixth A#	Augmented Third E#
Imperfect Fifth G	Minor Tone D	Major Sixth A	Major Third E	Major Seventh B	Diatonic Tritone F#	Small Limma C#
Imperfect Third Eb	Dominant Seventh Bb	Perfect Fourth F	Unison C	Perfect Fifth G	Major Tone D	Imperfect Sixth A
False Octave Cb	Minor Fifth Gb	Diatonic Semitone Db	Minor Sixth Ab	Minor Third Eb	Minor Seventh Bb	Imperfect Fourth F
Diminished Sixth Abb	Diminished Third Ebb	Diminished Seventh Bbb	Diminished Fourth Fb	Diminished Octave Cb	Diminished Fifth Gb	Great Limma Db

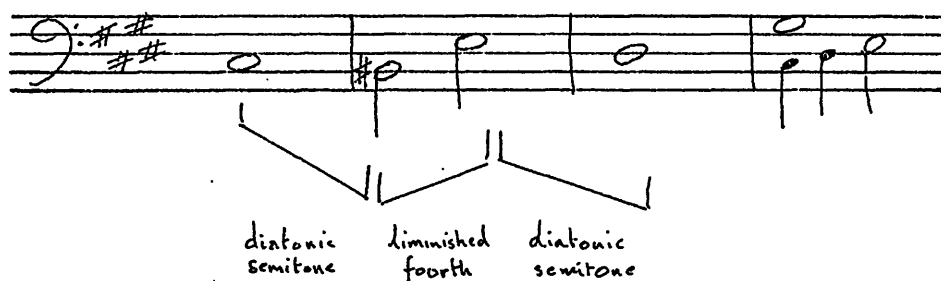
Fig. 8

The somewhat complicated nomenclature simply reflects traditional usage, which is rather unsystematic. This constitutes an account of the intervals usually recognised in music, and Longuet-Higgins (1962(b)) gives examples of the use of all of them. For the present we shall just look at a couple, which will serve to show the straightforward manner in which the space reflects the kind of things that a musician says about tunes. The first of these is the example that was used before to illustrate the imperfect fourth

- Elgar's "Land of Hope and Glory".



We can see at a glance the remote nature of the imperfect interval, the strong temptation for the singer to misinterpret it as the adjacent perfect version,⁷ the fact that this hugely emphasised excursion returns to the region in which it started, and that it does this via the harmonically rather indefinite minor tone, which does nothing to disabuse the interpreter who mistook the descending fifth for a perfect interval. We might contrast this use of a remote interval, so well suited to its subject of Pomp and Circumstance, with another.



The harmonically remote diminished fourth between the leading note and the minor third is absolutely unmistakeable, solidly anchored at each end by close and unambiguous semitones.

By now it should be clear that we can attach an objective meaning to the term "remote", which we have just been applying to the imperfect fourth and the diminished fourth, and the converse term "close" which was used of such intervals as the diatonic semitone. As we move away from the central point of the space in fig. 8, the unison, or interval of zero harmonic "distance", in all directions toward the periphery of the region illustrated, the intervals through which we move become increasingly "remote", in the sense in which the musician uses the word to describe the intervals which are most difficult to pitch correctly in a performance. Thus the space not only distinguishes the variety of musical intervals which traditional notation and equal temperament (and some theorists) tend to obscure: it also affords a measure of their difference in equating distance in the space with the subjective harmonic distance.

The simplest definition in the discrete space of harmonic relations defined by Longuet-Higgins is as the "Manhattan" distance between the points concerned - that is the sum of the number of steps in any sense on the two axes that are made in travelling along a minimal path from one to the other. This measure seems to mirror the intuitive musical dimension accurately, and in all the work to

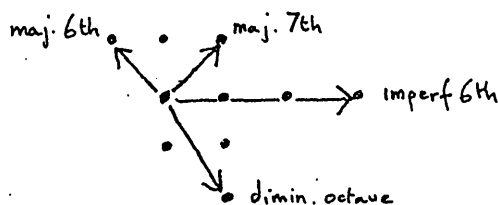
be described here, it has never seemed necessary to use any more complicated metric.

Thus, for example we see that the major third and the perfect fifth are intervals of size 1, and very different from such "homophonic" relatives as the diminished fourth of size 2, and the even more remote imperfect fifth, of size 4. Their inverses, the minor sixth and the perfect fourth, are intervals of the same size, since it is a consequence of the way the space was generated that an interval and its inverse appear in symmetrical positions relative to the origin. The two principal versions of the tone, on the other hand, seem likely to be a much more confusable pair. One, the major tone, has size 2, while the other, the minor tone, is of size 3. The major sixth and the major seventh complete the set of intervals that the degrees of the major scale make with the tonic: they are of size 2.

Whilst in the earlier stages of the argument, we appealed to musical facts which everyone must agree were true statements of the case, many people might not regard the above statements as obvious. It might seem surprising to suggest that the major seventh and the minor third are intervals of equal harmonic distance, since in terms of other musical qualities such as consonance they are so different, one being popularly viewed as a harsh sound, the other as sweet. Nevertheless, consonance, which is a property concerned in such phenomena as the "sweetness" or otherwise of an interval

or chord, the different character that they take on in various inversions, and the differing degree to which we can tolerate inexactness in their tuning, is a quite separate aspect, and is explained quite differently.

At no stage in formulating his theory did Longuet-Higgins appeal to any intuitions concerning the closeness or distance of intervals on the dimension of consonance or dissonance. The separation of consonance and dissonance from harmony is continued in the theory itself. The major seventh is like the major sixth in that it is a very unambiguous interval when played in Equal Temperament: there is little danger of confusing the diminished octave with the major seventh, (see below ,) just as there is little chance of confusing the imperfect sixth with the major sixth. The harmonic space tells us more than just that they are alike in this respect: it tells us why.



In both cases the distant homophone is more remote than the close version.

Similarly, although the very concordant major and minor triads appear in the space as the closest of clusters, there are two other equally close clusters which many would regard as discordant. Nevertheless, these clusters are extremely harmonically close and

unambiguous, as the theory suggests. In fact, with the benefit of this theory which clearly distinguishes harmony from the secondary phenomena of concord, we are in a very good position to appreciate some facts which would otherwise be puzzling, such as the wide variability of conventions as to what is and is not considered concordant. There are plentiful examples of music which uses the two "mirror images" of the major and minor triads mentioned above as concords as stable and resolved as the triads themselves, and many of them are quite acceptable to the general public, as for example, the Bossa-Nova tune 'The Girl From Ipenema'.



Some modern jazz has taken the matter even further, and often appears to have completely subordinated concord to pure harmony as it appears in the space. Thus Miles Davis, in his theme 'So What', feels free to treat a chord progression from tonic minor to flattened supertonic minor, a movement along the diagonal in much the same way as more conventional music treats a chord progression along the horizontal, say from tonic to subdominant. Having described Longuet-Higgins' theory of harmony, and made the all important distinction between harmony and concord, the way is prepared for an examination of some earlier theories, and of the

reasons why they are unusable for the purpose in hand.

2.2 Earlier Theories

The theory described here is so simple that it is hard to believe that it has never been formulated before. The facts that it accounts for have been known at least since the time of Morley (1597). Nevertheless, the article on Harmony in Grove (1940) is not only of a principally historical nature, but seems explicitly to abjure the construction of a systematic set of principles for harmony, when its author says (presumably with such theorists as Rameau or, more recently, Hindemith in mind) "The theory of the generation of chords and harmony from the harmonic series ... must be discarded once and for all."

Some early formulations come rather close. In particular the phonetician and translator of Helmholtz's (1862) work on tone perception, Ellis (1874, 1875, 1885) used exactly the same diagram as Longuet-Higgins, in talking about the various temperaments, which were his main concern in the field. He also described modulation, and ideas close to the idea of key-frames, in terms of this ⁵ diagram, which he termed the "Duodenarium". He did not, however, appear to think of it as a theory of harmony, nor as adding anything to Heltholtz's account (see below) except in providing a diagram in which to illustrate the details of temperaments.

Ellis (1885) acknowledges the influence of Weber (1817), who

also produced a similar diagram. In his case however the diagram was of keys, rather than their notes, and again his concern was with modulation. But if major and minor keys are rolled into one, the two diagrams become identical, with two axes corresponding to the fifth and the third. It may be that this theorist influenced not only Ellis, but Schoenberg (1954) who also expressed the degrees of freedom in the relation between the harmonic 'regions' with a similar diagram. However, this last theory is really very distant from the one in hand.

Weber is also of interest in the present context in that he thought that the kind of questions that have been investigated here were important, and tried to answer them, and make his theories relevant to the simple tasks involved. Thus, he asks (pp.332) "how ... is the ear determined to perceive this or that harmony as the tonic harmony", and asks how, afterwards, a modulation is detected. He couches his answers in terms that are startlingly close to the kind of procedural principles that are embraced in this work, when he says that such definitions are based on simplicity of description, and a principle of inertia, or insistence on very compelling evidence before the necessity for a change will be accepted. These statements amount to something rather like the Principle of Congruence, as he shows when he invokes them to explain why pieces so often begin with the tonic. He remarks with some surprise that this approach has not been taken

before. A hundred and fifty years later, we might feel the same surprise that this straightforward approach is just as rare.

There have been attempts to produce systematic accounts of Harmony from first principles after the same fashion as Longuet-Higgins, by first accounting for the relationships between the notes of music, and building on that sure foundation an account of chord function, then of chord sequence. One of the more recent ones is Hindemith's (1937).

Like Longuet-Higgins, Hindemith's first step is to infer a definition of the set of musical intervals, and to be able to assign to each its exact frequency ratio in Just Intonation. Both of them do this in exactly the same way, inferring the definition from the same set of simple ratios, the harmonic series, making the same observation, that ratios involving such factors as seven are not part of western tonal music, and generalising in the same way, from the set of intervals made by the harmonics whose ratios to the fundamental involve only factors of two, three and five, to intervals between those harmonics. In particular, the six ratios allow Hindemith to explain and identify the frequency ratios of the twelve semitone degrees of the diatonic scale. He does, and with the possible exception of his assertion that the "minor seventh" must always be the dominant seventh (which he justifies in terms of relations in a chromatic scale, rather than to the tonic) or subdominant or subdominant, any musician must agree with his statements, as Longuet-Higgins certainly does. (We shall see later

and can see now, just by looking at the space of figure 4, that the "minor seventh" (e.g. C to B $\flat$) is a particularly ambiguous interval harmonically). He calls the list of twelve ratios that results "Series 1" (Hindemith (1937)p.56). Hindemith does this by using the ratios of the first six harmonics as factors in generating the full set of musical ratios. These are only a redundant specification of the three dimensions, so it is clear that he can use them to correctly identify the full range of musical intervals of the other theory, and assign the correct frequency ratio to each, although since there will be more than one derivation for each, we may anticipate trouble when we get to the question of a metric upon the intervals he has defined. If we look upon Series 1 simply as a set of musical relationships, then, on the terms of Longuet-Higgins' theory, it is a small region of the set of all possible musical relationships defined by the three dimensions of harmony. Thus we can say that the two theories proceed so far in a parallel fashion, with the definition of the set of musical relationships as their first concern. However, as its name implies, there is more to Series 1 than that: it is not just a set, but an ordered set.

For both theories the next concern is to establish a formal correlate of the musical idea of harmonic distance. We saw that, in Longuet-Higgins' theory, once he had identified the fact that there were exactly three degrees of freedom to the frequency ratios

of intervals in music, and hence made the identification of a musical interval as a vector in the three-dimensional space, then the definition of harmonic distance followed willy-nilly, as the simplest possible property of the space of harmonic relations thus defined. But Hindemith has not drawn from his original definition of musical relationships in terms of the first six harmonic ratios the further conclusion that the relation is one in several dimensions. In an unstated assumption as to the nature of harmony, he reduced the relation to essentially one dimension - the ordering of Series 1. The reduction is accomplished by means of an algorithm for generating the intervals in order from a fundamental. This algorithm corresponds to a "walk" in the harmonic space, starting from the tonic origin. Series 1 corresponds to the order in which the various points near the origin are encountered during this walk. Since he has not expressed his six ratios in terms of their prime factors, the walk visits the same points several times. In general this does not matter, for he just ignores the later redundant derivations. In one case, however, an early derivation of B is passed over, and the later derivation, hence a later position in the order, substituted without explanation. This however is much less serious than the fact that the algorithm contains several assumptions which are nowhere justified, the basic one being that a linear ordering is appropriate. The fact that the major sixth appears so early in

the series, earlier than the major seventh, and even earlier than the major third, is purely as a result of these assumptions which, in the light of Longuet-Higgins' theory, can be seen to be wrong.

It is not surprising that a theory so insecurely founded proceeds to get more and more complex, with the introduction of another series, 'Series 2', and a seemingly arbitrary choice of which series is to explain which phenomena.

The key to the confusion may be that Hindemith does not appear to separate the phenomena of Harmony from those of Consonance, the effect "sweet" or otherwise", of an interval. As we saw in the previous theory, the two are quite distinct: the minor seventh is harmonically close, but also rather dissonant, or "harsh". Hindemith never quite identifies the phenomena that he is accounting for, but the distinction does not appear to be maintained in his theory. This would explain why he considered the linear ordering to be appropriate, and why he was content to pass over the earliest derivation of the major seventh. It would also explain why his Series 1 is identical with Helmholtz's ordering of the intervals as to consonance, of whose work a brief account follows.

It two tones are sounded together, and the result examined with some non-linear device, such as the human ear or any real-life resonator, several additional "combinational" tones will in general be found besides the original two and any overtones that might have been introduced with them. In particular, there will

be tones of frequency given by various functions of the original two among which the difference, and to a lesser extent the sum, are the significant ones. Additional weaker combinational tones will result from the same non-linear combination of the driving tones, their harmonics and the combination tones themselves. Helmholtz produced an exhaustive and systematic account of the phenomena of consonance in terms of the interaction of all these components of musical intervals. In the words of his own summary, his "analysis of ... sound under ... principles cited leads to precisely the same distinctions between consonant and dissonant intervals and chords as ... the old theory of harmony." In Helmholtz's theory the set of musical intervals are both distinguished from the general set of arbitrary frequency ratios, and ordered among themselves as to consonance, by two principles.

The set of musical interval ratios are those made up of two frequencies some of whose harmonics (where present) and difference tones coincide exactly with each other and/or with the original two fundamentals. By definition this limits the set of frequency ratios under discussion to the rational fractions.

Among this set there are two ordering principles. Helmholtz sees these as imposed on the simple rule that rational numbers are musical intervals, as ^{he} shows by calling this section of the book

"The Disturbances of Harmony"*. The first is the degree to which the fundamentals and the various "by-products" - the harmonics and combinational tones - coincide among themselves. (If the harmonics of a fundamental are assumed to decrease in energy with their position in the harmonic series, then a measure of the influence of a particular coincidence of two harmonics can be derived from the inverse of the sum of their orders.)

The strongest combination tones are the various difference tones between the harmonics and the fundamentals of the two notes. These, and by extension any other combination tones, and their coincidences, can be ordered as to their influence in a similar way.

The second principle of ordering the intervals as to dissonance is the degree to which all these various frequencies interact to produce beats. Beats are produced when two frequencies which differ by a small number of cycles are [non-linearly combined. A difference frequency is produced which is so low as to be perceived not as a note, but as a rattle, disrupting whatever consonant quality the interval might have. (Helmholtz says that the disruptive effect is at a maximum when the beats are at the rate of about 33 a second, though to some extent this depends on the fundamentals.)

Thus, for example, the semitone is both weakly consonant, since, as the ratio involved ($16/15$) shows, only very high and weak

* Die Störungen des Zusammenklanges"

harmonics coincide, and also strongly dissonant, since the two fundamentals themselves, (in the usual musical ranges, produce beats. On the other hand, the perfect fifth ($3/2$) has coincidences between its most powerful and difference tones, and only the weakest interactions are close enough to produce beats.

Naturally, the details of consonance and dissonance depend upon the instrument involved, since different instruments have harmonics of different strengths. An attractive feature of this theory is that it explains how it is that we tolerate different degrees of imprecision in the intonation of musical intervals. If we change one frequency of an interval by a small amount, then all the by-products that were coincident will then beat. These beats will strongly disturb a consonant interval, where they involve high-energy tones, but will affect dissonant intervals much less, since they involve less energy and do not add significantly to the dissonance already present. This is the fact that the piano tuner takes advantage of, when he puts most of the imprecision of equal temperament into the thirds and more dissonant intervals, but keeps the fifths and particularly the octaves much more exactly intoned.

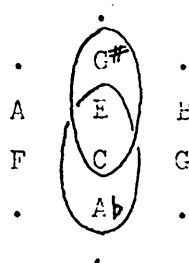
Throughout Helmholtz's development of the theory the ratios involving the prime factors seven and higher are treated just like the others. However when he goes on to build an account of chord function upon the basis of the foregoing, he needs to include some special remarks about tones derived from the seventh harmonic, since

in his system the "subminor seventh" ($4/7$) and the "subminor tenth" ($3/7$) are rather more consonant than the minor sixth and minor tenth. The special pleading he introduces is that the nature of chordal or harmonised music introduces a new constraint into play, namely that the relation of the consonances of C to each other must be reasonably consonant, and in particular that the inversion should be so. This is the case with the ratios involving only two, three and five but not for those involving higher primes. Thus in some sense he produces a reason in terms of consonance for the limitation to three dimensions that Longuet-Higgins' theory defines.⁷ However, he does not make this extension.

Helmholtz further develops his account to deal with the function of chords, in particular with the changes that the process of inversion introduces, and with the reason for the perception of one particular note as the root or harmonic centre of the chord. However he never identifies the harmonic relation in its own right, as distinct from the consonance relation, so his theory continues to be complementary to that of Longuet-Higgins, and cannot deal with the same phenomena. Most notably, he cannot deal with ambiguous chords, nor with the fact that the ambiguous intervals of equally tempered intonation are perceived in the full detail of their harmonic function, be it as the simplest or the most remote of interpretations. As we shall see, it is exactly these points which the harmonic space illuminates most clearly. It is Helmholtz himself who recognizes this limit on the power of his

theory, and poses the very problem that has been tackled in this project, when he says, in his examination of the consonant triads, and speaking of the triad C E A $\flat$, "On the pianoforte it would seem as if this triad ... must be consonant, since each one of its tones forms with each of the others an interval which is considered as consonant on the piano, and yet this chord is one of the harshest dissonances ... This chord is well adapted for showing that the original meaning of the intervals asserts itself even with the imperfect tuning of the piano, and determines the judgement of the ear."

It is only by identifying the three-dimensional space as the domain of harmonic relations implied by a chord that it can be seen how "the original meaning of the intervals" can "assert itself", and see what Helmholtz means by meaning. Although each interval in isolation could be interpreted as a major third or minor sixth, if they are examined together in the space it becomes clear that there is no way in which they can all "mean" major thirds: even when A $\flat$ is taken to "mean" G $\sharp$, one of the intervals has to be understood as a harmonically remote augmented fifth.



Moreover Helmholtz's phrase "the harshest of dissonances" is made clearer in this light: the configuration by taking $G\sharp$ with C and E looks very much like the one obtained by taking it as $A\flat$. In short, the "meaning" of the $A\flat/G\sharp$ with respect to the other two seem to be ambiguous, and of course any musician is familiar with this ambiguous nature of the "augmented fifth" chord, and would agree that this is what gives it its distinctive character.

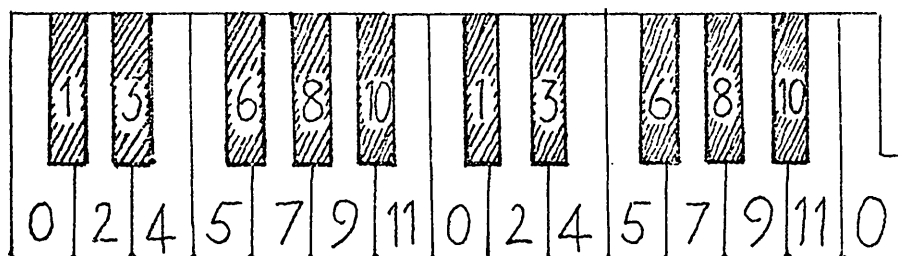
This kind of analysis, of purely harmonic phenomena, is very close to the kind that will be required for the identification of key, and is the next concern.

2.3 Music In Three Dimensions - The Disambiguation of Equal Temperament

So far, when we have talked about the nature of such things as imperfect intervals and chords, it has been more or less taken for granted in each case that each note could be unambiguously associated with a single position in the space. But when music is played in equal temperament each note is ambiguous, and is, at first glance, associated with several positions in the space (indeed with infinitely many, since the space is indefinitely extending). Our musical notation maps the points of the harmonic space onto points on a cylinder, by equating the points of figure 4 which have the same name. Equal temperament maps the space onto a torus of twelve points, but further equating those points separated by three major

thirds, or three vertical steps. The mapping can be represented by redrawing the harmonic space with the positions labelled according to the keys that would be used to play them on an equally-tempered keyboard. A convenient labelling is provided by the numbers 0 (for C, B $\sharp$ etc) to 11 (B, C $\flat$ etc.) as shown in Fig. 9.

8	3	10	5	0	7	2	9	4	11	6	9	8
4	11	6	1	8	3	10	5	0	7	2	5	4
0	7	2	9	4	11	6	1	8	3	10	5	0
8	3	10	5	0	7	2	9	4	11	6	1	8
4	11	6	1	8	3	10	5	0	7	2	9	4
0	7	2	9	4	11	6	1	8	3	10	5	0



Major And Minor Key Frames:

						11						
9	4	11				5	0	7	2			
5	0	7	2				8	3				

Fig. 9

From the discussion of that imperfect interval in Elgar's melody, it is clear that people do interpret the notes of melodies played in ambiguous equal temperament in the full space of harmonic relations. And when the concept of key was defined earlier, this was in terms of the harmonic interpretations that would be laid upon equally tempered notes. The task in hand is that of modelling people's ability to identify the key of a melody. If we can write rules that will disambiguate the notes of equal temperament, and assign them to unique positions in the space, then we shall be at least on the way to achieving that goal.

The measure of harmonic distance that the space^{affords} offers an obvious suggestion as to how this disambiguation may be achieved. [It is simply that an interval or a chord or other group of notes heard in equal temperament are to be assigned positions in the space in such a way as to minimize the sum of their distances from each other.]

This idea can be applied to chords, when it leads to a simple account of chord function and perception in equal temperament. For example, the common major chord, familiar by now as an L-shaped group in the space, is clearly in no danger of losing its identity when subjected to the ambiguities of equal temperament, if the notes are disambiguated in the above way.

8	3	10	5	0	7	2	9	4
4	11	6	1	8	3	10	5	0
0	7	2	9	4	11	6	1	8
8	3	10	5	0	7	2	9	4
4	11	6	1	8	3	10	5	0
0	7	2	9	4	11	6	1	8

Relative to an arbitrary chosen interpretation of the note 0, there is an obvious interpretation for the notes 4 and 7, as its immediate neighbours. In other words, looking back to the space of note names of Figure 4, having once interpreted 0 as the central C, we are very unlikely to want to interpret the other notes as F $\sharp$ or F $\times$, which would involve very strange interpretations of the chords function, but will rather interpret them as the straightforward major third E and fifth G in the simple major triad. The same is true for any major or minor triad, and reflects the musical commonplace that the common triads are solid, unambiguous chords. Chords such as the dominant seventh or ninth, which the musician might describe as "unresolved", appear as more diffuse configurations, while for certain others the algorithm stated above gives no single closest grouping, but many, all of which are equally close. The principal examples are the

diminished and augmented chords. The diminished is shown below:

8	3	10	5	0	7	2	9	4
4	11	6	1	8	3	10	5	0
0	7	2	9	4	11	6	1	8
8	3	10	5	0	7	2	9	4
4	11	6	1	8	3	10	5	0
0	7	2	9	4	11	6	1	8

The fact that these chords above all are recognized by the musician as completely ambiguous and dependent on their context for their interpretation lends extra support to this theory, and to the measure of harmonic proximity.

We shall see that such chords, whether they are truly ambiguous or just rather vague, require further rules for their interpretation in actual melodies. But the harmonic space identifies their ambiguous nature, as opposed to the unambiguous major triad.

The way in which our vision organises points arranged on a square lattice, in perceiving them as forming groups, is so well matched to the above 'cluster-forming' algorithm, that we have been able to take advantage of it in arranging a real-time display of the harmonic events of music played on a keyboard. This has been done, simply by making each key of an electronic organ operate a switch connected to those of a square array of light-bulbs whose

position in the array corresponds with the various interpretations of the note in question.



When a note is played on the instrument, its interpretations are shown in the array. When two or more are played, they are perceived as clusters in the ways that have been described.

This machine offers a rich vein of intuitions about our perceptions of harmonic music. All the algorithms that will be discussed in later sections start from the assumption that, when we listen to music, we identify the pitches of each note to the nearest semitone, then map them onto the full harmonic space in exactly the way that the 'light organ' maps the keys onto its display of the space. Our interpretation of the harmonic aspects of the piece are then carried out in the space.

Armed with the space, or its embodiment in the 'light organ', and these naive ideas about harmonic analysis in terms of proximity in the space, we are already in a position to say some quite interesting things about real live music. There is a rather odd harmonic effect achieved with a simple sequence of major chords in

the opening tutti of Beethoven's 4th piano concerto. When examined in the harmonic space it becomes quite clear how the trick is done. After the solo piano has established a firm footing in the key of G in the first few bars,



the strings come in with the chord of the major third, B major, emphasize the shift in tonality with a brief excursion to its dominant, F# and back to B, then proceed in unmistakeable leftward steps through the major chords of E, A and D, to the chord of G major - but it is a different G. It is with some surprise that we find that this tremendous journey has taken us to what the next few bars clearly treat as the tonic, and the surprise is well-founded as we see in the 'strip cartoon' of Fig. 11.

Precisely the same device is common in early Jazz and Swing music, and the middle section of Basin Street Blues, and Sweet Georgia Brown are only two of countless pieces involving such a sequence. Perhaps the analysis suggests a reason for its immediate success and interest at the time.

①

 F#
 . . . C . D A . . .

②

 B
 . . . (C) G D

End of solo: D7

③

 F#
 D A

 D

④

 D#
 B F#

TUTTI: B

⑤

 A#
 F# C#

 E#

⑥

 D#
 B F#

 B

⑦
 G#
 E B

 E

⑧
 C#
 A E

 A

⑨
 F#
 D A

 D

⑩
 B
 G D

 G

Fig 11: The progress in harmonic space of the opening Tutti of Beethoven's 4th Piano Concerto.



However the algorithm that the above remarks imply for the analysis of chords and chord sequences, interpreting chords in the way which gives the tightest clusters, and in the position closest to their predecessors in the space, will not do in general. In the above case, all the chords are the unambiguous major triads, and all the movements are by one step horizontally or vertically. With the more ambiguous chords, however, other things must be taken into consideration if our algorithm is not to do violence to musical good sense. For example, if the above algorithm, using simple Manhattan distances, is applied to the First Prelude of Bach's "Well Tempered Keyboard", treating each arpeggio as a chord, the very first bars illustrate the deficiency. The first chord is a simple triad - that of the key, C major. The second is the major triad of the subdominant, or fourth, F major, with an extra note, D. It is clear where the three notes of the triad belong. The question is, which D? The algorithm picks the D which is closest to the other notes of the chord - the version related to the F by a North-Westerly major sixth. However, we might suspect

that we have not yet left the key of C, in which case the D should be notated as the supertonic of C, off to the East of the subdominant. The suspicion would immediately be strengthened by the next chord, that of the dominant of C, G major, with an extra F. This chord also contains a D, one which is unambiguously the supertonic of F. But there is no reason to suppose that it is not the same as the D in the preceding chord, so it seems likely that the previous interpretation was wrong - and the fourth bar, a repetition of the original C major chord, makes this quite certain. In fact, the second bar represented a 'passing chord', a transition between the chords of the first and third bars, and not the independent harmonic statement that our naive algorithm took it as. On the other hand, if there had been some reason to suspect a modulation or shift of key to the subdominant then the other analysis of the D would have been right. If we are to write better rules for the analysis of such music, then they will have to recognise such 'passing chords', and embody the kinds of ideas that were implicitly invoked when it was said that we might suspect the analysis because we were in a particular key, and that the piece involved a transition between the tonic and dominant, in which the second bar was just a passing chord, in the absence of any evidence of a modulation. In short, we have been trying to run before we can walk in the foregoing analyses, and we look at some more elementary phenomena of music before we can return to these interesting problems. In particular, we see yet another reason to

believe that the problems connected with the nature of key, and with its identification are of central importance in our interpretation of music, so we shall get back to them immediately.

3. First Ideas For Key Identification In Sole Melodies: Two Wrong Algorithms.

3.1 Introductory

In section § 2 of Part II the Principle of Congruence was formulated, as the guiding principle of all the algorithms advanced in this work. The principle casts such properties of a piece of music as its metre and key in the role of "frameworks" of "expectations", against which the actual events of the piece are viewed and interpreted. When events conform to the framework as predicted, they are said to be "congruent"; otherwise they are "non-congruent". The principle then states that, if the events of the piece are thought of as arriving for interpretation in order,

"No non-congruent event will occur until such a point in the piece that the event can be perceived unambiguously to be non-congruent."

This is to say that our algorithms for discovering these frameworks will assume the earliest events in the melody to be consistent with, or congruent to, the framework that is really there. In other words, it is of the nature of the music we are discussing, that syncopations will not occur until such time as the earlier events of the piece have established sufficient metrical framework for the syncopation to be perceived as such, and similarly, that accidentals will not occur until it has been made clear that they are accidentals. Therefore (once we define the criterion of

congruence in question) we can draw strong inferences from the early events of the tune, ones that perhaps we could not draw from those in an arbitrary section of the piece. (The reader is referred back to Part II for examples.)

This principle may be simple, but it is not a truism. There are conceivable algorithms for key identification that would implicitly deny it - an identification of the tonic based on a simple frequency count, for instance. If it is obvious, then this is all to the good, since it is the principle that supplies the methodology for the investigation.

The principle is not a theory. We get a theory from it when we produce a definition of congruence, and this theory can be falsified if, when we examine pieces of music according to the definition we find that the definition does not ascribe conformity with the principle to the melody. In that case, since the principle is a statement of the nature of the music, and the model that we have enshrined in a definition of congruence does not reflect that nature, we must abandon the model in favour of a new one, and a new and better definition of congruence.

Thus the principle itself is not directly falsified upon the outcome of the predictions on an individual theory, but only insofar as a persistent failure to produce successful theories under it leads us to suppose that it does not represent the nature of the music under discussion. This somehow ceases to be worrying when we

reflect that it is really a sketch of the algorithms and programs that we are going to produce as theories, a sort of control routine. We have already seen it in this guise, in those parts of the metrical algorithms which assumed that early long-notes, dactyls and so on, could be used with confidence to infer metre, that is, could not under the principles be non-congruent syncopations. Later events could not contradict these inferences, since under the same principle, they could be non-congruent. In the algorithms that follow for the task of key-identification we shall see a very similar control structure, and shall find no reason to suppose that the Principle of Congruence does not reflect the very nature of at least the kind of music that we have chosen to test our theories against.

The far-reaching consequences of the principle, both as to the methodology outlined above, and as to the actual form that the theories resulting from it take on, constitutes the 'Artificial Intelligence' content of the research. The computer, besides acting as a fast and reliable version of pencil and paper verification, and besides providing a formal language for the expression of the theory, also provides the form of the theory itself.

3.2 Clustering

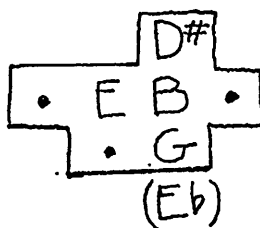
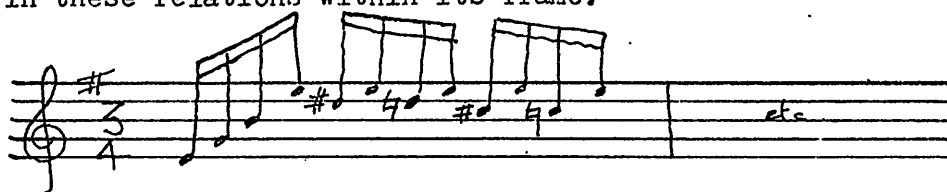
The earlier discussion of spatial clustering in the Harmonic

space as a rationale for analysis of chord function and chord sequence, although it had to be abandoned as being an oversimplification for that purpose, offers a suggestion that key-identification might be done by a simple procedure of clustering. The algorithm would run something like this.

"As each note arrives, identified to within equal temperament, write down the corresponding set of possible interpretations in the space. Choose an arbitrary interpretation (position) for the first note. Take as a set of hypotheses all those keys for which this position is non-accidental, that is, falls within the major or minor frame corresponding to the notes of the key, as in Figs 6 and 7. Choose interpretations for succeeding notes so as to minimise the sum of their distances from the cluster of predecessors. When each succeeding note has been interpreted and added to the cluster, reject those keys for which the new position falls outside the frame, under the Principle of Congruence. If only one hypothesis survives, then that is taken to be established as the key of the piece, and the procedure terminates. Otherwise the next note is examined.

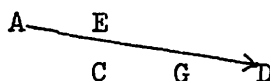
(We should have to spell out some further conditions as to cases where a note is accidental to all going key hypotheses, and where the end of the melody is reached before a unique hypothesis has been arrived at. We should also have to slightly refine the definition of a minor key. All of these questions will be dealt

with later, in the context of another algorithm. To both^{er} with them now would be to overcomplicate an algorithm which is only being set up in order to be knocked down.) This attractively simple algorithm would also solve the other half of the original problem, that of notating the melody, since it notates the melody once and for all, and infers the key of the melody from that notation. For Fugue 10 of Book I, the first four notes are correctly clustered as the minor triad, while the fifth note, placed closely beside them, is identified as the leading note, and leaves E minor as the only key containing all of these notes in these relations within its frame.



Just previously, the Principle of Congruence was described as providing a general framework for the algorithms to be discussed, and that they could be compared on the basis of various definitions of congruence that they involved. The congruence in question here is basically whether or not the interpretation of the note involved which the clustering procedure assigns to it is in the key frame (non-accidental) or accidental. However the matter is

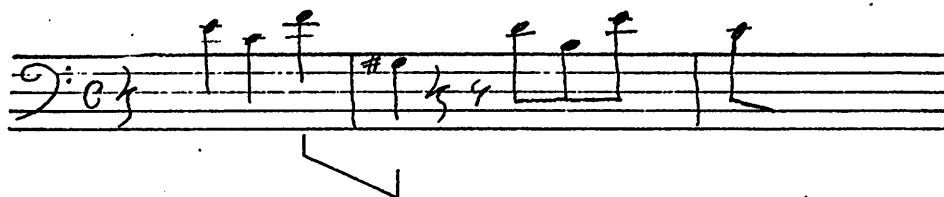
complicated slightly by the fact that this interpretation is carried out independently of any hypotheses as to the identity of the key; the clustering procedure itself involves a part of the congruence component. This part is a little hard to state in musical terms - which in itself might make us smell a rat - but it adds up to something like a rule saying that the closest interpretation of an interval must be the right one, until sufficient context has been given, in the form of a "cluster" of preceding notes, to force such an interpretation. For example, it is only with some preparation by such a full context as a major triad that the imperfect fifth down from the major sixth to the second can occur.



The fugue subjects of the forty-eight offer several examples of melodies to which the algorithm assigns the wrong interpretation, and hence the wrong key. More needs to be said about tones, for example, since it is not the case that a subject never begins with the marginally more remote minor tone, rather than the closer major tone (Fugue 11 of Book I has an initial minor tone).

Moreover, if we continue the operation after the key decision, to carry out the second part of the task in notating the melody, we find that more must be said about the notation of semitones. But one of the algorithm's wrong decisions in particular demonstrates

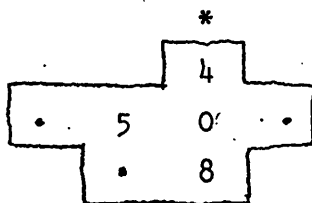
its basic shortcoming. The A minor subject, no. 20 of Book II of the forty-eight, includes an early diminished seventh, between the third and fourth notes.



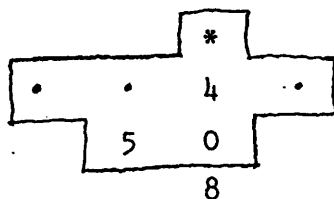
Notation of the fourth note, nine keyboard semitones away from its predecessor, by clustering with respect to the first three notes, has to choose between the two starred positions (among others) in interpreting the note.

*
4
5 0
*

By the clustering measure, the lower alternative is chosen. Thereupon, the only key-frame which fits is that of F minor,



and the true solution, A minor, is rejected.



This is a very unsatisfactory state of affairs, because this is not a mistake that a human interpreter would be likely to make: there seems to be little temptation to think of the melody as beginning on the major seventh, or leading note of F, and every reason to feel that the first note is the dominant of A, notwithstanding the rather remote diminished seventh that such an interpretation implies. In other words, while there may be some errors that we would be prepared to put up with from the algorithm, there are strong musical reasons why this should not be one of them.

What appears to be wrong is that part of the definition of congruence that militates so over-confidently against early imperfect intervals. The two alternatives that were investigated in algorithms for key identification are described in the following sections.

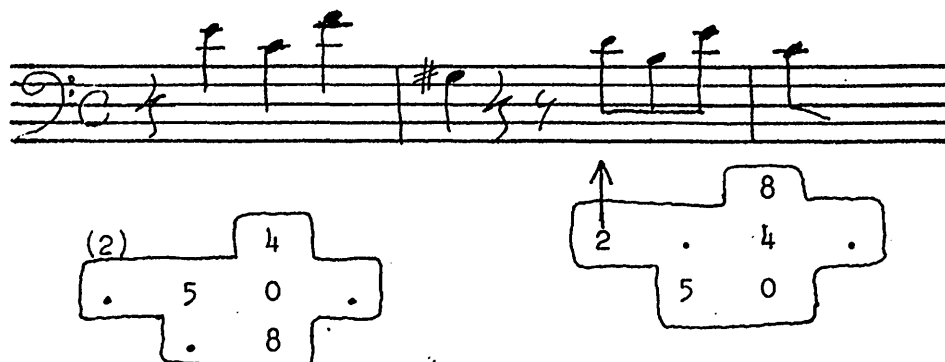
3.2 The 'M.I.6' Algorithm

In the Clustering 'algorithm', the embodiment of the idea of key was as a very 'passive' sort of thing, a conclusion drawn at some stage from an interpretation of the notes in the making of which it played no part at all. But when the concept of key was discussed earlier, it was defined in a very 'active' sense as a definition of the way in which a piece was to be interpreted, as regards the harmonic relations of its notes. The clustering algorithm, while it represented a key as a set of such relations

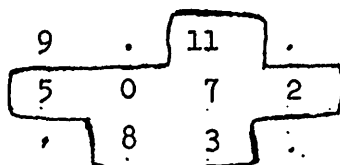
in the form of a key frame, was at the same time less than faithful to this definition in that it carried out the business of interpretation in the harmonic space without any reference to the particular key-hypotheses in hand. In the case of that A minor subject, No. 20 of Book II for example, it seems a little odd to hold the hypothesis that the key might be A minor, yet to reject it out of hand, simply because the closest interpretation in the space is not the interpretation that it involves. It would perhaps be more reasonable to associate the business of interpretation in the space with the keys that are being considered, where we have said that it belongs - in short, to associate an interpretation, in the form of a list of positions in the space, with each key hypothesis.

This was the solution adopted in the algorithm described in the second part of Longuet-Higgins and Steedman (1971). In this algorithm the order of the two parts of the task, notation of the melody and key decision, were reversed: notation was made completely contingent upon the decision as to key, rather than the other way round as in the previous algorithm. Like that algorithm, this one examines a set of key hypotheses under a criterion of congruence with respect to the notes of the melody taken in order, until only one survives. At each stage in the decision, a separate interpretation or notation is associated with each of the key-hypotheses. As before, hypotheses may be rejected when

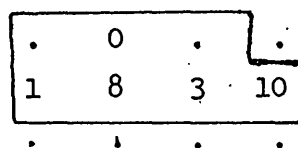
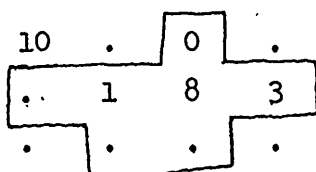
they impute a non-congruence to the melody, in favour of ones which do not, under the familiar principle. However, in this case, the criterion of congruence is much weaker. It is simply that a note is congruent to a key if it can be interpreted as being in its scale, and non-congruent only if it must be accidental. In other words, as far as the algorithm for key identification goes it might just as well ignore the space altogether and work in equal temperament. We only bother to keep an interpretation in association with each key because we shall be concerned later with what now becomes the separate problem of carrying out this notation. In the clustering algorithm, a note had to be not only possible to interpret in the key, but actually interpreted as being in the key by the independent process of clustering. So, for example, when the algorithm deals with the A minor subject that proved such a stumbling block for the previous algorithm, since the fourth note could be in both A minor and F major, albeit with different interpretations of its relation to its predecessors, both hypotheses are maintained, in association with those different interpretations. It is only with the fifth note that one of the two is rejected by the Principle of Congruence. The fifth note, of value 2, can only be accidental to F minor. It can be interpreted as D, the fourth of A minor, so A minor survives at the expense of F minor.



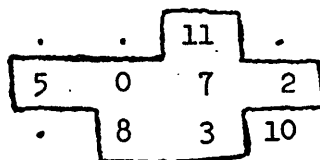
To make the algorithm work in general for melodies of this kind in minor keys, there is a slight complication to be made to the simple picture of the minor key as the frame familiar from such examples as that above. In Bach's time, if a composer wished to write an ascending or a descending scale in a minor key he would use the major sixth in place of the minor sixth of the frame, if the scale were ascending, and would use the minor seventh in place of the major version, if the scale were a descending one. That is to say, in C minor (C minor) the upward scale 7, 8, 11, (G, A $\flat$, B) would become 7, 9, 11 (G, A, B) and the downward scale 0, 11, 8 (C, B A $\flat$) would become 0, 10, 8 (C, B $\flat$, A $\flat$). This convention, which we call the 'melodic convention', means that the major sixth may be considered as belonging to the minor key if it is preceded by the fifth, and succeeded by the major seventh.



An example is offered in the C# minor subject of Book II. The eighth note is the major sixth (10) of the true key, 11 minor. It is essential to the success of the algorithm on this subject that the non-accidental nature of that major sixth be appreciated, since there is another key which includes all of these first eight notes, namely that of the dominant 8 major.



Similarly, the minor seventh may be considered as belonging to the minor key when it is preceded by the tonic and succeeded by the minor sixth, when its normal interpretation would be as 'minor third of dominant'.



accidental to all the keys which have been maintained up to that point, 9 major, 6 minor, 1 minor and 4 major.



In this latter case there is little point in waiting for further information from the rest of the melody, since it soon wanders off into even more remote regions. Instead we deal with both of these kinds of ambiguity with a supplementary rule. This says that, in such dilemmas, if there is a hypothesis that interprets the first note of the melody as its 'tonic', then that hypothesis is to be chosen. Failing that, one which interprets the first note as the dominant is picked. This rule was called the 'tonic-dominant preference rule', because it turns out that we do not need to elaborate it any further: in the forty-eight, all the subjects that invoke it begin on one or the other. Later on some strong criticisms of this rule will be voiced, based on the fact that it involves an underhand appeal to rules of the nature of fugue, rather than a general rule of melodies and their key-identification.

A program embodying these rules alone assigns all forty-eight fugue subjects to the keys that Bach indicated in his key-signatures, in the way summarised in Table 3. (See Appendix III for details and worked examples of many of the subjects that have been mentioned). For the moment we note just two things about the operation of this

Fugue	Key found	Book I		Book II	
		Number of notes	Decision at note	Number of notes	Decision at note
1	0 major	16	16*	24	24*
2	0 minor	22	5	9	9*
3	1 major	22	16	4	4*
4	1 minor	5	4	22	12
5	2 major	15	15*	9	9*
6	2 minor	12	8	24	15
7	3 major	25	12*	21	21*
8	3 minor	13	13*	14	9
9	4 major	13	12	6	6*
10	4 minor	25	7*	46	18
11	5 major	19	6	23	19
12	5 minor	14	4*	27	7
13	6 major	18	8	29	13*
14	6 minor	18	5*	18	18*
15	7 major	31	15	43	16
16	7 minor	12	4	21	21*
17	8 major	7	7*	22	22*
18	8 minor	17	5	25	25*
19	9 major	8	8	20	20*
20	9 minor	32	5	11	6
21	10 major	38	14	24	9
22	10 minor	7	7*	30	6
23	11 major	14	11	12	12*
24	11 minor	21	7	25	6

Table 3.

* indicates that the tonic-dominant preference rule was invoked.
 This table differs from that in Longuet-Higgins & Steedman(1971)
 in that rests are taken into account in totals of notes.

part of the program, concerning key-identification. Firstly, the tonic-dominant preference rule operates in nearly half of them, that is in twenty-two cases, of which seventeen are brought about by the end of the tune being reached before the rule of congruence has reached a unique conclusion. (This is a consequence of the more general property of the algorithm which is the price we pay in avoiding the hasty conclusions of a clustering procedure: the algorithm tends to reach its decisions at a very late stage in the tune.) The frequency with which it is used suggests that it needs a rather more searching examination than it received when it was so airily introduced.

This suspicion hardens into alarm if we reflect that the rule implies that no fugue subject may both start on the dominant and not contain the subdominant to obviate the possibility that the key could be that of the dominant. If there is such a subject, then the rule will deliver the wrong answer. [Virtually the first Bach fugue that one thinks of outside the Forty-eight - the St. Anne organ fugue - has a subject of this kind, and the algorithm gets it wrong.



C	G	D	.	.
A $\flat$	E $\sharp$	B $\flat$	F	(C)

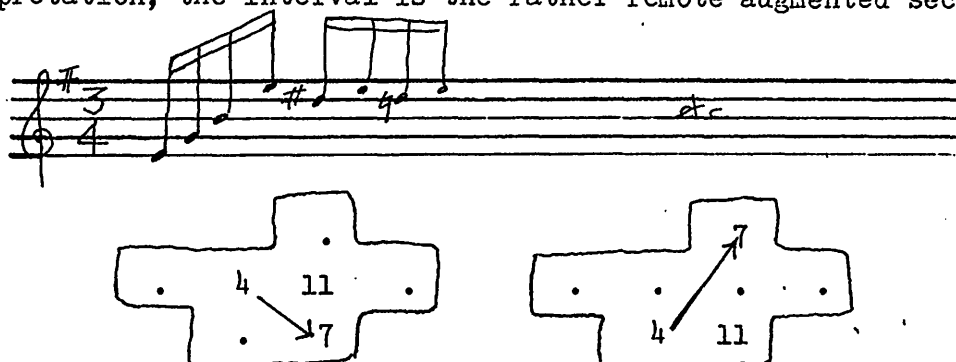
This is a particularly damning error, since the human interpreter that we are claiming to model would never make this particular error. Among other things, he would find it quite impossible to believe that the interval between the second and third notes was an imperfect fourth, which is the interpretation associated with the key of the dominant, B $\flat$.

In short, the hundred percent success of the algorithm in identifying key signatures for the Forty-eight is rather a fluke. One thing that is at fault is the tonic-dominant preference rule, which, as the St. Anne fugue makes clear, is not a statement about the nature of melody, but rather an inaccurate statement about the statistics of fugue subjects. The development of a subject in a fugue is to a great extent 'about' the relationship of tonic to dominant. The demands of the counterpoint make it extremely likely that the subject will begin on one or the other, and more likely the former, but it is not a requirement, even of fugue, and much less of melodies in general, or of our ability to interpret them. Even among the Forty-eight, one of the subjects begins on the supertonic; fortunately for the algorithm, it does not require any appeal to the preference rule.

Since it is used so very frequently in the algorithm's dealing with the subjects, clearly something must take its place.¹ An examination of the finer details of the program's analyses suggests that what is needed is a completely different kind of algorithm, rather than a similar rule to be used in conjunction

with this one.

For example, in the case of the E minor subject, No. 10 of Book I, besides the correct hypothesis, which involves the interpretation of the first interval as a minor third, there is an incorrect hypothesis, of G# minor, which is maintained, since both of the notes could be in the key. However in the associated interpretation, the interval is the rather remote augmented second.



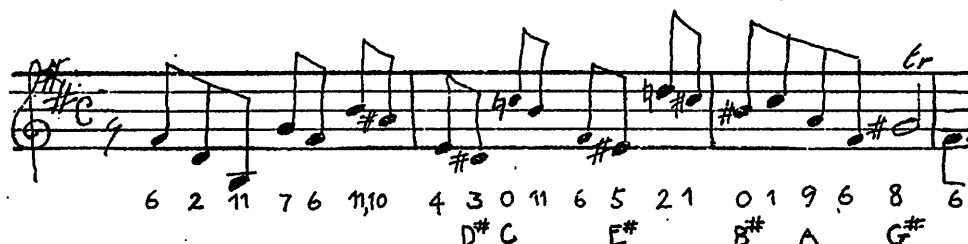
This is a very unlikely interpretation, in the absence of any preceding context to establish it, and particularly in view of the fact that it also demands that the next interval be interpreted as a diminished fourth, an even more remote version of the interval which under the correct hypothesis is interpreted as a completely straightforward major third. This extraordinary idea is maintained until the seventh note precipitates the infamous preference rule. It seems scarcely credible to think that people are so very cautious in their dealings with tunes that they entertain for a moment any hypotheses which interpret those four notes as anything but a minor triad, or to think that they are in any doubt as to the key after the first three notes. Perhaps in that case, rather than finding

a replacement for the preference rule in order to bolster up the existing algorithm, we should formulate a new kind of algorithm which will match our intuitions more. In particular, we might look for some kind of middle ground between the clustering algorithm, (which in its simple form was as over-hasty in rejecting as an interpretation exactly the same pons asinorum of the augmented second in Fugue 20 of Book II as this algorithm is over-cautious in entertaining), and this one.

But before turning to this, the algorithm of section § 4, we must pay some attention to the second part of the problem that was set at the beginning, that part dealing with the problem of notating a melody.

The notation, or interpretation in the full harmonic space, of the ambiguous intonation of equal temperament is trivial while the key is being decided, since by definition all notes up to that point must be non-accidental, and we have seen that for each key, including the one which will eventually prove correct, the notes are interpreted according to the frame. But, ~~once~~ the key has been established then accidentals may occur, and must be interpreted in their relation to the notes of the key, if the program is to come up to the standard of the music student in our paradigm of the dictation class. A striking example of the kind of problem facing us is in the subject of the great B minor fugue, No. 24 of Book I. The first part of the program has no difficulty in

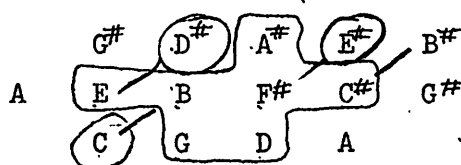
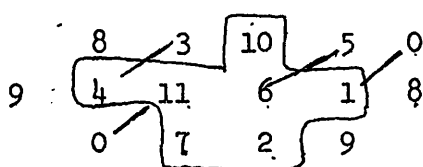
identifying the key correctly from the notes of the first bar, but in the next bar there are three notes - 3, 0 and 5 - which do not belong to the key, and which the composer has written as D#, C# and E# respectively. In the third bar there are three more extraneous notes - 0, 9 and 8 - of which the 0 has been written this time as B#. The question arises: how can the listener tell these six notes are to be interpreted as Bach wrote them, and in particular that the first 0 is a C# and the second a B#?



For the first four of the notes under consideration the solution is provided by another rule, which seems to describe Bach's use of chromatic scales, that is, note sequences of the form $(n, n+1, n+2, \dots, n+m)$ or $(n, n-1, n-2, \dots, n-m)$. We state this rule - the 'semitone' rule - as follows: in a chromatic scale the first two notes are always related by a diatonic demitone, and so are the last two. A 'diatonic semitone' is an interval such as that between the B and the C in C major - a move in harmonic space of one step to the left and one down, or its inverse.

There is a further discussion of the status of this rule in a later section § 3.41. There are three chromatic scales in the second bar of the subject above, namely (4,3), (0,11) and (6,5).

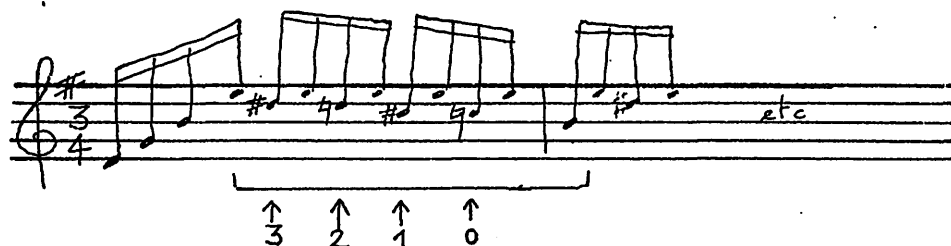
By the above rule, these are certainly diatonic semitones, and moreover the notes 4, 11 and 6 are all within the established key of 11 minor. Therefore the 3, the 0, and the 5 are to be interpreted as $D\sharp$, $C\flat$, and $E\sharp$ respectively.⁷ But in the third bar the 0 is the last note of the chromatic scale (2,1,0) and the first note of the scale (0,1). It must therefore be related to 1, which belongs to the key, by a diatonic semitone, and this demands the quite different interpretation as $B\sharp$, rather than $C\flat$



A problem arises with the first 9 in the third bar, which must be accidental to 11 minor, but is not part of a chromatic scale. We need one more rule - our last - for such an eventuality. We find it in the old idea of clustering: the accidental is placed as closely as possible to the preceding note-positions in the space. This is the rule also used in notating accidentals in the middle of chromatic scales.⁷ For lack of a better formulation, the clustering is taken to include all the preceding notes of the tune.

This is far too extended a context. Although it correctly interprets the 9 in question as the lower right-hand A, when we try to give the 8 in the same bar the same treatment, we find that there are two different readings of the note which are equally close to the already-occupied positions in the space. Both the right-hand or the left-hand $G\sharp$ in the above diagram give the same

measure. This is the only case among the forty-eight in which this clustering by the Manhattan distance measure fails to settle the interpretation of an accidental - though even here, it tells us how to write it down in the traditional notation (and, furthermore, would have settled the matter correctly if we had given it all the notes of the trill that Bach wrote on the note). However, to say this is really to quibble, since everybody knows that that note is not really ambiguous. By this time the piece has modulated solidly into the key of the dominant, and the program is simply an inadequate reflection of our interpretative performance when it includes the part of the melody that precedes the modulation in the context of the note. In another case where this rule applies its inadequacy is revealed in actual error, albeit one of a rather fine distinction. In the E minor fugue of Book I that we looked at before there is a long descending sequence which we realise is achromatic scale, although it is interleaved with repetitions of the top E, and the program cannot recognise it as such.



decided by the preference rule, it places this accidental by clustering with respect to the preceding six notes. The resulting interpretation is as the lower righthand D. However there are compelling musical reasons to believe that this is wrong, and it causes a further conclusion by the algorithm that certainly is. We know it is really part of a descending scale and therefore likely to be associated with the next note in that sequence, 1, as a diatonic semitone. This note has to be the major sixth C#, both because the key we are in is E, and because otherwise the adjacent E notes will be related to it by distant imperfect intervals. However, not only does the program take no account of these facts, but when it comes to interpret that 1, has so biased the cluster toward the right, by choosing the notation of the previous minor seventh, that the note is interpreted as the imperfect major sixth, notwithstanding the consequent wild imperfection ascribed to its relation with the surrounding E notes. In short, the real musical context in which this interpretation should be carried out is very much more local than that used by the clustering rule. When we next look at the notation problem in the context of another algorithm, we shall make a great deal of this failure. It should also be remembered that, because of the nature of the semitone rule, the same result would have occurred if the descending sequence had been a simple chromatic scale, without the alternating E notes. Thus both the notation of accidentals and of chromatic scales are less than perfect in these forms.

4. The Triangle Game

4.1 Introduction

4.11 Introductory Proposal

The kinds of mistake that the previous algorithm made suggest that a completely different kind of algorithm was required as its successor. Just as the 'Clustering' procedure was introduced as an introduction to that algorithm, a simple example of its type, so there is an extremely simple idea of the new kind, which will illustrate the nature of the second algorithm for key-identification. This idea attempts to capture the remarks made earlier about the 'obvious' interpretations of the first few intervals in the E minor, Fugue 10 of Book I, and the St. Anne fugue, although it does include an assumption that the melodies it deals with will begin either with tonic or with dominant. This is an even stronger assumption than that involved in the old tonic-dominant preference rule, and will be one of the first features of this simple algorithm to be discarded. The idea is put forward in Longuet-Higgins (1972) as an outline, without the rather odd extra rules that will be included here to make a complete algorithm, as a suggestion as to the form of such theories, rather than as a serious algorithm. It will be used for the same purpose here, though with certain liberties taken with the original form of the proposal, in order to make points clearer.

The intervals between the very first two notes of the E minor and the St. Anne fugues both have obvious interpretations relative

to the tonic, since they are in its chord, and it is the failure of the algorithm to embody this realisation that makes its analyses so unlike our own. The new proposal is to capture in rules the idea that the first interval can be associated with a particular interpretation, say as moving from dominant to mediant, as in the St. Anne, relative to an implied key centre. That key centre is to be taken as the key of the piece. Not all the subjects begin with the particularly clearly interpretable intervals within chords. Some begin with seconds and sevenths, but perhaps we can make some strong assumptions about the interpretations relative to a tonic that should be laid upon them, in some set of rules such as the following.

1. Two hypotheses are to be entertained:

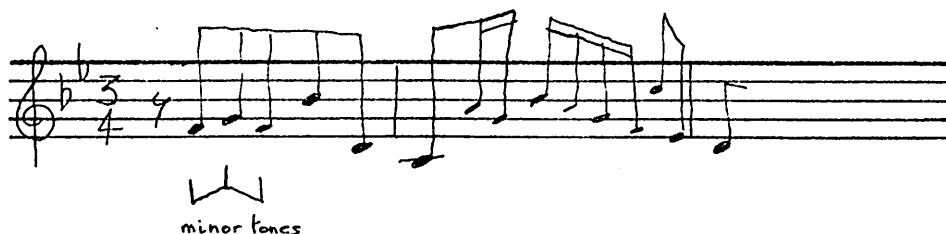
a) the first note is the tonic, and the key is that key, in either minor or major mode.

b) the first note is the dominant, and the key is that of the tonic a fourth below, again in either mode.

2. The first interval, made by the first two notes of the melody, or in certain cases another early interval, confirms one or other of these two hypotheses (the mode may be decided forthwith or later). It does this in the following way. (And here we go beyond the outlines of Longuet-Higgins (1972) for the sake of illustration.)

a) if the interval is a unison, then it is passed over, and this procedure is applied to the succeeding interval. If the interval is an ascending tone, and is followed by a descending

tone, then both are passed over and this procedure is applied to the next interval after the pair. This is necessary in Fugue 21, Book I.



- b) Otherwise the interval is interpreted as perfect, including the tone (interpreted as a major tone). (The minor tone is the only "imperfect" version of an interval which could have been entertained under the two hypotheses.)
- c) Given the two initial hypotheses, and the interval selected by (a) and (b) above, the rules for confirming a hypothesis are as follows. Each class of interval is identified by its size in keyboard semitones, ignoring octaves, so that, for example, a descending semitone is identified with the major seventh, of size 11.

size of interval in semitones	vector notation in harmonic space	implication as a chord in harmonic space	interpretation of first note	mode
0	.	-	-	-
1			dom	minor
2			ton	-
3			ton	minor
4			ton	major
5			dom	-
6		-	-	-
7			ton	-
8			dom	minor
9			dom	major
10			dom	-
11			ton	-

So, for example, in Fugue 17 of Book I, since the first interval is a ascending fifth, seven semitones in size, and as the table shows, a perfect fifth is always interpreted as $\overset{2}{\underset{\wedge}{\text{bing}}}$ from tonic to dominant, then the first note must be the keynote. As it happens, the next note is the major third of the scale, so the further question of whether the key is major or minor is settled then: the key is A $\flat$ major.



Less obviously, the first interval of Fugue 9 of the same book is a tone. The table shows that every ascending tone is to be interpreted as starting on the tonic when it is not followed by a descending tone, so that the key of this subject must be E. This is right, although because of the melodic convention, the decision as to the major or minor character has to wait until the second bar.



In general this model would work well on the subjects of the forty-eight and it gets them all right, except for two. It also does fairly well on subjects of Bach's organ fugues. But its errors are more revealing than its successes. We can disregard its error in believing that II, 5 starts on the dominant rather than the tonic - nothing short of a full-blown analysis of the subject including the important and still untouched concept of cadence is going to get this right, since it starts with the notes of the subdominant triad.

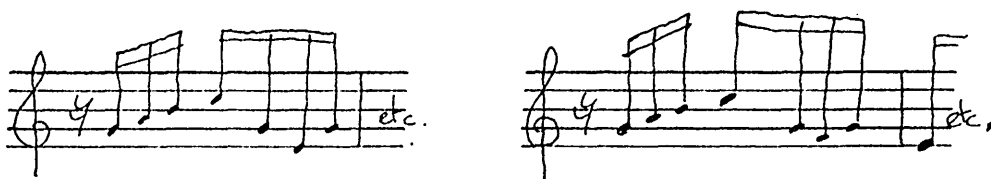


Other errors are more serious.

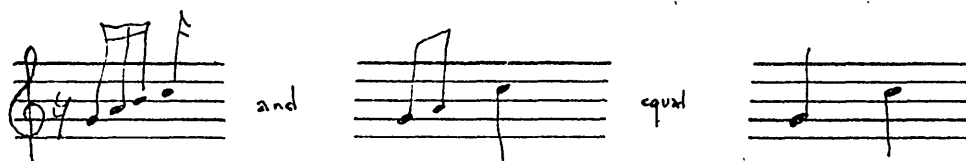
4.12 Three Interesting Errors Of The Introductory Proposals

4.121 The Initial Ascending Tone

In section 2,(b) of §4.11 it was stated that, in addition to the constraint ex hypothesis of section 1(a) and (b) that all first intervals other than the tone must be perfect, the tone must also be perfect (the major tone) rather than imperfect (the minor tone, say from dominant to submediant). Unless one takes the curious view that the rules one uses for fugue are radically different from those for other melodies, this is intuitively very odd. There are countless examples outside fugue where a melody begins on the dominant, goes to the submediant (a minor tone) and thence to the tonic. Such melodies are, for example, September in the Rain, She'll be Coming Round the Mountain, A Room with a View and Pop tunes too numerous and trivial to mention. Fortunately there are also some Bach fugues beginning on the dominant with an ascending minor tone:



In all of these examples there is no question of the melody being in the key of the dominant: in all cases their various opening sequences are equivalent to a movement from Dominant to tonic. In other words,



Conveniently enough, the fugues state this most clearly of all. There is no doubt in the hearer's mind that these sequences 'state' the triad of the true key note, C, and no amount of success on other material should persuade him that there is anything out of the way about these melodies.

Can we explain why this evidence should be taken seriously, and why the rarity of such sequences among fugue subjects should not determine the form of our model? A central idea in fugue is the reflection of tonic by dominant and dominant by tonic among the various voices. This is the principle which underlies the phenomenon of tonal answer, (Higgs). In the answer, (cf. Part I) to a given fugue subject, early and salient dominants in the subject are

answered by the tonic a fourth above (or a fifth below) while other notes are generally answered at a fifth above (or a fourth below). Such subjects as the above present a problem for the composer in writing his answer, for the answer should reflect the structure of the subject as closely as possible in a good fugue. A tonal answer to the above subject would destroy the structure of the run, since the initial tone would be replaced by a third. On the other hand a real answer would contradict the basis of fugue by not reflecting dominant by tonic. (In this case Bach chooses the latter alternative, presumably because the initial dominant is short and the run only a grace to the tonic.)

In this light we can see that an initial minor tone will be a rare occurrence in fugue, but that this is an artifact of the form, and may have nothing to do with the way in which we establish a sense of the key in a melody.

If we accept this view then we must accept the fact that we may have to look further into the tune than the first two notes to establish the key, and may have to entertain more than one hypothesis as to the harmonic function of the notes, (e.g. minor tone versus major tone) while we are doing it. This can hardly surprise us. Moreover if we are to capture that description of the above fugue subjects which we claimed was intuitively obvious, we shall have to talk about such constituents

of melodies as runs, and consider them as units.

4.122 The Tonic/Dominant Hypothesis

It is a consequence of such aspects of fugue as those discussed above, that fugue subjects will have a strong tendency to begin on tonic or dominant. However, just as in the case of the initial minor tone, a huge range of melodies from all periods suggest that this is not in general true of melodies, and makes our chosen subjects an unrepresentative sample. Such melodies as the carol In the Deep Midwinter, Whence is that Goodly Favour Flowing?, Jeannie with the Light Brown Hair, Chattanooga Choo Choo, etc. start on such degrees as the mediant and submediant and seem to present no complexity in their implication of key. For this reason, if we are going to continue to use fugue subjects (which we want to do for other reasons) then we must pay quite considerable attention to the few which do not start on tonic or dominant. Fortunately Bach provides us with a few. Among the forty-eight, Book II, 21 has the supertonic as its first note. Intuitively it is clear that the four notes of which this is the first constitute mere decoration to the tonic, and that what the first bar "means" is a tonic, major, triad arpeggio. It is clear from this that we are going to want to recognise such decorative sequences in a special way. The idea of "parsing" a melody into sequences that are treated as units is familiar from the earlier work on metre. We must break the tune up into such constituents in order to

understand the harmonic sequence of events.

Having reduced this subject's beginning to a special case of starting on the tonic, we have not advanced the cause of a truly general key identification algorithm, for this feature could easily be incorporated into the tonic/dominant algorithm, just as the repeated unison already is. The next example, a C minor organ fugue, is not so easily dealt with.



We will not attempt to deal with this fugue at this stage, merely point out that it begins on the minor submediant, that the first note is not in any simple sense part of a decoration or either dominant or tonic, and that the initial semitone will be completely misleading under the tonic dominant rules. We now proceed to some positive proposals.

4.2 Playing Games With Intervals

4.21 The Concept of Key

The idea behind the model which follows is that the concept of tonality of a sequence of notes is intimately involved with the concept of the triad of that tonality. That is to say that when a sequence of notes suggests a tonic to us, in terms of which its interval relationships are interpreted, the way it does this is by

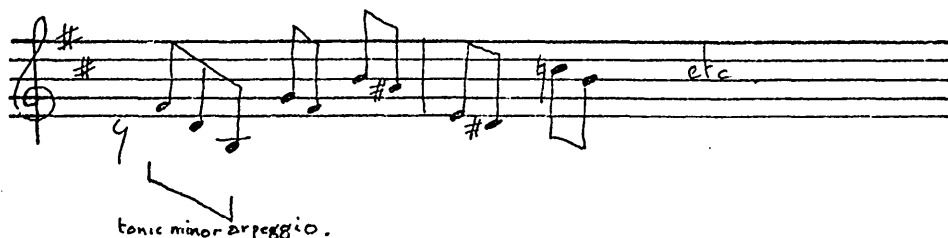
implying a triad. The most obvious case in point is when the sequence begins with an arpeggio and actually states a triad, as does the E minor fugue 10, Book I of the forty-eight. In this case we clearly want to say that the first three or four notes make us certain that we are in the minor key of the first note, for those notes are all in that triad and no other. By the third note we no longer want to en^etertain the possibility that we are in the closely related major of which the first note is the mediant, much less in the absurd key of G[#] minor which the old algorithm allowed. However, few melodies are as obliging as to start with an arpeggio, and many do not even start with one of the intervals within the triad. We shall see that we can write rules which interpret the actual intervals of the melody in terms of the intervals which make up triads, the 5th, the 3rd and the minor 3rd, and their inversions.

Of course most melodies pass through many "triad regions", or tonalities, whilst remaining in one key. It is an assumption of the following model that the first tonality established in the melody is that of the key signature. This is of course an oversimplification, again taking unfair advantage of the fact that we are dealing with fugue subjects, but it is very often true. We shall see that when it lets us down we can draw some inferences concerning the nature of a better model.

4.22 Implication And Confirmation Of Triads

4.221 Some Introductory Examples

To describe the assumptions and nature of an algorithm for key identification we can begin by looking at three fugue subjects - the B minor fugue 24 and the E major fugue 7, both of Book I, and the C major subject already quoted earlier in this section.



The B minor begins with an arpeggio, and this seems to establish the key. It must, because the accidentals soon begin and we must have a key framework within which to interpret them. (We know that the tonic dominant rule from the old algorithm is inadequate because of the 'St. Anne'.) We have already shown that the tonic/dominant hypothesis will not do, so the initial descending major third (from F# to D) cannot simply be interpreted as dominant to minor mediant of the keynote B: it could be interpreted as mediant to tonic of D major. On the other hand no other interpretations of this first event of the melody seem remotely likely. It was noted in the discussion of the previous algorithm's analysis of the E minor fugue 10, Book I that the interpretation of the, in that case, minor third as an augmented second seemed ridiculous. Here no such 'imperfect' interpretation is possible.) We express these statements in the following tentative rule:

The first interval is perfect (we will deal with the tone later). If it is a 5th, 3rd, minor 3rd (or their inversions) then it proposes for the first tonality of the melody (and hence the key) the two triads in which it occurs.

Although the first interval of Book I, 24 does not uniquely determine the key, the second interval seems to settle the matter. It confirms one of the two hypotheses, because it completes its triad, and it rejects the other. We express the idea that the second interval 'confirms' a hypothesis proposed by the first interval in the following equally tentative rule:

If the second interval is a perfect 5th, 3rd or minor 3rd, which lies in a triad proposed by the first interval then that proposal is confirmed.

It seems as if the E $\flat$ major fugue no. 7 of Book I of the forty-eight is a similar case, having an initial broken chord, (albeit a major one) which establishes the key beyond further doubt:



This is because notes 2, 3 and 4 seem to constitute a decoration of G, rather than having any particular force of their own. The "meaning" of the first five notes seems to be E $\flat$ - G - E, with the F in some way subordinated to the Gs which surround it. We capture this intuition by saying that the configuration:



behaves like a single note, (we are not at the moment concerned with the duration of this implied note) and we call it an inflection. It is a rather "well-behaved" constituent of melodies, and we found it necessary in talking about answers. Later in this section we shall consider some contextual restrictions on its definition.

The C major fugue which was introduced earlier in this section, to show that the first two notes of a Bach fugue subject might be a minor tone apart, suggests another constituent of melodies, another kind of sequence which must be considered as a whole. This time the constituent corresponds not to a single note, as did the inflection, but to a single transition between two notes, or constituents corresponding to single notes.

It was said in the introduction that we cannot make a hard and fast rule as to the major versus minor character of an initial tone. Moreover there is nothing in the first few ascending seconds to determine which we are dealing with in this case. It was suggested that the question was settled rather later, by the fact that the first bar seems to simply state the tonic triad. To capture this intuition we say that the initial sequence of rising seconds is treated in a block as a movement or transition from the first note of the sequence, C, to the last, G. This kind of transition is termed a run, for obvious reasons. It is treated as a block just like the simple first interval of the B minor: that is, it

proposes hypotheses for later confirmation. Not only is the idea of this kind of constituent familiar from the earlier work: this very category was required in the treatment of metre. Thus the run can be said to be well-behaved across these two domains. The first fugue of Book I shows that this account of runs is not yet complete. Book I, 1, also starts with a run of an ascending 4th, but from tonic to subdominant, not dominant to tonic.



With the simple definition of a run as a transition, to be treated just like a simple interval, the true hypothesis that the fugue begins on the tonic is never advanced, so we can never get the right answer by the process of confirmation outlined earlier. We could make a special rule that whatever happens the keys of the first note are always included in the set of hypotheses set up for decision. But this seems a bit arbitrary, and still leaves unexplained the events in the C minor organ fugue quoted before. This does not begin on tonic or dominant, neither does the first run, between $A\flat$ and $E\flat$, propose the triad of C minor. There does not seem to be anything wildly out of the ordinary about this subject: its settling down into C minor does not come as a great surprise, so we would like it to be subsumed under the rules we

need for the other fugues. Until we have amplified our rules for proposing triads on the basis of intervals to include the 'implications of seconds' we cannot define the rules for 'implications of runs'. We shall do this next, and for the moment merely say that the implications of runs are more complicated than those of simple intervals, and that, besides the implications of the transition between the end notes of the run, its internal structure may propose some further hypotheses, by rules of implication and confirmation similar to those we have already dealt with.

Some tentative rules have been phrased for key decision in terms of the successive transitions between notes in the melody. The first transition proposes hypotheses as to the key of the melody, subsequent transitions decide between these until only one is either confirmed or maintained. We have seen informally that transitions may be of two kinds: the run, an ascending or descending sequence of seconds, and the jump, an interval which is not a second involved in a run. We have also seen that the end points of transitions can be of more than one kind, can be more complicated than just a single note. The necessity of considering the inflections as such a point-configuration has already been suggested, and others will be introduced below. These definitions of constituents are logically separate from the interpretative rules of hypothesis proposal and confirmation. We might therefore identify them with a partial syntax of melody. In the following discussion both of the algorithm and the program which embodies it

we shall maintain this distinction of the syntactic and interpretative components.

This business of proposal and confirmation of hypotheses might be seen as a board game, in which the players are the hypotheses proposed in the first stage as to the key of the melody. Each plays on the 'board' of the harmonic space of Figure 4. They each have a number of 'sticks', corresponding to the diads of the thirds and the fifth which occur in triads. The aim is for a player to put sticks down corresponding to his key. The few syntactic rules described so far say how he may do this. We shall see more in the sections that follow. If a player is unable to put a stick down on one or another side of his triangle, then he is eliminated, under the principle of Congruence. Play continues until only one player is left: 'he' is the key of the piece.

Because of this metaphor, the algorithm being developed here is called "The Triangle Game". In the sections that follow it will be shown that the events of the melody may be such that several 'players' of this game can take them as support, and even that new players may be created to join the game in the middle.

In the interpretative domain we have only considered the properties for implication and confirmation of triads for the major and minor 3rds and the 5th transitions.

In fact transitions of all sizes are associated with triad implications, although since the rest of the intervals are not

directly present within triads, this association is less direct.

In the next two sections the definitions of these two components will be filled out. Since the syntax definitions of constituents of melodies arise out of the desire to simplify the interpretative component, we shall look at interpretation first.

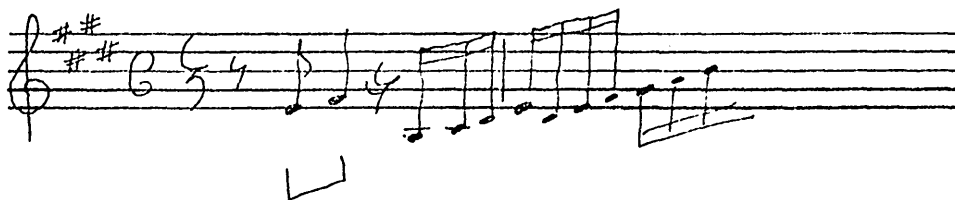
4.222 Interpretative Rules

4.2221 Interpretation Of Transitions As One Of The Chord Diads

The thirds and the perfect fifth, (and their inversions) imply the two triads apiece which they take part in. The remaining intervals are the minor and major tones, and the semitone, and their inversions the sevenths. (Imperfect intervals other than the minor tone have no implications. Hence if a hypothesis requires an interval to be imperfect then it will tend to be eliminated at the expense of any which do not.)

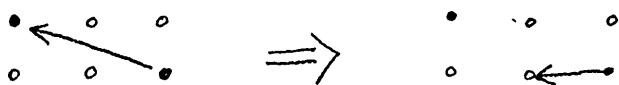
These seconds and sevenths are interpreted as implying thirds and fifths which in their turn imply the triads in which they occur. Unlike the intervals of the triads, the seconds are not all identical in their implications with their inversions the sevenths. The implications are based on the harmonic space.

The major second - two steps to the right in the harmonic space - implies the perfect fifth - one step to the right. Thus the initial tone of Book I, 9^f proposes a perfect fifth and hence the minor and major triads of the first note.



Its inversion the minor seventh - two steps to the left - implies the perfect fourth - one step to the left.

The minor second and the corresponding inversion, the "minor" seventh offer a slightly tricky problem. The ascending second -



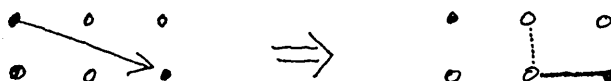
one step up and two to the left - clearly implies the fourth - one step to the left. So in fugue I, 21 the minor tone dominant to submediant sets up the true hypothesis that the melody begins on the dominant.



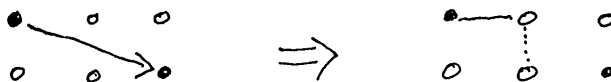
The inverse, the "minor" minor seventh is more complicated. Two steps right and one down, it can function as a grace note to the lower note, as in the first interval of 'Jeannie with the Light Brown Hair'.



In this case its implication is the fifth from the note one to the left of the lower notes:



implying in turn the above major triad. But because of the melodic convention in scales, we shall see when we come to deal with runs that the implication



must also be entertained in order that the internal structure of a run shall always suggest the key of its first end point.

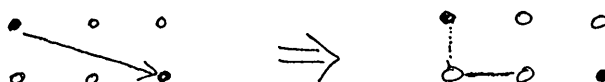
There is yet another interpretation that must be included for the minor seventh. Fugue 19 of Book II begins with several transitions between the tonic and the mediant of the true key, A major.



The first three transitions, therefore, leave the true key and its relative minor, F# minor, still 'in the game'. These two keys differ as to the interpretation of the "minor seventh" between C# and B.

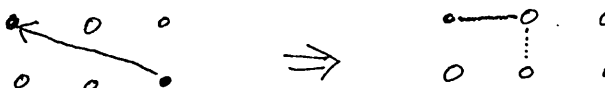
B	F#	C#	
	A	E	B

As was seen above, the 'major' version is interpreted as a fourth, so the F# hypothesis is maintained. In order that A major should live to fight another day (and win) the third implication of this interval



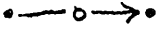
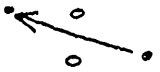
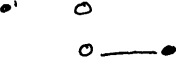
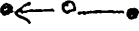
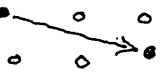
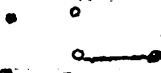
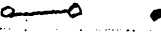
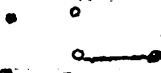
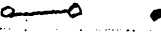
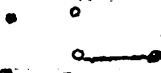
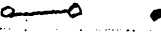
must be included.

As with the major tone and its inversion there is some temptation to restore the symmetry we see in the other intervals, and include all of these intervals among the implications of the minor tone (above). We do not do so because in Bach's music the implication



can never be true, not even by melodic convention.

The semitone is the most ambiguous interval of all. Both it

Size Of Trans.	Notation In space	Diad Interpretations In Space												
0	.	-												
1,11		<table border="1"> <tr> <td>a)</td><td></td><td></td></tr> <tr> <td>b)</td><td></td><td></td></tr> <tr> <td>c)</td><td></td><td></td></tr> <tr> <td>d)</td><td></td><td></td></tr> </table>	a)			b)			c)			d)		
a)														
b)														
c)														
d)														
2	a)  b) 	 												
3,9														
4,8														
5,7														
10	a)  b) 	 <table border="1"> <tr> <td>a)</td><td></td><td></td></tr> <tr> <td>b)</td><td></td><td></td></tr> <tr> <td>c)</td><td></td><td></td></tr> </table>	a)			b)			c)					
a)														
b)														
c)														
All Others		-												

Implications Of Transitions Through Various Intervals As
The Chord Diads, or Perfect 5th, 3rds, etc.

Table 4.

The true key of C# minor is supported in particular by an interpretation of the initial descending semitone as the neighbouring perfect fifth from C# to G#, because it is in the chord of C# minor. Having once made this interpretation, it is only natural to regard the next transition, to E, as being in the same harmony, and to disregard the fact that the first note of the transition is B#, since it has already been interpreted as a dominant G#. This idea will be elaborated later. For the moment it is just said that such an 'implied' end-point of a transition is called a virtual note, and that every transition is to be taken as beginning on the virtual end point of its predecessor, whether or not that is the real note with which it begins.

4.222 Further Interpretative Rule for "Run" Transitions

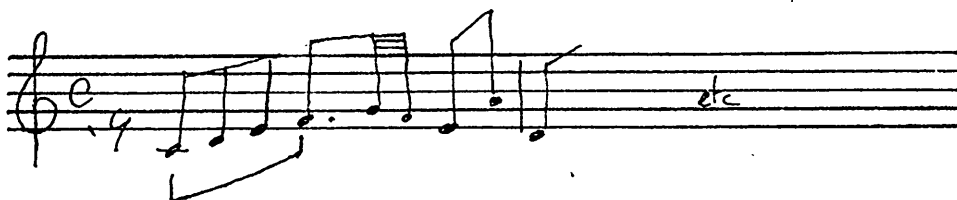
We are now, having set up the rules for interpreting seconds as implying triads, in a position to describe the rule for interpreting runs as internally implying triads over and above those already resulting from the status of a run as a transition. Although we shall not formally define runs until the later section on syntax, we can use a working definition as follows: [a run is a sequence of intervals such that each interval is a second, each interval is of the same duration, and each interval is in the same direction.]

The basic idea is that a run is illustrated just like a

separate melody within the main body of the tune. That is, the first second of the run proposes a number of hypotheses and the later ones may confirm one or more of them. However, the triads "confirmed" within the run are merely implications outside in the context of the whole melody. They are like any other implication in that they may propose hypotheses (if the run is the first transition of the melody) or confirm those already proposed (if the run occurs later).

However, since the seconds and sevenths are ambiguous, both as to their notation and as to their implications, the internal implications of scales are many and complicated. Accordingly, a simpler statement, which amounts to much the same thing, is made. It is that, in addition to the implications set up by the transition of a run between its two end points, the run carries the further implications of its first second or seventh, each such implication being subject to the usual requirement that the notes that make up the run are notationally consistent with it.

A full illustration of this rule of 'internal implication' of runs will have to await a fuller description of the algorithm which applies it. For the moment it is just noted that in the cases of the C major Fugue 1 of Book I,



and the C minor organ fugue (which was used to show that even Bach fugue subjects can begin on the submediant, and be left high and dry by the tonic dominant assumption),



the 'internal implication' from the first interval of the run in both cases is that of the fifth from tonic to dominant of the true key. In both cases this allows the true hypothesis to be at least advanced, so that, (as we shall eventually see) it may be in the end confirmed in both these cases.

As in the case of simple seconds and sevenths, we have to say what effect such an implicit interpretation of the transition has on the further interpretation of the piece, and in particular whether the next transition is to be considered in its own right, or as starting on the implicit end-point of the former. As in the previous case a full discussion is postponed, but there is a clear hint from the argument concerning seconds and sevenths that the successor will be considered as beginning on the implied, or virtual note.

4.3 "Syntactic" Rules

4.31 Introductory

In section § 2.21 we saw the lineaments of a syntax of melodies,

a set of constituents upon which the interpretative rules operate. We saw that there are two kinds of constituent we need for the task in hand: groups of notes equivalent to a single note, which we called points, (e.g. the inflection) and groups of notes corresponding to a movement between two points. We called these latter transitions, and the two kinds we talked about were runs and simple jumps. In this section this "syntax" is filled out further.

In the introduction we saw another example of a configuration of several notes acting as a single point, in II, 21. We called this configuration the turn:



and add this to the list of points. The following organ fugue, one of many similar, shows the necessity of yet another kind, also implicitly mentioned in the introductory section, when it was said that, if the first interval were a unison, further evidence had to be sought.



triad

Under the interpretative rules, in order to have any hypotheses at all we must wait until the descending min. 3rd, G to E, and pass over the successive unisons between the first four notes. The most

natural way to do this in the present framework is to make them a point. We do so, and call this type a repetition, though with some reservation as to whether a transition is never a unison. Any sequence of two or more notes of the same pitch (or at the octave) count as such a point, so the first bar of the above example is describable as a triad sequence of repetitions. The four types of point - the turn, the inflection, the repetition and the single note where it is not involved in one of the above - complete the class of points,

We have already informally identified the only two types of transition that concern us in the task of key identification. They are just the run and the jump. Having established the two classes of object and their members, and having shown in common-sense terms why we need them, we must refine their definitions.

4.32 Further Syntax

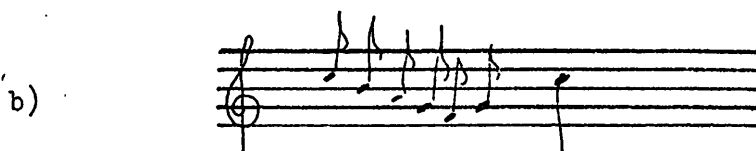
4.321 Methodological

The above rough definition of points and transitions overlap. For example, we have so far made so few restrictions on the definitions of runs and inflections that it is unclear whether

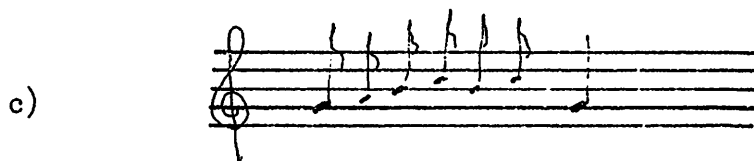
a)



should be parsed as "a run from a note on C down to a note of F, followed by a run from that F note to a C note", or as "a run from a C note to an inflection on G, etc.", or even as "a run from a C note to an F turn, " or as all of these. When we have decided that question, we shall similarly have to make the definitions clear in such cases as



and



It is now the time to do this.

We would like to make these decisions in the kind of way in which we set up these syntactic categories in the first place. That is to say in such a way as to make the rules for inferring key as simple as possible, and in fact confine them to the single one of triad completion. However, the evidence concerning these finer points is very much slighter, at least among the forty-eight fugue subjects, and is really inadequate to decide many of them. This does not matter very much, for the following reasons.

In the whole approach to the problem of modelling key-inference

presented here several simplifying assumptions have been made. The most gross of these was to assume that the metric structure of the melody could be ignored by the algorithms for this task. The assumption stands or falls on the performance of these algorithms, but whatever their success, it is nevertheless clear that a theory which makes no distinction between



and



lacks something in descriptive power. In the latter case it may well be that the initial notes stand as a unit for C, but in the former it is clearly not so. This evident incompleteness, even at this very low level of melodic analysis, coupled with the lack of any syntactic categories above this level of the minutiae of melody, means that the syntax has a rather odd status in the model. The best that can be expected is that we may lay hands on a few of the more well-behaved categories, with simple definitions, in the full expectation that these may be a little vague. We shall just have to put up with the vagueness, at least until new evidence is forthcoming. This new evidence is most likely to emerge from the operation of the full algorithm

for key-inference, which can be regarded more as a testbed for such subcomponents as this, than as being defined by them. (In the end, it turned out that the rest of the algorithm had sufficient power to get the right answers (for the forty-eight at least) even when the inflection category of point was missed out altogether, although at the cost of some speed and perspicuity in its analyses.)

4.322 Definitions

In keeping with this light-hearted approach the questions raised above are decided as follows.

All ambiguity in the definitions is to be removed: the syntax will always define a unique derivation for every tune. This is done by restricting the conditions under which the point configurations were to be parsed as points, in the following ways.

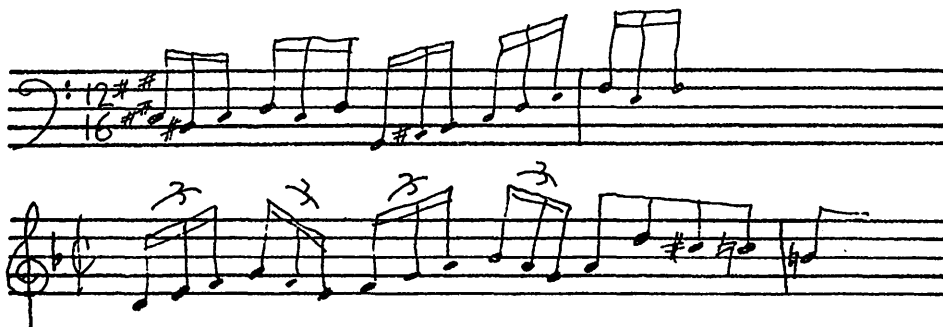
In all cases except one, (discussed below) the run definition takes precedence over that of the inflection: that is to say that in all cases such as the above examples a and b, where the inflection configuration



occurs in a context where one or more of the seconds that it comprises could be included in a run, they are to be included in the run, and the configuration in question is not an inflection. The exception to this general rule is for sequences of the following kind.



Where the inflection configuration is succeeded but not preceded by an ascending second of the same duration as ~~the second of the same duration~~ as the seconds of that configuration, (when it would come within the scope of the above rule) it is nevertheless to be parsed as an inflection. (This inflection will begin a transition of type jump if there is only one such succeeding second, but will begin a run if there are more than one such.) This exception to the general rule "if it could be a run then it is", is included for reasons which can only be fully explained in the context of the full algorithm ^{in its} application to our subject matter, the forty-eight. However some intuitive support can be drawn from examination of the following two fugue subjects, numbers 4 and 6 of Book II.



In no. 4 the first four notes constitute a case of the exception, which is made more plausible in that a description of the first

six notes as "a triplet standing for C# , followed by a triplet standing for D#, in a step of an ascending second" conforms with our intuitions as to the structure of the fragment. Similarly, the first bar of no. 6 affords two examples of the general case, with the sixth and twelfth notes of the sequence having the character of solid destination for the scale runs that precede them, and seeming not to be in any danger of being dominated by the surrounding notes. Unfortunately the effectiveness of this argument in favour of these rules is considerably weakened by the fact that our intuitions as to the structure of these melodies are probably strongly influenced by their metric structure. It has already been admitted that the omission of any metric considerations from these definitions is a serious lack, and a limit on the seriousness with which they can be regarded. They stand only as working definitions.

Similar restrictions must be made on the definitions of turns in the context of seconds of the same duration. In this case the need is to capture the idea that such sequences as that which surrounds the sixth note of fugue no. 6 above must not be parsed as turns. For simplicity's sake the turn is restricted to those contexts in which it is neither preceded nor succeeded by a run of seconds of the same duration and where none of the seconds which make it up could possibly be part of a run.

In a similar spirit of simplification the definition of the

inflection is restricted so as not to overlap that of the turn. That is to say that if the descending and ascending seconds of the inflection occur in a situation in which they could be part of a turn, then they are to be parsed as such.

Many of the above contextual restrictions on definitions of points and transitions are most naturally expressed (with an eye to a parsing algorithm) as precedences of one definition over another. In particular, the restriction of the inflection definition to contexts where it cannot be defined as part of a turn is easily expressed as a precedence of the latter definition over the former. Both they and the repetition take precedence in their definitions over that of the single note or interval. Similarly a rule which has so far been left unexplicit, namely that a second is to be parsed as a jump only in contexts where it cannot be included in a run, is most naturally expressed as a precedence of the definition of the run over that of the jump. In a parsing algorithm this idea of precedence ordering in the definitions will be expressed in an ordering of testing for the constituents of the structural description.

This completes the informal account of the syntactic definitions, and the next task is to restate them simply and formally, in the form in which an obvious parsing algorithm will need them. This rather tedious business, together with that obvious algorithm, is treated in Appendix IV, section § 2.

4.333 Parsing Algorithm

Once the definitions have been cast in this form it is straightforward to specify the parsing algorithm for the syntax. To parse a melody according to the grammar, is to repeatedly fill out the structure of fig. 12 by assigning values to the features of its nodes that were specified in section § 2.21, Appendix IV, according to the constraints set out in section § 2.22 of the same Appendix.

However, the terminals of the 'grammar', the intervals of various sizes, (unison, second, third, ...) are not the elements of the input string. An interval of a certain size such as a second, is an equivalence class over the musical intervals defined by the harmonic space. We can show the mapping from vectors in the harmonic space onto these equivalence classes quite simply by showing side by side a region of the space around a central origin, and the same region with each point re-labelled with the interval equivalence class of the vector that it defines relative to that origin. (In the following figure, on the next page, interval classes are represented with the integers 0 to 6, where 0 is "unison", 1 is "second" and so on.)

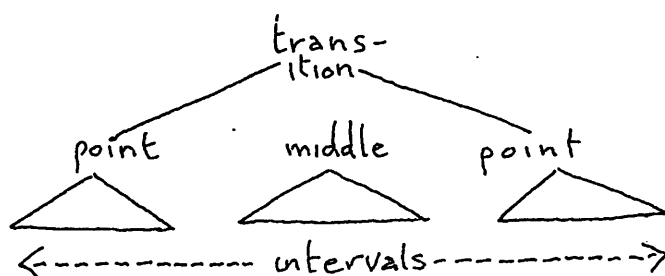


Fig.12

1	5	2	6	3	0	4
6	3	0	4	1	5	2
4	1	5	2	6	3	0
2	6	3	0	4	1	5
0	4	1	5	2	6	3
5	2	6	3	0	4	1

We can clearly define this mapping from vectors in the space onto interval names as a simple rule.

If the input to the algorithm were a melody "played" in just intonation, then this rule would be the only additional component that the parsing algorithm would need in order to parse the string.

However, the task in hand is slightly more complicated than this.

The task is to produce an algorithm, of which this parser is a part, that will decide key from an input "played" in the ambiguous equal temperament, where each interval of the string maps onto several vectors in the harmonic space, and hence onto several equivalence classes. (For example, the equally-tempered interval from keyboard note C, (C, B $\sharp$, D $\flat$, etc.) maps (among more remote alternatives) both onto the class of seconds, as the augmented second C to D $\sharp$, and onto the class of thirds, as the minor third C to E $\flat$). It is at this point that we become aware of the decomposition of the problem that was hinted at earlier. The problem of key decision, and that of parsing are logically quite separate from the further

problems raised by the ambiguity of equal temperament. In this section and previous ones we have discussed parsing and key decision independently of the problems that equal temperament introduces - such problems as that of identifying an interval as a second or a third, or a perfect fifth, which carries implications of tonality, from the imperfect, which does not. In the next section, concerning notation, we shall see that we can look at the problems that the ambiguity of equal temperament introduces, quite independent of the problem of deciding a key, and of such things as parsing.

4.4 Notation

4.41 Notation Of Equal-Tempered Melody As a Separable Problem

All of the preceding discussion of interpretative rules and parsing only makes sense when applied to melodies input in just intonation, whose notes are identified unambiguously as to pitch, and therefore as to position in the harmonic space, so that the class of the intervals (second etc.) can be identified. It was assumed that the additional problem that we have set ourselves can be identified in trying to model key identification in the ambiguous equal temperament could be assumed to be the separable one, that is the problem of getting from equal temperament to just intonation. This is the problem of 'notating' the melody, where, as usual here, the term is used in its fullest sense of

complete harmonic identification.

When the concept of key was being refined, in the introduction to this part, it was said that a key defines a set of expected pitches, and the harmonic functions of those pitches as defined by the harmonic space. We saw there that another way of putting it was to say that a key defined the way in which an interval in ambiguous equal temperament would be interpreted, or disambiguated. The example given was that of the interpretation of the tone from C to D in the light of the two keys of C major and A minor. Each key meant a different interpretation of the ambiguous interval in the harmonic space, 'and in just intonation.' This means, conversely, that the idea of notation of a melody is inextricably bound up with the idea of its key, to such an extent that we can only talk about rules for notating a melody according to a given key. (This complete identification of the key with the notation was involved in the first key identification algorithm, in that it associated a separate notation with each hypothesis it held as to the key of the piece. However, as we shall see in a moment, the identification was not complete, and in particular, the old semitone rule, (which defined the notation of semitone sequences independently of any key hypothesis was inconsistent in this respect.) In the following discussion we shall look at the problem of notation according to a given key, quite independently of our larger purpose of déciding that key. Afterwards we shall return to that problem, and resolve

the difficulty of how we can decide a key on the basis of a parse and a notation which we can only make in the light of already (in some sense) knowing the key. (Clearly we shall get around this problem as we did in the first program, by having several "hypotheses", and carrying a separate notation along with each of them, but that is getting ahead of the story.)

4.42 Notation In A Key

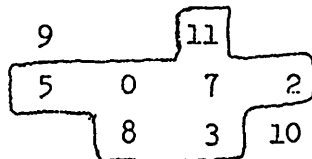
4.421 Key Frames Again.

To identify the idea of notation, or disambiguation of equal temperament, with that of key, the latter idea must be clear. In the introduction, a key was defined as a set of notes, specified as to relative position in the harmonic space. The idea of this "frame" of notes in the scale is already familiar from the discussion of the earlier algorithm. If we are thinking of tunes in just notation, then the frame is a set of expectations as to the notes, and the harmonic relations between them, that may occur in the melody. In the realm of equal temperament, the frame defines the interpretations that will be chosen to resolve the relational ambiguity of those notes. The first and simplest rule for achieving the notation of a melody in a given key embodies this idea of key. It is that if for some note of the melody one of the alternatives onto which it maps in the harmonic space is within the

frame of seven scale degrees, then that is the notation adopted.

For example, when in the key of C minor, it is not necessary to look any further for an interpretation of the Equally Tempered note-value 3 than to the E $\flat$ which is in the associated key frame, for this is what it means to be in this key.

In addition, the melodic convention applies here as it did in the old algorithm: in the appropriate contexts of ascending and descending minor scales, the major sixth and minor seventh are included in the frame, and therefore in the above rule, as the mediant of the subdominant, and the minor mediant of the dominant respectively.



4.422 The New Semitone Rule

4.4221 The Old Rule

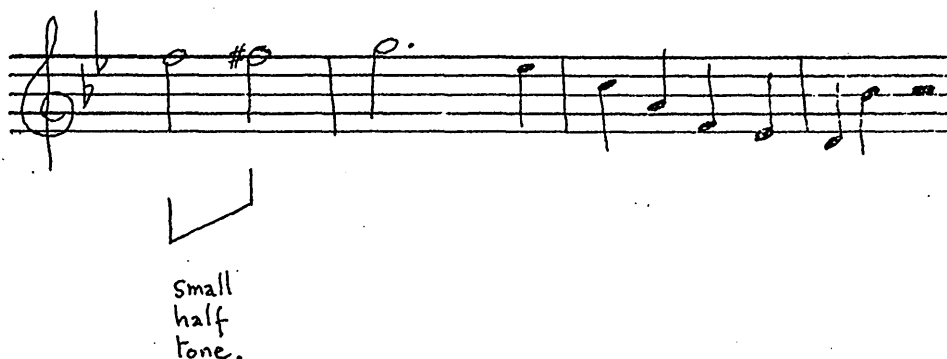
In the first algorithm for the identification of key, accidentals were treated in two ways. Those which occurred in the context of chromatic scales, ascending or descending sequences of semitones, were treated separately from ordinary accidentals, and their notation was achieved according to a special rule, the

semitone rule. This rule said that "in any sequence of notes separated by semitones, the first two are always separated by a diatonic semitone, and so are the last two". As has been hinted already, this rule has two serious inadequacies, which will be revealed by a good hard look at it in the light of our new decomposition of the problem into two domains of key decision and notation in that key.

Firstly, we can now see that this form of the semitone rule obscures the distinction drawn above, as between the business of key decision on the one hand, and notation on the other, and which is otherwise preserved for the most part in the early algorithm.

Secondly, it is incomplete, in that it simply will not deal with all possible chromatic scales between all degrees of the scale. Some of these, for example the descending ones onto the minor submediant, do not end with a diatonic semitone, and hence would be wrongly notated under this rule. Admittedly, it seems to be a feature of Bach's style that he never in fact does have such a sequence. That is one reason why the rule works, despite the confusion of its status, remarked upon above. However, this is not a sufficient excuse to maintain this formulation. We know that this rule is not a feature of composers' style in general, even ones as close to Bach's own period as Beethoven, as can be seen from one of his fugues for Piano and String Trio, where the first three notes, played in various

octaves by the trio, are a chromatic sequence beginning with a chromatic semitone.



It is as unsatisfactory to appeal to features of Bach's style at this very trivial level, before deciding the interpretation, as it was to appeal to the features of fugue that they start on either tonic or dominant) in the old ^{sic}tonic-dominant rule. It is one thing to say that Bach could never have composed the above - it is quite another to say that he could not have written it down.

If we take the two complaints together then it is clear that this formulation must be abandoned. If we are to accept the first objection, then we have to make a semitone rule, where the rule for notating the semitones is separated from the criterion for deciding key on the basis of that notation. But if the notation must be

done before deciding whether or not the key is a reasonable one, then the rule may have to notate sequences according to some very unreasonable keys. In other words, it must be complete and cover any sequences of semitones for a given key. Some of these may be unreasonable enough as to require non-diatonic semitones at the end of chromatic sequences. (An example from the forty-eight where an intermediate hypothesis, later rejected, requires this, will be given in due course.) Although it has already been said that Bach's style ensures that such keys will immediately afterwards be rejected, if the above objections to the old semitone rule are valid, then it is necessary to have a much more powerful and general semitone-notating rule, and consider stylistics later when a full description of the melody has been formed. Fortunately, it is possible to formulate such a rule.

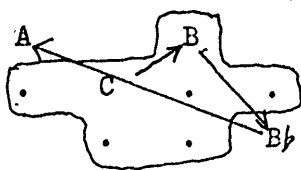
4.4222 The New Rule

The way this rule was derived was by separating two components of the old rule. On the one hand there is a component which reflects the fact that there are constraints in interpreting semitone sequences, and in particular that certain semitones are interpreted, with respect to a given key, as diatonic. On the other hand, there is a stylistic component, which says something like "If a melody has been interpreted as having a chromatic sequence which begins or ends with a chromatic semitone, then it probably wasn't written by Bach." We keep the former part, but

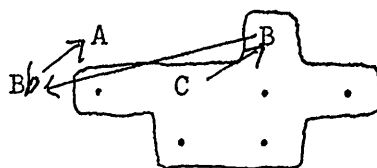
leave out the latter, on the grounds that when people deal with melodies, they do things like deciding key and harmonic interpretation first, then decide on this basis who the composer might be, rather than needing to know the composer in order to be able to interpret these simple aspects of the piece.

This former part is reflected in a much stronger statement about chromatic scales than the old semitone rule. Every note in a chromatic scale except the first and last is involved in two semitones, one on each side of it. The new idea is that every note in such a scale, except the first and last, is involved in at least one diatonic semitone.

The reasoning behind this suggestion is twofold. First of all, all paths in the space made by two semitones which do not involve at least one diatonic semitone are very long and rambling. Secondly, because of this they imply that the harmony associated with the interpretation contains some wild leaps: in the following example of a descending chromatic scale from tonic through major sixth in C minor the interpretation which involves the B $\flat$ in two chromatic semitones implies an implausible and unsupported momentary dominant harmony between the tonic and subdominant harmonies which really are implied. The interpretation of the B $\flat$ as related to the major sixth by a diatonic semit^{one}, subsuming it within the single harmony of the subdominant, is the right one.



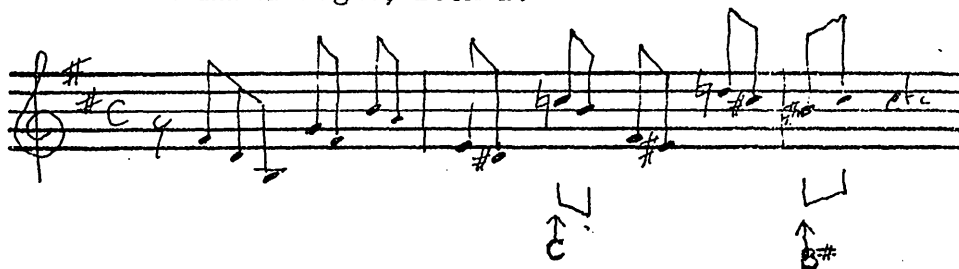
wrong.



right.

Thus the problem in notating a chromatic sequence is to decide which of each pair of semitones if any is chromatic. This depends on the key. In the previous example, if the key had been G minor, the 'South-Easterly' B $\flat$ would have been right, and the 'Westerly' one wrong. What is the nature of a rule to notate chromatic scales in a given key?

The old semitone rule was designed to cater for such notational contingencies as shown in the two semitone pairs marked below in the subject of the B minor fugue, Book I.



The keyboard note-value 0 appears in both pairs, but in one case it is to be interpreted as C $\flat$ and in the other B $\sharp$. The old algorithm was designed to notate these two differently with respect to the single key of B minor. It achieved this by invoking the semitone rule to say that each pair had to be a diatonic semitone. While there is complete truth in the statement that those semitones must be diatonic, there is also a complete nonsense about notating

the latter pair with respect to the same key-centre as the former: a modulation to the dominant has already been signalled, by its leading note, so the 0 is to be interpreted relative to that, as the leading-note B $\sharp$ of its dominant C $\sharp$, rather than in the key of B minor.

The solution adopted in this algorithm is to make the rule for notating semitone sequences very simple indeed - so simple in fact that we have to pay the price of not being able to notate that last passage of the B minor using it alone, but must rather have a separate rule for detecting modulation, and notate such passages with respect to the new tonal centre, rather than to a single key signature for the whole piece.

We do this by simply prescribing the notation of any note in a chromatic scale for a given key with a frame, like a key frame, of all twelve notes.

	9	4	11	6
10	5	0	7	2
	1	8	3	

For a given key, the notes of the scale are notated in this way. Rules for detecting modulations (or for defining 'well-formed' Bach) should be written to apply to the result of this notation. Thus, in the B minor a modulation might be detected by the occurrence of the leading note of dominant, causing a shift of the tonal centre to the dominant, and allowing the correct notation to be derived from the frame, relative to that new tonal centre. It

is only fair to state immediately that such rules have not been written, so that this algorithm, for melodies that modulate at least, will not carry out the notation part of the problem as originally framed. All that can be said is that perhaps it constitutes more sensible framing of the question, even if it does not give the answer.

It should also be noted in this connection that the task that the algorithm is performing is rather different in principle from that attempted by the old algorithm. The program no longer detects the key of the piece, nor notates the piece in that key. Rather, it detects a harmony associated with a sequence of notes, a much more local tonal centre which might be termed a tonality rather than a key, and notates the sequence with respect to the local tonality, rather than with respect to a global key. The algorithm only functions as a key-identifier by grace of a logically quite separate assumption that the first tonality is the same as the key of the piece.

If it seems that a great deal of power has been sacrificed, it should be reflected that the old algorithm's power in this respect was to some extent illusory. It did make errors in the notation of semitone sequences, and was not, in such analyses as that of the B minor of Book I, reflecting the kinds of facts about the melody that it should have done.

4.423 Accidentals

In the first algorithm for key identification, accidentals not included in semitone sequences were notated by choosing the interpretation which had the least total of its Manhattan distances in the harmonic space to all the previously encountered notes of the tune. It was clear all along that, whatever the extent of the context that should be taken into account, it was more restricted than this rule would imply. In one case, where the extent of previous melody was large, the algorithm was led into a situation in which two interpretations were equal-valued simply because it took into account notes from a previous modulation, which were in no musical sense the context of the note in question. Much more seriously, any errors in notation induced by this algorithm will tend to be cumulative. The old algorithm produced a particularly sad case of this in its analysis of the E minor fugue of Book II, as was described at the very end of section § 3.2. In that case, a trivial error caused by the old rule, that of notating the minor seventh as minor third of dominant, instead of the correct subdominant of subdominant, so shifted the centre of gravity of the tune to date to the East of the harmonic space that the next note but one, the major sixth, was notated as the extremely remote easterly imperfect sixth. Accordingly, in this new algorithm it was decided to avoid this danger of cumulative errors by simply placing accidentals as closely as possible to the previous note. We may expect that in the case of the rather

ambiguous tone this may lead to errors, since it will always notate an accidental which is a tone from the preceding note as being a major tone from it. (This is correct in the case of the E minor fugue.)

This rule is a little odd. Intuition would suggest that accidentals should be notated with respect to their successor(s), since usually an accidental signals a change of tonality, in relation to which it is to be understood. But at present we have no rules to express this, and so make do with this admittedly unsatisfactory rule.

If there is an error arising from this rule, it will only lead to further errors when there are long sequences of accidentals. Once again, this will only be the case when a modulation has taken place, and, as has been argued above, there is no reason why this algorithm should work when it is mistaken as to the true tonal centre.

There is one interval for which this rule for notating accidentals does not yield a unique answer. This is the interval made up of six keyboard semitones, and it has two interpretations with the same minimum Manhattan distance in the space, those of the augmented fourth and the diminished fifth. If there is such an interval, and its second note is an accidental, then clearly the algorithm must take some further context into account. The question is, what further context? Since we saw in the case of the E minor that notation really depends on such statements of structural description as "this note is in the middle of a broken chromatic

scale", rather than such an arbitrary context as ^{that involved in this rule,} it might be said that there is no useful answer to the last question. In the actual implementation the algorithm is simply arranged to draw the user's attention to the fact that it cannot decide the notation. In the event it turns out that the question does not arise: none of the fugue subjects used contains an augmented fourth ending on an accidental, though there is no particular reason why tunes in general shouldn't.

4.43 Notation Algorithm

Because the new semitone rule notates all semitones by rote, without reference to any context, the sub-algorithm for notating intervals in general is extremely simple. It simply checks first to see whether the interval is a semitone or not, and if it is, notates according to the semitone-rule frame. If not, then if the note is non-accidental, including being in the local context of melodic convention, then it is notated according to that frame, otherwise the interpretation closest to the predecessor is chosen. Since the notation proceeds by interval, and each note takes part in two intervals, at each stage (except for the first interval) the first of the two notes has already been decided. (See Appendix IV, Section § 3 for details.)

4.5 Interpretation And Identifying Key

4.51 Introductory

4.511 Recapitulation Of Parsing And Notation

We have just seen in Sections § 3 and § 4.4, that for some given key, a unique notation, that is a disambiguation in full of harmonic function in the context of that key, can be assigned to each of any string of notes, input in equal temperament. When this notation is known, a syntactic description of that string for that key can be derived. This description consists of a segmentation of the tune into its constituent transitions. Each transition defines a vector in the harmonic space (among other things).

4.512 Recapitulation Of Interpretation

In section § 4.222 we saw how to get from the vector in harmonic space defined by a transition to one of the three diads, the major third, minor third and perfect fifth, plus their inversions, that occur in the major triads, (see section § 4.2221). We also saw, in section § 4.221 how to get from these diads to implied triads, which were identified with the idea of a tonality, by the single obvious inference rule that a diad carries the implication of the two triads in which it occurs.⁷ In the immediately succeeding section, § 4.52 we shall look again at these two kinds of interpretive rules, define them better, and look at an algorithm for applying them to melodies.

These rules of notation, parsing, and interpretation all apply

to a melody whose key is known (or assumed). They only tell us what the melody would convey if it were in that key. There are quite separate rules which, given such an interpretation of the melody, concern the degree to which this message makes "good sense". These rules, given several interpretations (corresponding to several different keys) will say which of these make most "sense". This is clearly going to be at the heart of the algorithm for key-identification.

4.52 Interpreting A Parse In Some Given Key

4.521 Implied Intervals And "Virtual Notes"

This and the following section, § 4.522 are an extension of section § 4.2221 concerning the interpretation of transitions as the diads that make up the major and minor triads.

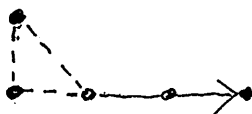
In that earlier section the idea was advanced that transitions of seconds and sevenths, through intervals other than the fifth, the thirds, and their inversions, were to be interpreted as implying transitions through those intervals. In some case, notably that of the semitone, the original transition would imply several such diads. In these sections the concern will be to fill out these two aspects of the implication on diads by transitions - in defining what is meant by implication, and how often multiple ambiguity of this implication can be handled.

To say that a transition, which defines a vector in harmonic

space, is to be taken as implying another vector in that space is to say something about the possible harmonies of which that original interval is a realisation. Thus if we say that the perfect fifth implies the major and minor triads of its lower note, then we are talking about the harmonies of which the fifth is a realisation.

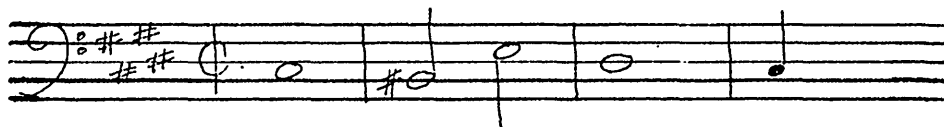


If we want a melody to "state" some other harmony then we had better not use that interval. When it is said that the major tone implies the perfect fifth, which in its turn has the implications that we have just seen, a similar thing is meant: if you want a subdominant harmony, say, then don't use the major tone from tonic to supertonic, because that isn't what it means.



To say that a major tone from the supertonic implies a perfect fifth from tonic to dominant raises a further problem. Where is the succeeding transition to be considered as starting from? It was shown in section 4.221 that to be consistent with the idea that the earlier transition stood for a fifth to the

dominant, we should consider the next transition as beginning on the dominant. The example given there was that of Fugue 4 of Book I,



where the 'reading' of the first semitone as a fifth to a 'virtual' G# , and of the subsequent diminished fourth as a "minor sixth" from the virtual note to E, leads to an interpretation of the first three notes as 'meaning' the tonic triad, and hence correctly establishes the key.

Such an implied note as this dominant, which does not appear in the melody, but only exists as a consequence of interpreting the major tone as a fifth, was called a virtual note. Since a virtual note derived from some such interpretive process must form the starting point for the next transition, the algorithm must keep a record of such a virtual end point in order to correctly analyse the next transition. It goes without saying that this extra storage requirement will, like the notation and the description, be associated with the particular hypothesis under which this interpretation is made.

In the table of implications of Section § 4.22 the imperfect intervals have undefined implications. They correspondingly

give rise to undefined virtual notes. These are dealt with in a special way, as described in § 4.6 below.

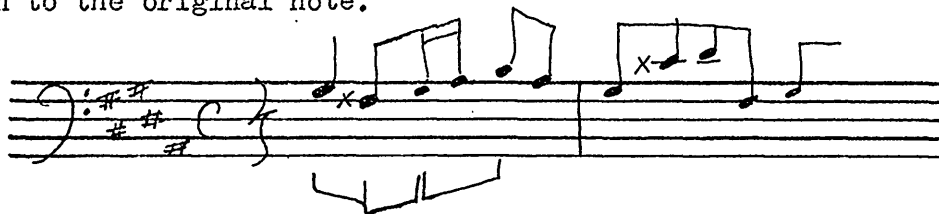
It was said in the earlier section § 4.222 that the same thing applies to the 'internal implications' of runs: since a virtual note generated by such an implication is really an identification of the harmony which the hypothesis implies, it is natural to take the virtual note as the beginning of the next transition for that hypothesis, disregarding the 'real' end point of the run.

It can happen that the 'virtual interval' set up in this way is itself a second etc. In this case the interpretation of the virtual interval is exactly as if it were a real one.

In all cases, if the new transition is a run, the internal implication is taken on the basis of the first of its seconds etc, as it really is, regardless of virtual beginnings, on grounds of simplification.

One more observation completes the details of virtual notes.

It may happen, as in Fugue¹⁸ of Book I, that a second etc., implying a virtual note, is immediately followed by its inverse, in a return to the original note.

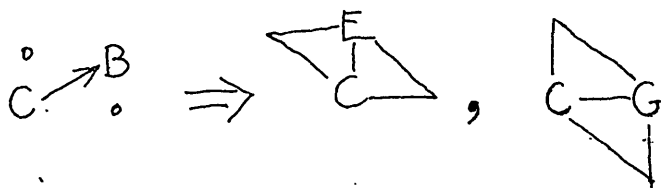


There is an additional constraint that the interpretation of the inverse interval must be the same as the former. The point of

the restriction is to prevent wholesale proliferation of hypotheses from these ambiguous intervals, in a manner discussed in the next section. It also applies when the interval which is the inverse of its predecessor begins a run.

4.522 The Ambiguity Of Implications Of Seconds And Sevenths

On the above view of what virtual notes really mean in the theory, it should be clear that when it is said that the vector corresponding to a transition through a second or seventh can stand for several diad vectors, this is simply to say that it can be the realisation of several harmonies. For example, among the alternatives that the diatonic semitone of a transition from C to B can stand for are the following:



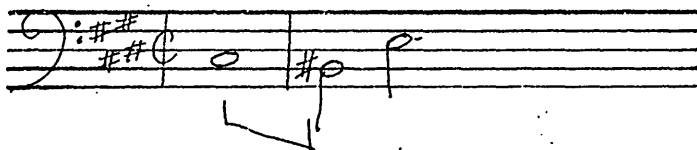
It can be interpreted as the major third from C to (virtual) E, or as the perfect fifth from C to (virtual) G. In the first case the associated triad harmonies are C major and A minor, in the second case they are C major and C minor.

Each of these corresponds to a different description of the melody. When we are forming a description of a melody according to a key the ambiguity of the interpretation of the transition must be reflected in an ambiguity in description: there will be

as many descriptions as there are diads that the transition may stand for, and when we come to such an interval in the formation of a description, a copy of the description must be created for each such alternative interpretation. (See Appendix IV Section § 4.)

In terms of the metaphor of the "Triangle Game" it can now be seen that the "players" are interpretations of the melody, and that an ambiguous second or seventh causes new 'players' to enter the game, each a copy of some earlier one.

For example in Fugue 4 of Book I,



for the true key-hypothesis, C# minor, as for many others, the initial descending semitone is diatonic, and therefore has four interpretations. A copy of the original hypothesis is made for each, and associated with each interpretation of the interval in that key. Of the four, one interpretation actually supports the key hypothesis, and is preserved while the others are discarded. As we have seen, the 'virtual' interpretation of the next interval is unambiguous support for this hypothesis and no others, and ^{this leaves C# as the only survivor} since ~~they represent the same key,~~ the problem is taken as solved.

4.6 Determining Key By Comparing Descriptions In Different Keys - The Principle Of Congruence Again.

(For all worked examples below see Appendix IV Section 5 for corresponding program output and detailed analyses.) In previous sections a certain type of algorithm will have become familiar. It runs as follows. "Examine the melody from left to right under all going hypotheses in parallel. If at some stage in the left-right progress one hypothesis is contradicted under some criterion, while another is maintained, then the former is rejected and is not considered further."

This kind of algorithm has been typical of all stages of the work, and was enshrined in the even more general "Principle of Congruence", which states that a non-congruence, be it of rhythm or harmony, may not occur in the melody before a point at which it can be unambiguously seen to be a non congruence. The algorithm presently under discussion is of this type. Like the earlier algorithm, at each stage of the key identification there are a number of hypotheses as to the key. The tune is examined under all of these in parallel, and at each stage in the pass from left to right the attempt is to eliminate some of these according to the Principle of Congruence. The process as usual terminates when only one hypothesis survives.

In such algorithms there are two important parameters. Firstly, how are the hypotheses first proposed? Secondly, what is the criterion under which congruence is assessed, and hypotheses rejected or maintained? In earlier algorithms, both for identification of key and metre, there has been ^aconsiderable variety of ways of

handling these questions. In the metrical algorithms the Principle applied in its strongest form: only one hypothesis was ever entertained at one time, erected and elaborated upon the earliest pieces of evidence that the algorithm came across. In the early key-identifier, the set of hypotheses was originally set up as the set of all possible keys, and the criterion for the application of the Principle was the weak one that the occurrence of an accidental under some hypothesis was to be taken as a non-congruence, and the hypothesis eliminated. (We have seen that this is too weak a criterion. The point of the current algorithm is to establish a better sense of congruence.)

The algorithm under discussion here is like the earlier algorithm in reaching its conclusion by the gradual elimination of a set of initial hypotheses. As in the earlier version, these are the set of all possible keys, (more exactly all those keys to which the first note of the melody is not accidental.) However, its criterion of congruence is very much more powerful, and in particular, means that nearly all of these are immediately eliminated upon the examination of the first transition.

This criterion is twofold. Firstly, exactly as in the old algorithm, accidentals constitute non-congruence, and hypotheses which involved early accidentals in their description of the melody are rejected in favour of those which do not. (This time, if all hypotheses involve accidentals, as happens in fugue 12 of Book I,

then all are maintained under this part of the criterion.) But the strength of the new algorithm lies in a further assumption of congruence. We have talked earlier about the harmony or triad that a transition implies. The further definition of congruence is that a harmony is congruent to a key if it is that of the triad of the key concerned. This is the assumption introduced in section § 4.21. Under the Principle of Congruence it follows that the first harmony of the melody will be that of its key, and that the subsequent events will continue to imply harmonies congruent to the key until that key is uniquely established, (and non-congruence can hence be seen as such).

This is an extremely strong assumption, and we shall see that it is something^h of an over-simplification and causes the only errors that the algorithm makes. However, in a huge majority of cases it seems to hold good.

If transitions had unique implications of harmony/triad, then the very first transition of the melody would usually settle the question of key, since there would usually be only one hypothesis maintained under the Principle: the one whose triad was implied by that first transition. All others would be eliminated as non-congruent. This occasionally happens. If the first transition is an ascending run from the tonic to minor third, as in Fugue 14, of Book I, then the only hypothesis which both contains all three notes as non-accidentals, and contains the minor third in its triad is the true key of F# minor.



(In fact F# minor is also confirmed by the internal implication of the first interval of the run. The major tone F# to G# implies the perfect fifth F# to C# also confirming F# minor. The algorithm is designed to note the fact that the two interpretations both arise from the same hypothesis as to the key, and to terminate, giving that as the answer. See Appendix IV Section 5, as for all examples here.)

A similar thing happens in No.7 of Book I, where the first transition is a jump from dominant to a mediant inflection. There is only one key in which all of these notes can be accommodated, and it is also one which the transition supports, so it is immediately returned as the answer.



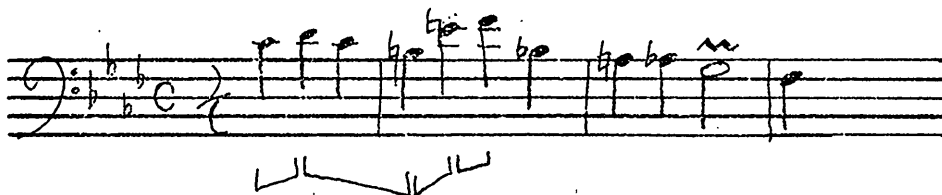
More generally, even the simplest jump of a fifth between two simple notes carries the implication of two triads, let alone the further proliferation of implications introduced by runs and jumps involving ambiguous seconds and sevenths. This means that in general the first transition will not settle the question of key,

and the earliest possibility of a decision will be at the second transition. An example is the A $\flat$ major no. 17 of Book I.



The initial fifth causes two hypotheses of A $\flat$ minor and major to be set up. One of these is immediately confirmed by the second interval to the major third, at the expense of the other, and since it is the only survivor, the corresponding key is yielded as the answer.

Where the first transition is highly ambiguous, the analysis may need to be extended through several more transition. The strongly chromatic subject of the F minor, Fugue 12 of Book I, is a striking example.



The first transition, being a single semitone jump, advances many hypotheses as to the key, of the piece, including two interpretations of the interval associated with F minor. The second transition is the chromatic scale from the minor sixth to the leading note of the dominant of the true key. This is interpreted in various

ways according to the various hypotheses, but in all cases the transition contains an accidental to the key involved. This case raises some details of the algorithm that has not been mentioned so far.

Transitions which include an accidental, or are through imperfect intervals, have undefined implications, and therefore give rise to undefined virtual notes. In such cases the point at which the next transition should be considered to start is therefore undefined. In such cases, naturally enough the real starting point of the next transition is taken.

The second point raised at this point is that, just as in the first key algorithm, it may happen that all going hypotheses are eliminated by the criterion. This is allowed under the Principle of Congruence and, in this algorithm as in its predecessor, hypotheses are reinstated when this happens. However, just as the criterion of congruence in this algorithm is the twofold one where transitions must a) be non-accidental and b) actively support the hypothesis, so this reinstatement of failed hypotheses may happen in two ways. If no hypotheses involve the attribution of accidental status to the transition, as well as all being unsupported by it, then all are restored.

This last is what happens at this stage in the F minor, so it goes on to the next transition still with many hypotheses going.

But the accidental that caused this situation was the last note of the chromatic scale, so it is the first note of this

third transition, and so this transition too is accidental to all hypotheses. All are restored with the value of their current virtual note undefined, to indicate that it failed to draw any strong interpretation of that interval.

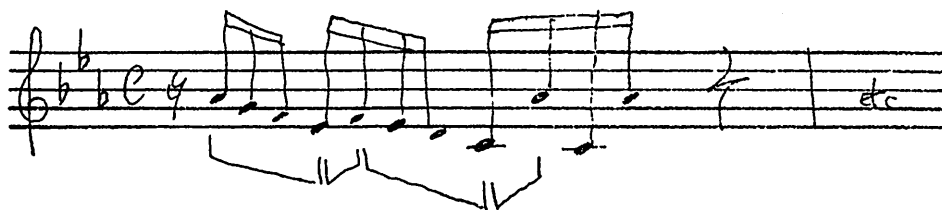
The next interval is really from the leading note of F to its tonic. Under all other hypotheses this semitone is chromatic with no implications, so although this semitone is as ambiguous in its implications as those that went before - out of the many hypotheses that have been maintained up to this point, only one is confirmed by one interpretation of this note, as the fifth from dominant to tonic of the true key, F minor.

The working of the algorithm on such a complicated example is far too complicated by the the number of different possibilities involved, and for the more detailed working of this example the reader is referred to Appendix IV, section 5. However, the 'reasoning' involved in the particular interpretation that serves as the key can be readily paraphrased as follows

"The first interval is a semitone from dominant to minor sixth, implying the harmony of the tonic minor. It is followed by a chromatic scale onto the leading note of the dominant, whose harmony is uncertain, but which does not suggest that this hypothesis is wrong. The next interval is also of uncertain harmony, but goes onto the leading note. The resolution of the leading note onto the tonic, with its cadential flavour of the fifth to the tonic, confirmed the idea that this is the right key."

This is admittedly a pretty 'free' translation from machine terms, but is intended to suggest that this analysis corresponds rather closely to the way in which a person would perceive this tune, and to suggest that, although the key decision required 6 notes, in 4 transitions, this must be about the time it takes us to really 'understand' this complex melody, and get our harmonic bearings.

The C minor organ fugue was originally introduced as an example of how even a Bach fugue subject could begin other than on the tonic or dominant.



The first transition is a run, beginning with a semitone. Of the many implications, the non-accidental ones are concerned with the keys 3 major and 8 major, besides the correct 0 minor. The second transition of a tone from $E\flat$ to F lends support to 3 major (because it is virtual in that interpretation, with three implications. 8 major is supported directly, and 0 minor is supported by a 'virtual' interval G to F, interpreted as G to C. Since a decision has not yet been reached, the next transition is examined, and only the hypotheses involving 0 minor and 3 major survive, since the transition is accidental to 8 major. The next lends no support to 3 major, but maintains C minor under virtual interpretation. The results of the program corresponding to this algorithm are summarised in Table 6, and further examples of its operation

are given in Section § 5 of Appendix IV. The details of the program are further elaborated in Section § 4 of the same Appendix. The remaining concern in this section will be to deal with two errors that the program makes. These two errors are both very similar, both arise from the same assumption of the algorithm, and suggest the form that its successor may take. Curiously enough, the trouble does not arise from the complicated rules for runs and ambiguous intervals, but from the simplest and most primitive assumption of the whole algorithm.



It was remarked even for the simple proposals that formed the introduction to this algorithm, that once we count the first three notes of the D major of Book II as one, (as we must) the subject begins with the notes of the subdominant triad. The algorithm is squarely misled, because it always assumes that the first harmony to be established is that of the tonic, and this harmony could hardly have been more directly established. Its analysis is in fact very reasonable: it thinks it's listening to something like this:



'La Cucaracha' differs only in that the repeated first note is at the lower octave, yet in this case the first few notes do describe the tonic. The different interpretation can only be arrived at by taking into account the further development of the melodies, and in particular the fact that the whole of this subject is an extended resolution, or cadence, onto the final harmony, which is then seen to be the tonic. Such an analysis, in terms of progressions of tonalities is beyond the scope of this work, but offers a strong hint about the nature of better algorithms.

In the rather similar case of the B major of Book II, the first few intervals are again the most straightforward of thirds and fifths, and again the first two between them make up a triad which is not that of the key, but of its relative minor. Whereas in the last case it seemed that the program correctly identified the harmony, but misinterpreted it as the tonic, in this case the limitation seems to be in the way it identifies the extent of the harmony. The third note does not belong with the first two, but rather with those that follow it, in the triad of the subdominant. This again raises the question of analysis of sequences of harmonies, but in a different way. It shows that

the way of identifying notes with a harmony which this algorithm uses, which is that if a note can be interpreted in the current harmony then it must be, will also have to be revised. Such a proposal, like the last one, also implies a quite different order of rules, in particular, ones taking account of contexts at least as extensive as those needed to explain these two melodies.

Fugue	Number of notes	Notes examined	Transitions examined	Key found	Remarks
1	24	5	2	0 major	
2	9	5	3	0 minor	
3	4	4	3	1 major	
4	22	15	3	1 minor	
5	9	5	2	7 major	*
6	24	4	1	2 minor	
7	21	5	4	3 major	
8	14	9	5	3 minor	
9	6	6	4	4 maj./9 maj.	amb.
10	46	3	1	4 minor	
11	23	8	3	5 major	
12	27	5	2	5 minor	
13	29	18	9	6 major	
14	18	3	2	6 minor	
15	43	4	3	7 major	
16	21	11	10	7 minor	
17	22	4	3	8 major	
18	25	3	1	8 minor	
19	20	8	5	9 major	
20	11	4	3	9 minor	
21	24	6	2	10 major	
22	30	4	3	10 minor	
23	12	3	2	8 minor	*
24	25	5	2	11 minor	

Table 5 (contd.)

(b) Book II

* indicates an erroneous result, amb. an ambiguous one.

Fugue	Number of notes	Notes examined	Transitions examined	Key found	Remarks
1	16	9	5	0 major	
2	22	5	2	0 minor	
3	22	6	2	1 major	
4	5	3	2	1 minor	
5	15	7	3	2 major	
6	12	4	1	2 minor	
7	25	4	1	3 major	
8	13	3	2	3 minor	
9	13	13	5	4 major	
10	25	3	2	4 minor	
11	19	6	2	5 major	
12	14	6	4	5 minor	
13	18	7	3	6 major	
14	18	3	1	6 minor	
15	31	7	4	7 major	
16	12	4	3	7 minor	
17	7	3	2	8 major	
18	17	5	3	8 minor	
19	8	8	8	9 maj./6 min.	amb.
20	32	6	1	9 minor	
21	38	4	3	10 major	
22	7	4	3	10 minor	
23	14	6	3	11 major	
24	21	3	2	11 minor	

Table 5 .
(a) Book I.

APPENDICES.

- Appendix I : First metrical Algorithm.
- Appendix II : Second Metrical Algorithm.
- Appendix III: First Key Algorithm.
- Appendix IV : Second Key Algorithm.
- Appendix V : Musical Terms.
- References.

First Metrical Algorithm

1. Principal Functions.

This Section is referred to in Section § 2 of Part II.

Function SUBEXEC is the principal function. It examines the melody TUNELIST, a list of NOTES, first for purely durational information, such as is given by the 'dactyl' configuration, under function MAIN. Then, in conjunction with the stored representation of the metrical aspects of the tune in METRIC, it examines for the 'isolated accent' condition, under function IISR.

A NOTE is a POP2 record.

A NOTE has a VALUE, a DURATION, and an OCTAVE.

```

FUNCTION SUREXEC TUNELIST;
VARS V FLAG T1 DT1 EXTRA WW XX;;

L1: IF TUNELIST.NULL THEN METRIC EXIT;
    LLL->V;
    ID(TUNELIST)->T1;
    FUNNY(T1)->DT1;
    1->FLAG;
    IF NOT(V=UNDEF)
        THEN IF VALUE(T1)="T" OR VALUE(T1)=UNDEF
            THEN 0->FLAG; GOTO L3
            CLOSE;
            IF DT1<V THEN 0->FLAG
            CLOSE;
        CLOSE;
    MAIN(TL(TUNELIST))->V;
    RESNABL2(V,LLL)->LLL;
    TL(3);.NEWPR;
L3: IF NOT(FLAG) THEN IISR(METRIC)->WW->XX;
    IF WW AND RESNABL1(XX,V)
        THEN LAST(METRIC)::TUNELIST->TUNELIST;
        ALLBUTL(METRIC)->METRIC;
        RESNABL2(XX,V)->LLL;
        NL(3);PR('**IAR APPLIES');
        .NEWPR;
    CLOSE;CLOSE;
    IF CMU.NULL
        OR TL(CMU).NULL THEN CMU->EXTRA; GOTO L2 CLOSE;
    IF LLL=HD(CMU) THEN TL(CMU)->EXTRA ELSE CMU->EXTRA; CLOSE;
L2:
    METRIC<>[%SUBSTRUC(MOVE(LLL,FLAG),EXTRA)%]->METRIC;
    METRUPD(METRIC)->METRIC;

    GOTO L1;
END

```



```
FUNCTION MAIN TUNELIST;
  VARS T2 T3 T4 DT2 DT3 DT4;
```

```
  IF TUNELIST.NULL THEN DT1 EXIT;
```

```
  HD(TUNELIST)->T2;
  DURATION(T2)->DT2;
```

```
  IF VALUE(T2)=UNDEF THEN DT1 EXIT;
  IF DT2>=DT1 THEN DT1 EXIT;
```

```
  IF TL(TUNELIST).NULL THEN DT1 EXIT;
  HD(TL(TUNELIST))->T3;
  DURATION(T3)->DT3;
  IF VALUE(T3)=UNDEF THEN DT1 EXIT;
  IF DT3>DT2 THEN DT1-DT2 EXIT;
  IF DT3<DT2 THEN DT1 EXIT;
```

```
  IF TL(TL(TUNELIST)).NULL THEN DT1 EXIT;
  HD(TL(TL(TUNELIST)))->T4;
  DURATION(T4)->DT4;
  IF VALUE(T4)=UNDEF THEN DT1 EXIT;
  IF DT4<DT3 THEN DT1 EXIT;
  IF RESNABL1(DT1+DT2+DT3,LLL) THEN DT1+DT2 +DT3; 1->FLAG;
  ELSE DT1-DT2-DT3;
  EXIT;
```

```
END
```

```
FUNCTION IISR METRE;
  VARS TV L SUBL;
  FUNNY(HD(METRE))->SUBL;
  U->TV;
  SUBL->L;
  L1: IF METRE.NULL OR TL(METRE).NULL THEN UNDEF;FALSE
    EXIT;
    IF TL(TL(METRE)).NULL
      AND HD(TL(METRE)).STRESS=1
      AND HD(METRE).STRESS=0
      AND TV
    THEN L+SUBL;TRUE EXIT;
    IF HD(METRE).STRESS=1 THEN SUBL->L; TRUE->TV
    ELSE L+SUBL->L
  CLOSE;
  TL(METRE)->METRE;
  GOTO L1;
END
```

This Section is referred to in Section § 2 of Part II.

The output yielded by the program as a 'trace' of its workings follows, for the examples of Fugues 2 and 15 of Book I, which illustrate the effect of both the durational rules, (both for 'long notes' and for 'dactyle'), and the 'Isolated Accent Rule'. In the traces that follow, the result of a successful application of the durational rules of function MAIN is indicated by the message "NEW METRE AT", followed by a number indicating the number of old groups that make up the new unit, followed by the words "AT GROUP", and a number identifying the ordinal number of the group of the old metre at which the new evidence was forthcoming, and hence identifying the phase. (The first group is numbered 1). A successful application of the Isolated Accent Rule is indicated in the same way, except that the message is prefixed with the message "***IAR APPLIES".

'NEW METRE OF' 2'AT GROUP' 3
'OF' 2./ 16

'NEW METRE OF' 2'AT GROUP' 4
'OF' 2./ 16

'**IAR APPLIES'
'NEW METRE OF' 2'AT GROUP' 5
'OF' 2. 2./ 16

' TIME SIGNATURE IS '* 2. 2. 2./ 16

' METRIC STRUCTURE IS

[[[([UNDEF 4)]]([(0 2)]]([(11 2)]]]]([(0 4)]]([(7 4)]]]
[[[([(3 4)]]([(0 2)]]([(11 2)]]]]([(0 4)]]([(2 4)]]]
[[[([(7 4)]]([(0 2)]]([(11 2)]]]]([(0 4)]]([(2 4)]]]
[[[([(5 2)]]([(7 2)]]([(8 4)]]]]([(7 2)]]([(5 2)]]]]([(3 2)]]([(0 2)]]]
[[(11 2)(UNDEF 14)]]

'**IAR APPLIES'
 'NEW METRE OF' 3'AT GROUP' 4
 'OF' / 8

' TIME SIGNATURE IS '* 3./ 8

' METRIC STRUCTURE IS '

```
[[[ ( 7 4) ] [ [ ( 9 2) ( 7 2) ] ] [ [ ( 6 2) ( 7 2) ] ] ]
[[[ ( 9 4) ] ] [ ( 11 2) ( 9 2) ] [ ( 7 2) ( 9 2) ] ]
[[[ ( 11 4) ] ] [ ( 9 4) ] [ ( 7 4) ] ]
[[[ ( 2 4) ] ] [ ( 0 4) ] [ ( T 4) ] ]
[[[ ( 11 4) ] ] [ ( 9 4) ] [ ( 7 4) ] ]
[[[ ( 9 4) ] ] [ ( 4 4) ] [ ( T 4) ] ]
[[[ ( 2 4) ] ] [ ( 4 2) ( 2 2) ] [ ( 0 2) ( 11 2) ] ]
[[[ ( 9 4) ] ] [ ( 0 2) ( 11 2) ] [ ( 9 2) ( 7 2) ] ]
[[ ( 6 4) ( UNDEF 8) ] ]
```

Second Metrical Algorithm

1. Principal Functions.

This Section is referred to in Section § 3.4 of Part II.

```

RECORDDEFNS("NOTE",[0 0 0 0 0])->TIE->SPACEPOS->OCTAVE->DURATION
->VALUE->DESTNOTE->CONSNOTE;
RECORDDEFNS("GROUP",[0 0])->GLIST->STRESS->DESTGROU->CONSGROU;
RECORDDEFNS("RUN",[0])->RUNLIST->DESTRUN->CONSRUN;
RECORDDEFNS("ALT",[0])->ALTLIST->DESTALT->CONSALT;
RECORDDEFNS("INT",[0 0 0 0])->DIR->DUR->SIZE->STARTVAL->DESTINT->CONSINT;

```

```

FUNCTION RMAIN REMLIST RECLIST CMUSIG;
VARB FST PERIOD RUNSAVE RUNRES;
NIL->RUNSAVE;
FALSE->RUNRES;
IF RUNTEST(REMLIST,CMUSIG) THEN
L1:FIRSTCMU(REMLIST,CMUSIG)->FST;.ERASE;
  IF ISRUN(RUNSAVE,HD(FST))
  THEN REPTTEST(REMLIST,RECLIST,CMUSIG,RUNRES)->PERIOD;
    RUNRES+1->RUNRES;
    IF PERIOD>1
    THEN RECUPD(REMLIST,RECLIST<>RUNSAVE,PERIOD,CMUSIG,
      RUNUPDFN);
    EXIT;
  RECUPD(REMLIST,RUNSAVE,PERIOD,CMUSIG,FIRSTCMU)
  ->REMLIST->RUNSAVE->CMUSIG;
  GOTO L1;
CLOSE;
CMUSIG;

```

```

RECLIST<>[F%CONSRUN(RUNSAVE)%];
REMLIST;
EXIT;
IF RECLIST.NULL
THEN IF ALTRTEST(CONSREST(NIL)::REMLIST,CMUSIG)
THEN CONSREST(NIL)::REMLIST->REMLIST;
CLOSE;
CLOSE;
IF ALTRTEST(REMLIST,CMUSIG)
THEN IF TL(OLDSIG).NULL AND CMUSIG.NULL
THEN RECUPD(REMLIST,RECLIST,2,CMUSIG,ALTUPFN1)
->REMLIST->RECLIST->CMUSIG;
CLOSE;
REPTTEST(REMLIST,RECLIST,CMUSIG,FALSE)->PERIOD;
RECUPD(REMLIST,RECLIST,PERIOD,CMUSIG,ALTUPFN2)
EXIT;
REPTTEST(REMLIST,RECLIST,CMUSIG,FALSE)->PERIOD;
RECUPD(REMLIST,RECLIST,PERIOD,CMUSIG,FIRSTCMU)
E ID

```

```

FUNCTION REPTEST REMLIST RECLIST CMUSIG MIDRUN2;
VAR3 MIDRUN2A CMU ICOUNT JCOUNT NOTECOUNT CURRLIST PREVLIST
OVERLAP MIDRUN1 AHEDLIST REMLIST3 REMLIST2;
0->ICOUNT;
;
RECLLENGTH(RECLIST)+MIDRUN2->JCOUNT;
0->MIDRUN1;

INITIAL: IF DEBUG THEN NL(1); PR('**LI:'); CLOSE;
IF RECLIST.NULL THEN 0
EXIT;
0->OVERLAP;
1->NOTECOUNT;
REMLIST->REMLIST3;
RECLIST->AHEDLIST;
FIRSTCMU(REMLIST, CMUSIG)->CMU->REMLIST2;
LISTNOTE(HD(CMU))->CURRLIST;
LISTNOTE(HD(RECLIST))->PREVLIST;
IF NOT(SANEDD(HD(PREVLIST), HD(CURRLIST)))
THEN FUNNYLEN(HD(RECLIST))+ICOUNT->ICOUNT;
TL(RECLIST)->RECLIST;
GOTO INITIAL;

CLOSE;
IF NOT(NICE(JCOUNT-ICOUNT))
THEN IF HD(RECLIST).DATAWORD="RUN"
THEN IF RUNLIST(HD(RECLIST)).LENGTH > 1
THEN 1->MIDRUN1
ELSE 0->MIDRUN1
CLOSE;
FUNNYTL(RECLIST)->RECLIST;
1+ICOUNT->ICOUNT;
GOTO INITIAL;

CLOSE;
FUNNYLEN(HD(RECLIST))+ICOUNT->ICOUNT;
TL(RECLIST)->RECLIST;
GOTO INITIAL

CLOSE;

```

Comment The above loop tests for the minimal conditions under which a repetition might be entertained for the note at the head of REMLIST, the remainder of the tune, the figure starting on a group of the preceding tune, held in RECLIST, whose number is given by ICOUNT. If these minimal conditions are not satisfied, attention passes to the successor in RECLIST. Otherwise a test is made for an actual repetition, as follows.

```

LVIRTUAL:IF DEBUG THEN NL(1);PR('**LV:MR1,MR2,MR2A-');
PR(MIDRUN1);PR(MIDRUN2);PR(MIDRUN2A); CLOSE:
IF NOT(SAMEDD(HD(PREVLIST),HD(CURRLIST)))
THEN FUNNYLEN(HD(RECLIST))+ICOUNT->ICOUNT;
TL(RECLIST)->RECLIST;

```

```

GOTO LINITIAL
CLOSE;
IF SAME SIZE(HD(PREVLIST),HD(CURRLIST))
AND NOT(MIDRUN1) AND NOT(MIDRUN2)
THEN GOTO LREAL
CLOSE;
TL(PREVLIST)->PREVLIST;
TL(CURRLIST)->CURRLIST;
IF PREVLIST.NULL
THEN FUNNYTL(AHEDLIST)->AHEDLIST;
IF AHEDLIST.NULL THEN
FUNNYLEN(HD(RECLIST))+ICOUNT->ICOUNT;
TL(RECLIST)->RECLIST;
GOTO LINITIAL CLOSE;
LISTNOTE(HD(AHEDLIST))->PREVLIST;
IF DATAWORD(HD(AHEDLIST))="RUN"
THEN 1->MIDRUN1 ELSE 0->MIDRUN1
CLOSE;
CLOSE;
IF CURRLIST.NULL
THEN IF NOT(MIDRUN2A) THEN 0->MIDRUN2 CLOSE;
IF MIDRUN2
THEN IF NOT(RUNTEST(REMLIST2,CMUSIG))
THEN 0->MIDRUN2A;
CLOSE;
CLOSE;
FIRSTCMU(REMLIST2,CMUSIG)->CMU->REMLIST2;
LISTNOTE(HD(CMU))->CURRLIST;
CLOSE;
GOTO LVIRTUAL;

```

Comment. Starting from the point in RECLIST indicated by ICOUNT, the function examines the tune RECLIST in 'bites' of a current metric unit, taken by the function FIRSTCMU. Each bite is examined first for being at least a virtual repeat of the earlier sequence, then, in the next loop, for being a real repeat. Several virtual units may need to be examined before a real part is found in the next loop, and attention may pass back and forth between the two for some time.


```

LREAL: IF DEBUG THEN NL(1); PR('**LR:'); CLOSE;
      IF NOT(SAMEHD(HD(PREVLIST), HD(CURRLIST)))
      THEN IF NOTECOUNT > 2
            THEN .REPRINT;
            JCOUNT- ICOUNT; RETURN;
            ELSE ICOUNT+FUNNYLEN(HD(RECLIST))->ICOUNT;
            TL(RECLIST)->RECLIST;
            GOTO LINITIAL;
      CLOSE;
CLOSE;
IF NOT(SAMESIZE(HD(PREVLIST), HD(CURRLIST)))
  THEN IF NOTECOUNT > 2
        THEN .REPRINT;
        JCOUNT-ICOUNT; RETURN;
        ELSE 1->NOTECOUNT; GOTO LVIRTUAL;
  CLOSE;
CLOSE;
TL(PREVLIST)->PREVLIST;
TL(CURRLIST)->CURRLIST;
1+NOTECOUNT->NOTECOUNT;
IF PREVLIST.NULL
  THEN FUNNYTL(AHEDLIST)->AHEDLIST;
  IF AHEDLIST.NULL
    THEN FIRSTCMU(REMLIST3, CMUSIG)->AHEDLIST->REMLIST3;
    1->OVERLAP;
  CLOSE;
  LISTNOTE(HD(AHEDLIST))->PREVLIST;
CLOSE;
IF CURRLIST.NULL
  THEN FIRSTCMU(REMLIST2, CMUSIG)->CMJ->REMLIST2;
  LISTNOTE(HD(CMJ))->CURRLIST;
CLOSE;
GOTO LREAL;
END

```

Comment. An exact repetition may nevertheless be a part of the virtual part if it contains very few notes, and so attention may return to the loop LVIRTUAL. If the exact repetition is long enough, then it is taken to be the required repetition, under the principle of earliest and longest figures.

This section is referred to in Sections § 3.2, 3.3 and 3.5 of Part III.

The output yielded by the program as a trace of its workings follows, for the examples of Fugues 15 and 20 of Book I and 4, 15 and 18 of Book II. Whenever a repetition is successfully established, the message "REPETITION OF FIGURE AT" is printed, followed by a number identifying the group of the old metre at which the figure, whether virtual or real, begins. (Again, the first group is numbered 1). This in turn is followed by the message "BY GROUP", again followed by a number, which identifies in the same way the group at which the repetition begins. If the repetition is overlapping, then the whole message is prefaced with the adjective "OVERLAPPING....".

I_2

REPETITION OF FIGURE AT GROUP 1 BY GROUP 3

TIME SIGNATURE IS 2. * 2. 2. 2. / 16

METRIC STRUCTURE IS

```

EEEE(UNDEF 4)JEE( 0 2)JE( 11 2)JJJEE( 0 4)JE( 7 4)JJJEE( 8 4)JEE( 0 2)JE( 11 2)
JJJEE( 0 4)JE( 2 4)JJJ
EEEE( 7 4)JEE( 0 2)JE( 11 2)JJJEE( 0 4)JE( 2 4)JJJEE( 5 2)JE( 7 2)JJEE( 8 4)J
JJEE( 7 2)JEE( 5 2)JJEE( 3 2)JE( 0 2)JJJJ
EE( 11 2)(UNDEF 14)JE(UNDEF 16)JJ
    
```

' REPETITION OF FIGURE AT GROUP ' 3'BY GROUP' 5
' REPETITION OF FIGURE AT GROUP ' 1'BY GROUP' 4

' TIME SIGNATURE IS ' 3. 2.* 3./ 8

' METRIC STRUCTURE IS '

[[[[[(7 4)]][((9 2)(7 2)]][[[((6 2)(7 2)]][[[[[((9 4)]][((11 2)(9 2)]][((7 2)(9 2)]][[[[[((11 4)]][((9 4)]][((2 4)]][((0 4)]][((4)]][[[[[((11 4)]][((9 4)]][((7 4)]][[[[[((6 4)]][((4 4)]][((4)]][[[[[((2 4)]][((0 2)(11 2)]][[[((9 4)]][((0 2)(11 2)]][((9 2)(7 2)]][[[((6 4)(UNDEF 8)]][(UNDEF 12)]][(UNDEF 24)]]

OVERLAPPING REPETITION OF FIGURE AT GROUP 1 BY GROUP 3

TIME SIGNATURE IS 2. * 2. 2. / 16

METRIC STRUCTURE IS

[[[UNDEF 4)]][(9 2)]][(8 2)]]]][(9 4)]][(11 4)]]]
[[[(0 4)]][(0 2)]][(11 2)]]]][(0 4)]][(2 4)]]]
[[[(4 2)]][(2 2)]]]][(0 2)]][(2 2)]]]][(4 4)]][(5 4)]]]
[[[(3 4)]][(4 4)]]]][(UNDEF 4)]][(9 4)]]]
[[[(11 2)]][(0 2)]]]][(9 2)]][(11 2)]]]]]][(0 2)]][(2 2)]]]][(11 2)]][(0 2)]]]
[[[(2 4)]][(0 4)]]]][(11 4)]][(4 4)]]]
[[[(9 3)]][(UNDEF 8)]]]

'REPETITION OF FIGURE AT GROUP ' 1'BY GROUP ' 4
'REPETITION OF FIGURE AT GROUP ' 1'BY GROUP ' 5

TIME SIGNATURE IS ' 4. 3.* / 16

METRIC STRUCTURE IS

[[[(1 2)]][(1 2)]][(1 2)]][(3 2)]][(1 2)]][(3 2)]][(8 2)]][(10 2)]][(0 2)]]
[[[(1 2)]][(3 2)]][(4 2)]]]]
[[[(5 2)]][(4 2)]][(6 2)]][(3 2)]][(8 2)]][(6 2)]][(4 2)]][(8 2)]][(6 2)]]
[(4 2)]][(UNDEF 2)]][(UNDEF 2)]]]]

First Harmonic Algorithm

1. Principal Functions.

This Section is referred to in Section § 3.2 of Part III.

Jo

```
RECORDERNS("KEY",0,[0 0 0 0 0])->KEYVAL->MODE->KPOSN->KEYLIST
->RELSpace->DESTKEY->CONSKEY;
```

```
FUNCTION HEXEC TUNELIST KCHOICE1 KCHOICE2 FUNOTATE;
VARS HYPLIST COUNTER TOPNOTE PREVIOUS V1;
1->COUNTER;
HD(TUNELIST)->TOPNOTE;
UNDEF->PREVIOUS;
L%APPLIST([0 1 2 3 4 5 6 7 8 9 10 11],
          LAMBDA X;CONSKEY(NIL,MAJOR,KEYPOS(X),"MAJOR",X);
          CONSKEY(NIL,MINOR,KEYPOS(X),"MINOR",X);
```

```
;END)%J->HYPLIST;
```

```
L1: IF TUNELIST.NULL THEN NL(2);PR('END OF TUNE');COUNTER-1->COUNTER;
      KCHOICE2(HYPLIST,TOPNOTE)->HYPLIST;
      OUTPUTH(HYPLIST);OUTPUT2(HYPLIST);HYPLIST;
EXIT;
IF TRACE1 THEN SP(1);NL(1);PR(COUNTER);NOTEPR(HD(TUNELIST));
      SP(1);CLOSE;
IF VALUE(HD(TUNELIST))=UNDEF THEN 1+COUNTER->COUNTER;
      TL(TUNELIST)->TUNELIST;GOTO L1;CLOSE;
KCHOICE1(HYPLIST,PREVIOUS)->V1;
IF V1.NULL THEN
      KCHOICE2(HYPLIST,TOPNOTE)->HYPLIST;
      OUTPUTH(HYPLIST);
      GOTO L2;
CLOSE;
IF HD(TUNELIST).VALUE=UNDEF THEN TL(TUNELIST)->TUNELIST;
CLOSE;
IF TL(V1).NULL THEN
      IF SSIGN(VALUE(PREVIOUS)-VALUE(HD(TUNELIST)),1)
      OR SSIGN(VALUE(PREVIOUS)-VALUE(HD(TUNELIST)), -1)
      THEN VALOF("V1")->HYPLIST;OUTPUTH(HYPLIST);
      INVTL(RELSpace(HD(HYPLIST)))->RELSpace(HD(HYPLIST));
      GOTO L2;
      CLOSE;
CLOSE;
HD(TUNELIST)->PREVIOUS;
TL(TUNELIST)->TUNELIST;
VALOF("V1")->HYPLIST;
IF TL(HYPLIST).NULL THEN OUTPUTH(HYPLIST);COUNTER+1->COUNTER;GOTO L1;
CLOSE; COUNTER+1->COUNTER;
GOTO L1;
```

Comment. Until a decision as to key is reached attention remains within L1 loop. The tune is examined note by note, and hypotheses selected by KCHOICE1, which eliminates accidental keys, and KCHOICE2 the tonic dominant rule. The function will either exit at the end of the tune, or having found one key, pass to L2.

```

L2: IF TUNELIST.NULL THEN NL(2);PR('END OF TUNE');
      OUTPUT2(HYPLIST);HYPLIST;
      EXIT;
      IF SSIGN(VALUE(PREVIOUS)-VALUE(HD(TUNELIST)),1)
      OR SSIGN(VALUE(PREVIOUS)-VALUE(HD(TUNELIST)),-1)
      THEN IF NOT(TL(HYPLIST).NULL) THEN GOTO L3 CLOSE;
            PREVIOUS::TUN=LIST->TUNELIST;
            COUNTER-1->COUNTER;
            INVTL(RELSPACE(HD(HYPLIST))->RELSPACE(HD(HYPLIST)));
      CLOSE;

L3: IF TUNELIST.NULL THEN NL(2);PR('END OF TUNE');
      OUTPUT2(HYPLIST);HYPLIST;
      EXIT;
      IF TRACE1 THEN NL(1);PR(COUNTER);;NOTEPR(HD(TUNELIST));
            SP(1);CLOSE;
      IF HD(TUNELIST).VALUE=UNDEF THEN 1+COUNTER->COUNTER;
            TL(TUNELIST)->TUNELIST;
      GOTO L3;
      CLOSE;
      FUNOTATE(TUNELIST)->TUNELIST;
      IF NOT(TL(HYPLIST).NULL) THEN GOTO L1 CLOSE;
      GOTO L3;

```

END

Comment. In the loop L2, the special case for notation of the semi-tone sequence is tested for. If this is the start of a semitone sequence, then notation of the sequence is carried out by the function SEMIFUN, which follows. Otherwise attention passes to L3, where the note is notated normally.

```

FUNCTION SEMIFUN SIGN;
VAR V;
IF TRACE1 THEN NL(2);PR('ENTERS SEMIFUN');CLOSE;
L%HD(TUNELIST),HD(TL(TUNELIST))%]->V;
IF TRACE1 THEN APPLIST(V,LAMBDA X;SP(1);
    NOTEPR(X);SP(1);END);
CLOSE;
HYPUPD(HYPLIST,V);
->HYPLIST;
IL(TUNELIST)->TUNELIST;
IF TL(TUNELIST).NULL THEN SEMIOUT(TUNELIST,SIGN) EXIT;
IF HD(IL(TUNELIST)).VALUE=UNDEF THEN COUNTER+1->COUNTER;
    TL(TL(TUNELIST))->TL(TUNELIST);
CLOSE;
IF TL(TUNELIST).NULL
    OR NOT(SSIGN(VALUE(HD(TUNELIST))-VALUE(HD(TL(TUNELIST))),SIGN))
    THEN SEMIOUT(TUNELIST,SIGN) EXIT;

L1: IF IL(TL(TUNELIST)).NULL THEN GOTO L2 CLOSE;
    IF HD(TL(TL(TUNELIST))).VALUE=UNDEF
        THEN 1+COUNTER->COUNTER;
        TL(TL(TL(TUNELIST)))->TL(TL(TUNELIST));
        GOTO L1;
CLOSE;
IF NOT(SSIGN(VALUE(HD(TL(TUNELIST)))-VALUE(HD(TL(TL(TUNELIST)))),
    SIGN))
    THEN L2: [%HD(TUNELIST),HD(TL(TUNELIST))%]->V;
        INVTL(RELSPACE(HD(HYPLIST)))->RELSPACE(HD(HYPLIST));
        IF TRACE1 THEN APPLIST(TL(V),LAMBDA X;NL(1);
            NOTEPR(X);SP(1);END);
            CLOSE;
            HYPUPD(HYPLIST,V)->HYPLIST;
            COUNTER-1->COUNTER;
            SEMIOUT(TL(TUNELIST),SIGN);
EXIT;
[%HD(TL(TUNELIST))%]->V;
IF TRACE1 THEN NL(1);NOTEPR(HD(V));SP(1);
CLOSE;
HYPUPD(HYPLIST,V)->HYPLIST;
TL(TUNELIST)->TUNELIST;
GOTO L1;

END

```

```

FUNCTION MAPVAL X Y;
COMMENT 'GIVES PIANO KEY VALUE OF A
POSITION (X,Y) RELATIVE TO ORIGIN C';
MOD(144+MOD(7*X,12)
+MOD(4*Y,12) ,12);
END

```

```

FUNCTION POSISHS NOTEVAL KPOS;
COMMENT 'GIVES THE SET OF INTERPRETATIONS IN
THEN SPACE FOR A PIANO KEY NOTEVAL(UE),
RELATIVE TO ORIGIN KPOS';
VARS POSHLIST W X Y Z;
NIL->POSHLIST;
-(XLIMIT)->Y;
L1:IF Y>YLIMIT THEN POSHLIST
EXIT;
(-XLIMIT)->X;
L2:IF X>XLIMIT THEN Y+1->Y;GOTO L1;CLOSE;
IF MAPVAL(X+FRONT(KPOS),Y+BACK(KPOS))=NOTEVAL
THEN CONSPAIR(X,Y)::POSHLIST->POSHLIST;
CLOSE;
X+1->X;
GOTO L2;
END

```

This Section is referred to in Section § 3.2 of Part III.

The output yielded by the program as a trace of its workings follows for the examples of Fugues 10, 12 and 24 of Book I.

In the first stage, of key-identification, the melody is examined note by note. At each note, its ordinal number in the melody is printed, followed by its keyboard pitch-value as an integer from 0, (C, B₄, etc.,) to 11, (B, C₅, etc.,). On the succeeding lines, the keys if any which are eliminated by its occurrence, if any are printed, each followed by the word "ELIMINATED". This continues note by note until one of three things happens. Either all the keys are eliminated, or the end of the time is reached. (In both cases the tonic dominant rule applied, and a message to that effect is output. In the latter case it is preceded by a message to the effect that the tune has ended). The last alternative is that only one key remains, which is then adopted as the answer. In all cases a message "DECISION AT NOTE NUMBER" is output, followed by the number of the note at which the decision was reached, and the key.

If any tune remains, it is notated according to that key. The trace of this merely identifies which note is being examined, and its keyboard value, and whether it is part of a semitone sequence. Notes within such sequences are preceded by the message "ENTERS SEMIFUN". They are not numbered, to avoid confusion, since they always begin with the note just preceding, but the numbering is obvious.

In all cases when the end of the tune is reached a message to that effect is output, followed by a sketch of the harmonic space around/

around the tonic as origin, points corresponding to notes that have occurred being marked with an X. Finally, a list of the notes of the tune with their notations in that space, relative to that origin, is printed, to give the same information more completely, although less transparently.

I, 10

1L 4J

0MINOR' ELIMINATED'
 1MAJOR' ELIMINATED'
 3MAJOR' ELIMINATED'
 3MINOR' ELIMINATED'
 6MAJOR' ELIMINATED'
 6MINOR' ELIMINATED'
 7MINOR' ELIMINATED'
 8MAJOR' ELIMINATED'
 10MAJOR' ELIMINATED'
 10MINOR' ELIMINATED'

2L 7J

1MINOR' ELIMINATED'
 4MAJOR' ELIMINATED'
 9MAJOR' ELIMINATED'
 9MINOR' ELIMINATED'
 11MAJOR' ELIMINATED'

3L 11J

2MINOR' ELIMINATED'
 5MAJOR' ELIMINATED'
 5MINOR' ELIMINATED'

4L 4J

5L 3J

0MAJOR' ELIMINATED'
 2MAJOR' ELIMINATED'
 7MAJOR' ELIMINATED'
 11MINOR' ELIMINATED'

6L 4J

7L 2J

4MINOR' ELIMINATED'
 3MINOR' ELIMINATED'

'TOLLIC DOMINANT RULE'

'DECISION AT NOTE NUMBER ' 7 4MINOR

7L 2J

3L 4J

9L 1J

10L 4J

11[0]
12[4]
13[11]
14[4]

'ENTERS SEMIFUN' [4] [3]

15[3]

'ENTERS SEMIFUN' [3] [4]

17[10]
18[1]
19[7]

'ENTERS SEMIFUN' [7] [6]

20[6]

'ENTERS SEMIFUN' [6] [7]

22[10]
23[6]
24[4]
25[2]

'END OF TUNE'

```

. . . . .
. . . . .
. . . . X X . .
. . . . X X X X .
. . . . X X X . .
. . . . . . . .
. . . . . . . .

```

'NOTATION OF SUBJECT'

```

[ 4[ 0 . 0]] [ 7[ 1 . -1]] [ 11[ 1 . 0]]
[ 4[ 0 . 0]] [ 3[ 1 . 1]] [ 4[ 0 . 0]]
[ 2[ 2 . -1]] [ 4[ 0 . 0]] [ 1[ 3 . 0]]
[ 4[ 0 . 0]] [ 0[ 0 . -1]] [ 4[ 0 . 0]]
[ 11[ 1 . 0]] [ 4[ 0 . 0]] [ 3[ 1 . 1]]
[ 4[ 0 . 0]] [ 10[ 2 . 1]] [ 1[ 3 . 0]]
[ 7[ 1 . -1]] [ 6[ 2 . 0]] [ 7[ 1 . -1]]
[ 10[ 2 . 1]] [ 6[ 2 . 0]] [ 4[ 0 . 0]]
[ 2[ 2 . -1]]

```

1[0]
 2MAJOR' ELIMINATED'
 2MINOR' ELIMINATED'
 3MINOR' ELIMINATED'
 4MAJOR' ELIMINATED'
 6MAJOR' ELIMINATED'
 6MINOR' ELIMINATED'
 8MINOR' ELIMINATED'
 9MAJOR' ELIMINATED'
 11MAJOR' ELIMINATED'
 11MINOR' ELIMINATED'

2[1]
 0MAJOR' ELIMINATED'
 0MINOR' ELIMINATED'
 3MAJOR' ELIMINATED'
 4MINOR' ELIMINATED'
 5MAJOR' ELIMINATED'
 7MAJOR' ELIMINATED'
 7MINOR' ELIMINATED'
 9MINOR' ELIMINATED'
 10MAJOR' ELIMINATED'

3[0]

4[11]
 1MAJOR' ELIMINATED'
 1MINOR' ELIMINATED'
 5MINOR' ELIMINATED'
 8MAJOR' ELIMINATED'
 10MINOR' ELIMINATED'

'TONIC DOMINANT RULE'

'DECISION AT NOTE NUMBER ' 4 5MINOR

3[0]

'ENTERS SEMIFUN' [0] [11]

5[4]

'ENTERS SEMIFUN' [4] [5]

7[10]

'ENTERS SEMIFUN' [10] [9]

[8]
[7]

11[5]
12[5]
13[7]

'ENTERS SEMIFUN' [7] [8]

'END OF TUNE'

.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	X	X	X	.	.
.	.	.	X	X	X	X	.	.
.	.	.	.	X	X	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

'NOTATION OF SUBJECT'

[0[1 . 0]] [1[0 . -1]] [0[1 . 0]]
 [11[2 . 1]] [4[1 . 1]] [5[0 . 0]]
 [10[-1 . 0]] [9[0 . 1]] [8[1 . -1]]
 [7[2 . 0]] [5[0 . 0]] [5[0 . 0]]
 [7[2 . 0]] [8[1 . -1]]

I, 24

1C 6J

0MAJOR' ELIMINATED'
 0MINOR' ELIMINATED'
 2MINOR' ELIMINATED'
 3MAJOR' ELIMINATED'
 5MAJOR' ELIMINATED'
 5MINOR' ELIMINATED'
 8MAJOR' ELIMINATED'
 8MINOR' ELIMINATED'
 9MINOR' ELIMINATED'
 10MAJOR' ELIMINATED'

2C 2J

1MAJOR' ELIMINATED'
 1MINOR' ELIMINATED'
 4MAJOR' ELIMINATED'
 4MINOR' ELIMINATED'
 6MAJOR' ELIMINATED'
 10MINOR' ELIMINATED'
 11MAJOR' ELIMINATED'

3C 11J

7MINOR' ELIMINATED'

4C 7J

3MINOR' ELIMINATED'
 6MINOR' ELIMINATED'
 9MAJOR' ELIMINATED'

5C 6J

6C 11J

7C 10J

2MAJOR' ELIMINATED'
 7MAJOR' ELIMINATED'

'DECISION AT NOTE NUMBER ' 7 11MINOR'

· 6C 11J

'ENTERS SEMIFUN' [11] [10]

8C 4J

'ENTERS SEMIFUN' [4] [3]

10[0]

'ENTERS SEMIFUN' [0] [11]

12[6]

'ENTERS SEMIFUN' [6] [5]

14[2]

'ENTERS SEMIFUN' [2] [1]
[0]

16[0]

'ENTERS SEMIFUN' [0] [1]

18[9]

19[6]

20[8]

'NOTATIONAL AMBIGUITY IN ' 11MINOR
'NOTE(S)' [8 16 3 UNDEF 8]

.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	X	X	X	X	.
.	.	.	X	X	X	X	.	.
.	.	.	X	X	X	X	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

21[6]

'END OF TUNE'

.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	X	X	X	X	.
.	.	.	X	X	X	X	.	.
.	.	.	X	X	X	X	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

'NOTATION OF SUBJECT'

[5[1 . 0]] [2[1 . -1]] [11[0 . 0]]
[7[0 . -1]] [6[1 . 0]] [11[0 . 0]]

[10[1 . 1]] [4[-1 . 0]] [3[0 . 1]]
[0[-1 . -1]] [11[0 . 0]] [6[1 . 0]]
[5[2 . 1]] [2[1 . -1]] [1[2 . 0]]
[0[3 . 1]] [1[2 . 0]] [9[2 . -1]]
[5[1 . 0]] [8UNDEF] [6[1 . 0]]

Second Harmonic Algorithm

1. Function DIADS.

This function, referred to in Section § 4.2221 of Part III, takes vectors in the space onto a list of possible interpretations as the special diad vectors, corresponding to the perfect fifth and major and minor thirds which are involved in the two common chords.

```

FUNCTION DIADS X1 Y1 X2 Y2;
COMMENT 'THE IMPLICATIONS OF A TRANSITION
INTERVAL FROM A POINT (X1,Y1) TO A POINT
(X2,Y2) ARE YIELDED AS A LIST OF LISTS
EACH OF FOUR ITEMS IDENTIFYING A DIAD
IN THE HARMONIC SPACE';
VARS X Y;
X2-X1->X;
Y2-Y1->Y;

COMMENT 'PERF. 5TH, MAJ & MIN 3RDS & INVERSES';
IF X=(-1) AND Y>=0 AND Y<1 THEN [%[%X1,Y1,X2,Y2%]]%] EXIT;
IF X=0 AND Y >=(-1) AND Y<1 THEN [%[%X1,Y1,X2,Y2%]]%] EXIT;
IF X=1 AND Y>=(-1) AND Y<0 THEN [%[%X1,Y1,X2,Y2%]]%] EXIT;

COMMENT 'DIATONIC MAJ 7TH & SEMITONE';
IF BOOLOR(BOOLAND(X=1,Y=1),BOOLAND(X=(-1),Y=(-1)))
THEN [%[%X1,Y1,X2,Y1%],
      [%X1,Y1,X1,Y2%],
      [%X2,Y1,X2,Y2%],
      [%X1,Y2,X2,Y2%]]%]
EXIT;

COMMENT 'MAJOR TONE';
IF X=2 AND Y=0
THEN [%[%X1,Y1,X1+1,Y1%]]%]
EXIT;

COMMENT 'INVERSE OF MINOR TONE';
IF X=2 AND Y=(-1)
THEN [%[%X1+1,Y2,X2,Y2%],
      [%X1,Y1,X1+1,Y1%],
      [%X1,Y2,X1+1,Y2%]]%]
EXIT;

COMMENT 'INVERSE OF MAJOR TONE';
IF X=(-2) AND Y=0
THEN [%[%X1,Y1,X1-1,Y1%]]%]
EXIT;

COMMENT 'MINOR TONE';
IF X=(-2) AND Y=1
THEN [%[%X1,Y1,X1-1,Y1%]]%]
EXIT;

COMMENT 'ANYTHING ELSE';
[%[%X1,Y1,X1,X1%],[%X1,X1,X1,X1%],[%X1,X1,X1,X1%],[%X1,X1,X1,X1%]]%];

```


2. Syntax and Parsing

2.1 Terminology.

This Section is referred to in Section § 4.3 of Part III. The natural way to talk about the syntax definitions has been in terms of a hierarchical structure of transitions and points, and, (in the case of, for example, the rules which restrict the definitions of inflections in the context of runs), in terms of 'context sensitivity'. However, there are no recursive definitions among the rules, and the contexts that have to be examined are extremely local. Although, therefore, the parsing can be carried out by a finite-state machine, that would be an unrevealing way of presenting the definitions. Instead, they are presented below in a form which relates closely both to the informal definitions that were used to introduce the ideas, and to a simple algorithm for parsing according to the definitions.

In order to present it in this way a few preliminary remarks are necessary.

Among the informal definitions given so far, some have been made in terms of the notes which make up the melody, while others have been made in terms of the types of intervals (second, etc.,) between the notes of the melody. From now on all definitions will be made in terms of intervals between notes. This simplifies some rules, but there is an obvious correspondence between intervals and notes, and it will usually do no harm to think of the intervals as notes in considering the rules.

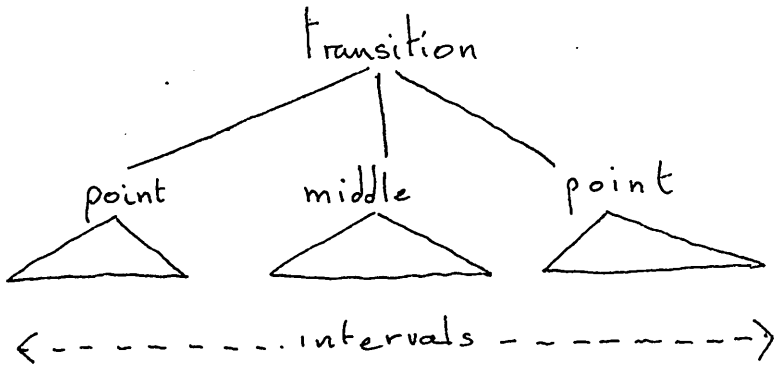
In the definitions below certain abridgements have been made upon earlier definitions. In particular, where for example a definition/

definition restricts a constituent to a context which is neither any second nor an ascending run, (say), we can express this completely by merely making the prohibited context that of the second, since the run is made up of such.

The following connectives are used to express the syntax definitions.

is made up of: This corresponds closely to the rewrite arrow of immediate constituent grammars of natural language. Like the rewritten arrow, it defines a set of structural descriptions for the strings of a language, and defines the set of legal 'sentences' of the language. Unlike the cases usually involved in the study of natural language, the one we are concerned with here is the universal language of all possible strings on the alphabet unison, second, etc., of terminal symbols. Another difference between this syntax and the usual grammatical paradigm, (but a trivial one), is that the structures generated are not strictly trees, since the end point of a transition is the same as the beginning point of its successor.

This relation is used to define one skeleton description of melodies, in which the string is made up of units called transitions each conforming to the following pattern.



Each node in the above descriptions has features attached to it, and it is upon the values of these features that the different descriptions of the transitions depend. These are dealt with next,

has: The relation of a symbol to the features associated with it.

Each feature is further defined by the following term.

is: Applied to a feature name, this identifies the values that the feature may take as members of a given set of such values.

2.2 Syntax

2.21 Part I. Definitions

A transition is made up of a point, (from which it starts), a middle, (through which it moves), and a point, (to which it moves).

A transition has a type.

A type is {run, jump}

A point is made up of a sequence of intervals.

A point has a type, a size, a unit duration and a direction.

A type is {turn, inf, rep, simple}.

A duration is a positive integer.

A size is {0, 1, ..., 6}.

A direction is {-1, 0, 1}.

An interval has a size, a duration and a direction.

A size is {0, 1, 2, ..., 6}.

A duration is a positive integer.

A direction is {-1, 0, 1}.

A middle is made up of a sequence of intervals.

2.22 Part II: Constraints.

Some shorthand conventions are used here. An interval whose size is 1 is referred to as a second; if its direction is 1 it is referred to as an ascending second; if it is -1 then as a descending one.

Similarly, points and transitions are occasionally referred to just by their type. It will be clear from the underlining of such (e.g. run) that they are syntactic objects, not just feature values.

A transition of type run has its first point of size 1 (a second) (and that point is of type other than turn) (followed by at least one second of the same duration and direction.) (Its last point is not a turn.)

A transition of type jump has as its first and last points ones whose features are unconstrained (except that they do not involve it in the above definition).

A point of type turn is a sequence of four intervals, constrained as follows. The first two are descending seconds, the third is an ascending second; the second and third seconds have the same duration as the first. The fourth interval of the turn can take any value for its feature, except that it must not be an ascending second with duration equal to that of the first three. The turn may not be preceded by a transition of type run, the first of whose seconds are of the same duration as those of the turn.

A point of type inflection is a sequence of three intervals, constrained as follows. The first is a descending second, the second is an ascending second, and the latter has the same duration as the first. The inflection may not be both preceded by a descending second of the same duration as the first interval of the putative inflection

point, and such that the third interval is an ascending second of the same duration. It may not be preceded by a run whose seconds are of duration equal to that of the first interval of the inflection.

A point of type repetition is a sequence of any number greater than 1 of intervals, subject to the constraint that all except the last have size 0, and : : have the same duration as the first.

A middle is a sequence of seconds, each of the same direction and duration as that of the preceding point if the transition that contains it is a run. Otherwise it is the null sequence.

A point of type simple is a single interval whose feature values are unconstrained (except that they do not involve it in any of the above definitions).

2.3 Implementation.

In the above definitions, the statements appended to each within brackets, usually of the form "(except where the values (of the features in question) would involve the constituent in question in one of the previous definitions)", are most naturally handled in the parsing algorithm by precedence. Thus by testing, (for example), for the presence of an inflection before testing for a point of type simple, (the unconstrained single interval), it is made impossible that an interval properly belonging to the point of higher precedence should be parsed as belonging to that of the lower precedence.

In the sections of program that follow, the constraints bracketted in the above definitions will be handled in this way, by precedence, and will be reflected in the programs as an ordering of certain tests. For example, since a jump is a transition whose only constraint is that it should not be part of a run, the function TRANS corresponding to the definition of a transition above simply consists of a test for type run, followed by the test for type jump. (See section § 2.4 below).

In the two syntactic categories, of transitions and points, the precedences of various types of each can be simply expressed, where the symbol > is to be read as "is greater precedence than".

Among transitions,

run > jump

Among points,

turn, inflection, repetition > simple

It/

It has already been mentioned how function TRANS below reflects the first of these orderings. The function PCINTS in the same Section § 2.4. reflects the second.

Once having stated the definitions in terms of precedence, the algorithm to parse strings according to them is straightforward, and involves "filling in" the feature values of the constituent-structure framework of Figure . For some transition in the melody, other than the first, its starting point has already been parsed, since it is also the end point of its predecessor. The type of the transition, its middle and end point are then established. This is done by the principal parsing function NEXTRANS, in Section § 2.4, which corresponds exactly to the aforesaid Figure .

It will be noted that, while the integrity of the various categories of constituent and feature are preserved in this selection of functions, (they all appear as identifiable bits of program,) the distinction between constituents and certain of their feature values is not, in-as-far as both appear as POF2 functions. For example, both the syntactic category transition, and the values of its feature type run and jump, appear as functions, TRANS, and RUN and JUMP, respectively. The distinction was only introduced in the first place for reasons of exposition to make it as clear as possible that this syntax is very simple. As far as its mere definition is concerned, either category could have been expanded at the expense of the other, and the program merely reflects the vacuousness of the distinction by representing them in the same way.

The functions corresponding to the various values of the type for points/

points are TURN, INF, REP, and NOTE, corresponding to the types turn, inflection, repetition and simple, respectively. Together with the transition types, functions RUN and JUMP, and given the implicit constraints of precedence, their form corresponds closely with the definitions of Section § 2.2.

2.4 Principal Functions.

In the following parsing functions, at any stage there is generally "partial parse" of some low-level constituent which must be repeatedly examined in the parsing of its higher-level parent-constituent. This lower object may either be some intermediate point or a terminal of the syntax, an interval. It is convenient to keep the partial parse at the head of the remainder of the melody, and to access it with the general function NEXTHING, which is like the standard function HD, except that where the next interval has not yet been defined, by the "notation" procedures of the next Section § 3 of this Appendix, the function involves those procedures.

[LOS TRIANGLE NEWPARSE]

```

FUNCTION TRANS;
COMMENT ' A TRANS IS OF TYPE RUN OR JUMP';
IF .RUN THEN EXIT;
.JUMP;
END

```

```

FUNCTION POINT;
COMMENT ' A POINT IS OF TYPE TURN, INF,
      REP, OR SIMPLE';
VARS NEW;
.NEXTTHING->NEW;
IF ID(NEW)="POINT"
THEI COMMENT 'THE NEXT POINT MAY ALREADY BE
      PARSED'; TL(TUNE);NEW;
EXIT;
IF .TURN THEN
ELSEIF .INF THEN
ELSEIF .REP THEN
ELSE .NOTE
CLOSE;
END.

```

```

FUNCTION NEXTRANS TUNE KEY;
COMMENT ' A TRANSITION IS MADE UP OF A BEGIN
      POINT, A MIDDLE & AN END POINT';
VARS BEGIN TYPE MIDDLE FINISH LASTDUR LASTYPE LASIDIR;
IF TUNE.NULL THEN FALSE EXIT;
.POINT;
->BEGIN->TUNE;
.TRANS;
->MIDDLE->TYPE->TUNE;
IF TUNE.NULL THEN NIL;UNDEF
ELSE
.POINT;
CLOSE;
->FINISH -> TUNE;
IF FINISH=UNDEF THEN NIL ELSE
FINISH::TUNE;
CLOSE;;
["TRANS",TYPE,BEGIN,MIDDLE,FINISH%];
TRUE;
END

```

L IJS TRIANGLE TRANS]

```

FUNCTION RUN;
VARS NEWDUR NEWSIZE NEWDIR NEW DUR DIR SAVE;
NIL->SAVE;
IF NOT (BEGIN.POINTSIZE=1)
THEN FALSE
EXIT;
IF LENGTH(TUNE)<2 THEN FALSE EXIT;
BEGIN.POINTDIR->DIR;
BEGIN.POINTDUR->DUR;
.NEXTHING->NEW;
SDD(NEW)->NEWSIZE->NEWDUR->NEWDIR;
IF NOT(.SECOND)
OR NOT( NEWDUR=DUR)
OR NOT( NEWDIR=DIR )
THEN FALSE
EXIT;
L1:NEW::SAVE->SAVE;
   TL(TUNE)->TUNE;
   IF TUNE.NULL THEN GOTO L2 CLOSE;
   .NEXTHING->NEW;
   SDD(NEW)->NEWSIZE->NEWDUR->NEWDIR;
   IF .SECOND
   AND NEWDUR=DUR
   AND NEWDIR=DIR
   THEN GOTO L1
   CLOSE;
L2:DIR->LASTDIR;
   DUR->LASTDUR;
   "RUN"->LASTYPE;
   TUNE;
   "RUN";
   REV(SAVE);
   TRUE;
END

```

```

FUNCTION JUMP;
BEGIN.POINTDIR->LASTDIR;
BEGIN.POINTDUR->LASTDUR;
"JUMP"->LASTYPE;
TUNE;
"JUMP";
NIL;
END

```

[LOS TRIANGLE NEWPOINT]

```

FUNCTION TURN;
VARB NEW NEWDUR NEWSIZE NEWDIR DUR SAVE;
NEXTHING->NEW;
NIL->SAVE;
IF LENGTH(TUNE)<4 THEN FALSE EXIT;
SDD(NEW)->NEWSIZE->DUR->NEWDIR;
IF LASTYPE="RUN" AND LASTDUR=DUR THEN FALSE EXIT;
IF NOT(.SECDISC) THEN FALSE EXIT;
NEW::SAVE->SAVE;
IL(TUNE)->TUNE;
NEXTHING->NEW;
SDD(NEW)->NEWSIZE->NEWDIR->NEWDIR;
IF NOT(.SECDISC)
OR NOT(NEWDIR=DUR)
THEN REV(SAVE)<>TUNE->TUNE;
FALSE
EXIT;
NEW::SAVE->SAVE;
IL(TUNE)->TUNE;
NEXTHING->NEW;
SDD(NEW)->NEWSIZE->NEWDIR->NEWDIR;
IF NOT(.SECDISC)
OR NOT(NEWDIR=DUR)
THEN REV(SAVE)<>TUNE->TUNE;
FALSE EXIT;
NEW::SAVE->SAVE;
IL(TUNE)->TUNE;
NEXTHING->NEW;
SDD(NEW)->NEWSIZE->NEWDIR->NEWDIR;
IF .SECDISC AND NEWDIR=DUR THEN REV(SAVE)<>TUNE->TUNE;FALSE
EXIT;

IL(TUNE);
L% "POINT", "TURN", REV(NEW::SAVE)%];
TRUE;
END

```

```

FUNCTION INF;
VARB NEW NEWDUR NEWSIZE NEWDIR DUR SAVE;
IF LENGTH(TUNE)<3 THEN FALSE EXIT;
NEXTHING->NEW;
NIL->SAVE;
SDD(NEW)->NEWSIZE->DUR->NEWDIR;
IF NOT(.SECDISC) THEN FALSE EXIT;
NEW::SAVE->SAVE;
IL(TUNE)->TUNE;
NEXTHING->NEW;
SDD(NEW)->NEWSIZE->NEWDIR->NEWDIR;
IF NOT(.SECDISC)
OR NOT(NEWDIR = DUR)
THEN REV(SAVE)<>TUNE->TUNE;
FALSE

```

```

EXIT;
NEW::SAVE->SAVE;
TL(TUNE)->TUNE;
.NEXTHING->NEW;
SDD(NEW)->NEWSIZE->NEWDUR->NEWDIR;
IF .SECASC
AND NEWDUR=DUR
AND LASTYPE="RUN"
AND LASTDUR =DUR
AND LASTDIR =(-1)
THEN REV(SAVE)<>TUNE->TUNE;
FALSE
EXIT;

```

```

TL(TUNE);
[%"POINT","INF",REV(NEW::SAVE)%];
TRUE;
END

```

```

FUNCTION REP;
VARS NEW NEWSIZE NEWDUR NEWDIR DUR SAVE;
NIL->SAVE;
IF TUNE.LENGTH<1 THEN FALSE EXIT;
.NEXTHING->NEW;
SDD(NEW)->NEWSIZE->DUR->NEWDIR;
IF NOT(.UNISON) THEN FALSE EXIT;

```

```

L0:NEW::SAVE->SAVE;
TL(TUNE)->TUNE;
IF TUNE.NULL THEN GOTO L1 CLOSE;
.NEXTHING->NEW;
SDD(NEW)->NEWSIZE->NEWDUR->NEWDIR;
IF NOT(.UNISON) OR NOT( NEWDUR=DUR)
THEN L1:TUNE.TL;
[%"POINT","REP",REV(NEW::SAVE)%];
TRUE;
EXIT;
GOTO L0;
END

```

```

FUNCTION NOTE;
VARS NEW;
.NEXTHING->NEW;
TL(TUNE);
[%"POINT","NOTE",[%NEW%]];
END

```

3. Notation.

This Section is referred to in Section § 4.4 of Part III.

ENDS TRIANGLE NEWINT1]

FUNCTION NEXTINT TUNE KEY;

COMMENT ' YIELDS THE NEXT INTERVAL OF THE TUNE,
TOGETHER WITH THE REMAINDER';

VARS DIR N1 N2 VAL1 VAL2 DUR1 DUR2 OCT1 OCT2 NLIS;

IF TUNE.NULL THEN UNDEF ;UNDEF;EXIT;

COMMENT 'IT MAY ALREADY HAVE BEEN PARSED';

TUNE.HD->N1;

IF N1.HD="INT"

THEN TL(TUNE);N1;EXIT;

N1.DEST.DEST.HD->OCT1->DUR1->VAL1;

IF TUNE.TL.NULL

THEN COMMENT ' THE LAST INTERVAL OF THE TUNE
IS A SPECIAL CASE';

NIL;[% "INT", VAL1, DUR1, UNDEF, UNDEF, UNDEF%];

EXIT;

COMMENT 'KEYBOARD SEMITONES ARE NOTED BY
SEMINOTE, OTHERS BY NOTATE';

TUNE.TL.HD->N2;

DIRDEF(N1,N2)->DIR;

N2.DEST.DEST.HD->OCT2->DUR2->VAL2;

IF SEMISIZE(N1,N2)=1

OR SEMISIZE (N1,N2)=(-1)

THEN SEMINOTE(VAL1,VAL2)

ELSE NOTATE(VAL1,VAL2)

CLOSE;

->NLIS;

NLIS.TL.HD::POSTLIST(KEY)->POSTLIST(KEY);

TL(TUNE);

[% "INT", VAL1, DUR1, NLIS, NLIS.SIZEOF, DIR%];

COMMENT ' AN INTERVAL HAS A VALUE A DURATION A NOTATION
A SIZE AND A DIRECTION';

END

```

FUNCTION NOTATE N1 N2;
COMMENT 'NOTATES THE INTERVAL BETWEEN TWO NOTES
        OTHER THAN A KEYBOARD SEMITONE APART';
VARS V POSLIST2 POS1 POS2 PLIST KLIST TRUTH;
IF I2=UNDEF THEN [%HD(POSLIST(KEY)),UNDEF%] EXIT;
IF KEY=UNDEF
THEN CLOSEST(POSISHS(N1),[%CONSPAIR(0,0)%]).HD.HD->POS1;
    GOTO LL;
CLOSE;
POSISHS(N2)->POSLIST2;
POSLIST(KEY)->PLIST;
KEYLIST(KEY)->KLIST;

COMMENT 'NOTATION OF FIRST IS KNOWN FROM PREVIOUS';
HD(PLIST)->POS1;

COMMENT 'MELODIC CONVENTION';
IF MELCOND(DIR)
THEN KLIST<>MELCONV(KEY,DIR)->KLIST;
CLOSE;

COMMENT 'TEST IF SECOND COULD BE IN KEY';
APPLIST(POSLIST2,LAMBDA X; IF NOT( TRUTH) THEN EXIT;
                                IF FIND(X,KLIST)
                                THEN X->POS2;
                                FALSE ->TRUTH;
                                EXIT;
                                END);
IF NOT(TRUTH) THEN [%POS1,POS2%] EXIT;

COMMENT 'IF OUT OF KEY, CHOOSE CLOSEST';
LL:IF POS1=UNDEF
    THEN CLOSEST(POSLIST2,[%HD(TL(PLIST))%]).HD.HD
ELSE CLOSEST(POSLIST2,[%POS1%])
    ->V;
    IF V.LENGTH>1 THEN NL(1);
    PR('+++++NOTATIONAL AMBIGUITY+++++');
    CLOSE;
    V.HD.HD;

CLOSE;->POS2;
[%POS1,POS2%];
END

```



```

FUNCTION SEMINOTE VAL1 VAL2;
COMMENT (NEW VERSION SEPT 16 72.  SINCE NEW FRAME
RENDERS ALL SEMITONES COMPLETELY
UNAMBIGUOUS, THE FUNCTION
NEED ONLY NOTATE ON SEMITONE AT A TIME.);

```

```

VARS TRUTHVAL SEMIFRAME KPOS KX KY POS2 POSLIST2;
KEY.KEYLIST.HD->KPOS;
KPOS.FRONT->KX;
KPOS.BACK->KY;
UNION (MAJOR(KPOS),MINOR(KPOS))
      <>[%CONSPAIR(KX-2,KY),CONSPAIR(KX-1,KY-1),
        CONSPAIR(KX+2,KY+1)%]
      -> SEMIFRAME;
POSISHS(VAL2)->POSLIST2;

APPLIST(POSLIST2,LAMBDA X;IF TRUTHVAL.NOT THEN EXIT;
      IF FIND(X,SEMIFRAME)
      THEN X->POS2;0->TRUTHVAL;
      EXIT;
      END);

[;%HD(POSLIST(KEY)),POS2%];

END

```

4. Other Interpretative Functions.

This Section is referred to in Section § 4.5 of Part III.

```

RECORDFNS("KEY",[0 0 0 0 0 0 0])->KEYREC->VIRTLAST->KEYTRIAD->KEYTUNE
->POSLIST->KEYMODE->KEYLIST
->DESTKEY->CONSKY;

```

```

FUNCTION MAIN TUNE;
COMMENT 'FINDS THE KEY OF TUNE';
VARS KEYL;
IF TRACE THEN NL(3);
LENGTH(TUNE)->TUNELENGTH; NL(1);
CLOSE;
COMMENT 'FOURTEEN KEYS INCLUDE SOME
INTERPRETATION OF THE FIRST NOTE';
POSSIBLE(HD(TUNE))->KEYL;
COMMENT 'SOME OF THESE ARE SUPPORTED
BY THE FIRST TRANSITION';
PROPOSE(KEYL,TUNE)->KEYL;
COMMENT 'THESE ARE ITERATIVELY EXAMINED FOR
SUPPORT FROM FURTHER TRANSITIONS, UNTIL
ONLY ONE SURVIVES.';
CONFIRM(KEYL,TUNE);
END

```

```

FUNCTION KEYLEXIT;
IF KEYL.HD.KEYTUNE.NULL THEN TRUE ELSE FALSE EXIT;
END

```

```

FUNCTION CONFIRM KEYL TUNE;
COMMENT 'TAKES A LIST OF KEYS, KEYL, AND
ITERATIVELY EXAMINES THE ONES GIVEN
ACTIVE SUPPORT BY THE TUNE';
IF TRACE THEN NL(3); PR("****")
;PR("CONFIRM"); STARLINE(50); CLOSE;
VARS SAVE;
NIL->SAVE;
LO: IF .KEYLEXIT
THEN COMMENT 'END OF TUNE'; KEYL EXIT;
IF KEYL.TL.NULL
THEN KEYL EXIT;
KEYL.TCOMPLET->KEYL;
IF KEYL.TL.NULL THEN KEYL EXIT;
TESTKEYS(KEYL)->KEYL;
GOTO LO;
END

```

```

FUNCTION TESTKEYS KEYL;
COMMENT 'TAKES A LIST KEYL OF KEYS, GIVES
      LIST OF SUPPORTED ONES.';
IF 4TRACE THEN NL(2); PR("***");
PR("TESTKEYS"); STARLINE(50); NL(2);
CLOSE;;
VARS SAVE2 SAVE; NIL->SAVE; NIL->SAVE2;
APPLIST(KEYL, LAMBDA X; VARS TR TU ACCFLAG; 0->ACCFLAG;
COMMENT 'FOR ALL, PARSE NEXT TRANSITION,
      NOTE IF ACCIDENTAL, & IF NOT KEEP IN
      LIST2.  TEST THE IMPLICATIONS OF THE
      TRANSITION, ADDING SUPPORTED HYPOTHESES
      TO SAVE';
      NEXTRANS(KEYTUNE(X), X).ERASE->TR->TU;
      IF TISINKEY(TR).NOT
      THEN 1->ACCFLAG;
      ELSE X::SAVE2->SAVE2; CLOSE;
      .SUBTEST;
      END);
COMMENT 'IF ALL UNSUPPORTED YIELD THOSE
      WHICH ARE NON-ACCIDENTAL IF ANY, ELSE
      ALL';
IF SAVE.NULL THEN SAVE2.VIRTFRIG ELSE SAVE EXIT;
END

```

LMJS TRIANGLE MAINTAIN]

VARS OPERATION 7 <=>;

FUNCTION PAIRED P1 P2;
IF P1.FRONT=P2.FRONT AND P1.BACK=P2.BACK THEN TRUE
ELSE FALSE;
CLOSE;
END

PAIRED-> NONOP <=>;

FUNCTION MAINTAIN IMPLIST KEY;
COMMENT 'TAKES LIST IMPLIST USUALLY CONTAINING
TWO KEY IMPLICATIONS , AND COMPARES EACH WITH
A KEY, GIVES TRUE IF SAME, ELSE FALSE.';
VARS CONFIRM;
U->CONFIRM;
IF IMPLIST.NULL THEN FALSE EXIT;
COMMENT 'SPECIAL CASES FOR RESTS AND UNISON';
IF IMPLIST.HD="RESTFRIG"
THEN HD(TL(IMPLIST))::VIRTLAST(KEY)->VIRTLAST(KEY); TRUE EXIT;
IF IMPLIST.HD="UNIFRIG"
THEN HD(TL(IMPLIST))::VIRTLAST(KEY)->VIRTLAST(KEY); TRUE EXIT;
IF ACCFLAG
THEN COMMENT 'NO KEY IS SUPPORTED BY A TRANSITION
' CONTAINING AN ACCIDENTAL'
; FALSE EXIT;
APPLIST(IMPLIST,LAMBDA XX;
IF CONFIRM THEN EXIT;
IF NOT(TYPE1(XX,KEY))
THEN EXIT;
UNION([%ISIDE(XX)%],KEYTRIAD(KEY))
->KEYTRIAD(KEY);
IVLAST1(XX)->HD(VIRTLAST(KEY));
IVLAST(XX)::VIRTLAST(KEY)->VIRTLAST(KEY);
TRUE->CONFIRM;
END);
CONFIRM;
END

FUNCTION IMPLICATS TRANSIT KEY;
COMMENT 'FOR A GIVEN KEY, A TRANSITION YIELDS A LIST
OF "DIAD INTERPRETATIONS", THEMSELVES LISTS,
NORMALLY CONTAINING TWO IMPLIED KEYS';
VARS OLIST FSTPT SECPT X1 X2 Y1 Y2 IMPS VLAST;
HD(VIRTLAST(KEY))->VLAST;
NIL->IMPS;
TRANSENDS(TRANSIT)->FSTPT->SECPT;
IF VLAST=UNDEF
THE IF FSTPT=UNDEF THEN [%[% "RESTFRIG",UNDEF%]] EXIT;
COMMENT 'THIS WILL BE V RARE - AN ACC. FOLLOWED BY A REST';

```

      FRONT(FSTPT)->X1;
      BACK(FSTPT)->Y1;
ELSE FRONT(VLAST)->X1;
      BACK(VLAST)->Y1;
CLOSE;
IF SECPT=UNDEF THEN [%["RESTFRIG",CONSPAIR(X1,Y1)%]] EXIT;
COMMENT 'THIS HAPPENS WITH A REST';

COMMENT 'OTHERWISE';
FRONT(SECPT)->X2;
BACK(SECPT)->Y2;

DIADS(X1,Y1,X2,Y2)->DLIST;COMMENT 'DIAD INTERPRETNS';
IF MTRACE THEN DLIST.REV->DIADPRINT;CLOSE;
IF MTRACE
THEN NL(1); IF DLIST.HD.HD=UNDEF
THEN PR(0) ELSE PR(LENGTH(DLIST)) ;CLOSE;
      SP(1);PR('EXTERNAL DIADS');
CLOSE;
LDIADS:COMMENT 'FOR ALL DIAD INTERPS, PUT THE TWO
      KEY IMPLICATIONS IN THE LIST "IMPS"';
      IF DLIST.NULL THEN GOTO LRUN CLOSE;
      DLIST.HD.DEST.DEST.DEST.DEST.ERASE->Y2->X2->Y1->X1;
      IF X2=UNDEF THEN
COMMENT 'NO DIAD INTERPS FROM AN
      IMPERFECT TRANSITION';
      .NIL EXIT;
      IMPTABLE(X1,Y1,X2,Y2)::IMPS->IMPS;
      TL(DLIST)->DLIST;
      .GOTO LDIADS;

LRUN:IMPS<>RUNIMPLY(TRANSIT,KEY);
COMMENT 'INTERNAL IMPLICATIONS OF RUNS';
END

```

L103 TRIANGLE SUBTEST

```

FUNCTION KEYCOPY KEY;
CONSKEY(
KEYLIST(KEY),
KEYMODE(KEY),
POSTLIST(KEY),
KEYTUNE(KEY).TUNECOPY;,
KEYTRIAD(KEY),
VIRTLAST(KEY).TUNECOPY,
KEYREC(KEY) );
END

```

```

FUNCTION SUBTEST;
COMMENT 'FOR EVERY DIAD INTERP., IF ANY,
      OF A TRANSITION TR WHICH SUPPORTS A KEY X,
      A COPY OF X IS ADDED TO A LIST OF SUPPORTED KEYS,
      IN THE GLOBAL "SAVE", BY SUBCALL OF SUBSUBT';

```

```

VARS
DIADPRINT FSTPT POSITS RJNIMPFLAG LASTART LASTFIN
THISTART THISFIN IMPLL DUPLICAT;

```

```

UNDEF->DIADPRINT;
IF MTRACE THEN NL(3);PR("#");CLOSE;
IF MTRACE THEN STARLINE(40);PR(KEYNAME(X));
PR("#");NL(2);
CLOSE;
IF X.KEYREC.NULL.NOT THEN

```

```

TRANSENDS(HD(KEYREC(X)))->LASTART->LASTFIN;
TRANSENDS(TR)->THISTART->THISFIN;
IF MTRACE THEN NL(1);PR("VIRTUAL");SP(1);
A4PRINT(VIRTLAST(X).REV,12);
NL(2);
CLOSE;
IF MTRACE AND PROPTACE
THEN
PR('TRANS , TYPE ');
PR(TR.TL.HD);PR(',FROM ');SP(2);PR(THISTART);
SP(2);PR("TO");SP(2);PR(THISFIN);
IF ACCFLAG
THEN NL(1);PR('TRANSITION CONTAINS ACCIDENTAL');
TR::KEYREC(X)->KEYREC(X);
TJ->KEYTUNE(X);
NL(1);PR("FAILS");
EXIT;
CLOSE;

```

```

IF NOT(THISFIN=UNDEF)
AND NOT(LASTART=UNDEF)
AND THISTART<=>LASTART
AND NOT(ACCFLAG)
THEN TR::KEYREC(X)->KEYREC(X);
TJ->KEYTUNE(X);
HD(TL(VIRTLAST(X))):VIRTLAST(X)->VIRTLAST(X);
X::SAVE->SAVE;
IF MTRACE

```

```

THEN NL(1);PR('MAINTAIN BY INVERSE');
CLOSE;

```

```

EXIT;
CLOSE;
COMMENT 'THIS PARA MEANS THAT A TRANS WHICH IS THE INVERSE
        OF ITS PREDECESSOR IMPLIES ONLY WHAT THAT PRECEDING
        TRANS IMPLIED.  JULY 5';

```

```

IF TR.TL.HD="RUN"
THEN TR.TL.TL.HD->FSTPT;
FSTPT.TL.TL.HD.LAST.INTPOS->POSITS;
IF NOT(LASTART=UNDEF)
AND POSITS.TL.HD<=>LASTART
AND NOT(ACCFLAG)
THEN 0->RUNIMPFLAG;
KEYCOPY(X)->DUPLICAT;
TR::KEYREC(DUPLICAT)->KEYREC(DUPLICAT);
TU->KEYTUNE(DUPLICAT);
DUPLICAT.VIRTLAST.TL.HD::DUPLICAT.VIRTLAST
->DUPLICAT.VIRTLAST;
DUPLICAT::SAVE->SAVE;
IF MTRACE THEN NL(2);PR('MAINTAIN BY INTERNAL INVERSE IN RUN');
CLOSE;
CLOSE;

```

```

CLOSE;
COMMENT 'THE SAME THING FOR INTERNAL IMPS. OF RUNS';

```

```

COMMENT 'OTHERWISE, THE LIST OF PAIRS OF KEYS
        ASSOCIATED WITH EACH DIAD INTERPRETATION
        OF THE TRANS ARE PUT IN IMPLL';

```

```

IMPLICATS(TR,X)->IMPLL;
IF IMPLL.NULL
THEN TR::KEYREC(X)->KEYREC(X);
TU->KEYTUNE(X);
IF MTRACE AND PROPTTRACE
THEN NL(1);PR('FAIL NO IMPS');
CLOSE;

```

```

EXIT;

```

```

L0: COMMENT 'FOR EACH PAIR OF KEYS IN IMPLL
        I.E. EACH DIAD, CREATE A COPY OF THE
        KEY AND TEST IT AGAINST THE DIAD IN
        QUESTION WITH SUBSORT, WHICH
        PUTS SUPPORTED ONES IN THE GLOBAL
        SAVE';

```

```

IF IMPLL.TL.NULL
THEN SUBSORT(X);
EXIT;
KEYCOPY(X)->DUPLICAT;
SUBSORT(DUPLICAT);
TL(IMPLL)->IMPLL;
GOTO L0;

```

```

E 10

```



```

FUNCTION SUBSUBT KEY;
COMMENT 'TESTS KEY AGAINST A DIAD INTERPRETATION
        OF A TRANSITION TR, IN IMPLL.HD, BY MAINTAIN';
    TR::KEYREC(KEY)->KEYREC(KEY);
    TU.TUNECOPY->KEYTUNE(KEY);
    VARS DIADTOPR;
    UNDEF->DIADTOPR;
    IF MTRACE AND NOT(DIADPRINT=UNDEF)
        THEN HD(DIADPRINT)->DIADTOPR;
            TL(DIADPRINT)->DIADPRINT;
    ELSEIF MTRACE AND PROPTACE THEN NL(1);PR("REST");
    CLOSE;
    IF MAINTAIN(IMPLL.HD,KEY) THEN KEY::SAVE->SAVE;
        IF MTRACE THEN NL(1);
            IF NOT(DIADTOPR=UNDEF) THEN DIADPR(DIADTOPR);
                CLOSE;NL(1);PR("MAINTAINS");
        CLOSE;
    EXIT;
    IF MTRACE AND PROPTACE
        THEN NL(1);
            IF NOT(DIADTOPR=UNDEF)
                THEN DIADPR(DIADTOPR);NL(1); CLOSE;PR("FAILS");
    ELSEIF MTRACE THEN NL(1);PR(0);
    CLOSE;
END

```

5. Examples.

This Section is referred to in Section § 4.6 of Part III.

The output yielded as a trace of its workings by the program follows, for the examples of Fugues 17 of Book I, 4 of Book II, 12 of Book I, the C minor organ fugue, introduced at the beginning of Section § 4.1, and the 'St. Anne' organ fugue.

In each case, the first stage of the process, announced by the messages "PROPOSE" and "TESTKEYS", in which all fourteen keys for which the first note is not accidental are examined, can be ignored, since it largely consists of indications as to why silly hypotheses are silly, and the interesting part of its outcome is available from the next part of the trace.

The entry to the principal iterative function CONFIRM for the subject in question is announced. Every time in the iteration that the main function for testing key hypotheses against the current transitions, TESTKEYS, is entered, during the course of CONFIRM, a message is output. Within TESTKEYS each hypothesis which has survived this far is examined in turn. For each, a message identifying its tonic and minor versus major character is output. There follows on the next line a label "VIRTUAL", which constitutes the past history of the piece, according to the hypothesis being considered, as a sequence of virtual notes. In particular the last of these will form the virtual beginning of the transition currently being studied for its support or otherwise of the key. There then follows a trace of the analysis of that key for the current/

current transition, couched in the following terms.

First a message identifies the type of transition in question as a run or jump, and its real end points, expressed as coordinates in the harmonic space relative to an origin of C, ^{follows} Next, if the transition contains a note which, to that key, is an accidental, then a message to that effect is printed. If the key is accidental, then no further investigation of it is made: the program prints a further message FAIL, and passes on to the next key-hypothesis.

For a key under which interpretation the transition does not involve an accidental, the diad interpretations of the transition as one or more of the thirds, fifths etc. which compose chords, are examined in turn. If there are no such interpretations, when the transition is through an imperfect or otherwise remote interval, then the message "FAIL NO IMPS" is printed, and the program passes to the next hypothesis. Otherwise a message identifying the number of external interpretations of the interval between the end points of the transition is printed, then, in the case of a run only, a message indicates how many internal interpretations there are from the run, and the fact that they follow the former. Having thus identified the origin of all the diad interpretations, a message is output for each one indicating whether it supports the key-hypothesis, or does not. If the hypothesis is supported by that interpretation, then the message "MAINTAIN" is printed. Otherwise, "FAIL" is the message.

When all interpretations have been examined in this way, the program indicates that it has passed on to the next key hypothesis by printing a message identifying the new key, and proceeds as for the last.

When all keys have been examined, the program has a list of key hypotheses/

hypotheses that have been actively maintained. This may be empty. If it is, then the message "NONE CONFIRMED" is printed. If some of the rejected keys none-the-less contained no accidentals, then these are restored at the expense of any that did, and the further message "NON ACC. KEYS RESTORED" is output. If, as happens in Fugue 12 of Book I below, all the keys were accidental, then the further message "ALL KEYS ACC. ALL RESTORED" is printed.

If the end of the tune has been reached, or there is only one key still represented in the resulting list of keys, the program exits from this confirmation stage. Otherwise it iterates the step that has just been described, prefacing the trace again with the name of the main hypothesis-testing function TESTKEYS.

When the program has exited from the confirmation stage, it prints its conclusion, which in general will be a single key, although occasionally the end of a melody may be reached before a unique key has been found, in which case all are printed, according to the following conventions.

It identifies the key(s) that it has found by printing their ~~name~~ name, two pieces of technical information called SIDES and VIRTUAL, which are of no general concern, although the latter has as has been seen before, to do with the history of the piece as a series of virtual notes, a message stating the number of notes that were examined in reaching the decision, a message giving the number of notes there are in the whole melody, and finally, one giving the number of transitions that were examined in arriving at the decision.

This is followed by a print labelled "PARSE", of the parse of the melody up until the decision. This is printed as a sequence of transitions/

transitions. Each is identified by its type - "RUN" or "JUMP" - followed by the type of its start point - "TURN", "INF", "REP", or "NOTE". Each of these is followed by an identification of the intervals that compose it, represented as lists containing in order, the word "INT", the keyboard value of the note with which the interval begins, its duration, the notation of the interval as a pair of x,y coordinates in harmonic space, its interval class, (second third, etc.) represented as an integer from 0 (unison, octave etc.) to 6 (seventh, etc.) and finally the direction of the interval as -1, (descending), 1, (ascending, or 0 neither). This list structure is suitable indented and arranged as to make it clear what is a part of what else.

The end-points of each transition are not printed, except in the case of the last, since they form the beginning of the next transition. In the case of a run the seconds other than those of its first point are printed immediately below those with no special marking other than of INTS. They are, of course, of size 1.

[8 4 2][3 4 3][0 4 3][8 4 2][5 4 3][1 4 3]
 [3 10 3]

***PROPOSE

TESTKEYS***

*****[11 MAJOR]*

1 'EXTERNAL DIADS'
 0

*****[9 MINOR]*

1 'EXTERNAL DIADS'
 0

*****[9 MAJOR]*

1 'EXTERNAL DIADS'
 0

*****[8 MINOR]*

1 'EXTERNAL DIADS'

'DIAD' (0,-1) TO (1,-1)
 MAINTAIN

*****[8 MAJOR]*

1 'EXTERNAL DIADS'

'DIAD' (0,-1) TO (1,-1)
 MAINTAIN

*****[6 MINOR]*

1 'EXTERNAL DIADS'
0

*****[6 MAJOR]*

0 'EXTERNAL DIADS'

*****[5 MINOR]*

1 'EXTERNAL DIADS'
0

*****[4 MAJOR]*

1 'EXTERNAL DIADS'
0

*****[3 MINOR]*

1 'EXTERNAL DIADS'
0

*****[3 MAJOR]*

1 'EXTERNAL DIADS'
0

*****[1 MINOR]*

1 'EXTERNAL DIADS'
0

*****[1 MAJOR]*

1 'EXTERNAL DIADS'
0

*****[0 MINOR]*

1 'EXTERNAL DIADS'
0

CONFIRM**

TESTKEYS***

*****[8 MAJOR]*

VIRTUAL [0 . -1][1 . -1]

'TRANS , TYPE 'JUMP',FROM ' [1 . -1] TO [0 . 0]
1 'EXTERNAL DIADS'

'DIAD' (1,-1) TO (0, 0)
MAINTAIN

*****[8 MINOR]*

VIRTUAL [0 . -1][1 . -1]

'TRANS , TYPE 'JUMP',FROM ' [1 . -1] TO [0 . 0]
'TRANSITION CONTAINS ACCIDENTAL'
FAILS

[8 MAJOR]

SIDES [3 1]

VIRTUAL [0 . -1][1 . -1][0 . 0]

'DECISION AT NOTE NUMBER ' 3

'LENGTH OF TUNE' 7

2 'TRANSITIONS EXAMINED'

PARSE

JUMP

NOTE

[INT 8 4 [[0 . -1] [1 . -1]] 4 1]

JUMP

NOTE

[INT 3 4 [[1 . -1] [0 . 0]] 2 -1]

NOTE

[INT 0 4 [[0 . 0] [0 . -1]] 2 -1]

[1 2 2][0 2 2][1 2 2][3 2 2][1 2 2][3 2 2]
 [8 2 1][10 2 1][0 2 2][1 2 2][3 2 2][4 2 2]
 [6 2 2][4 2 2][6 2 2][3 2 2][8 2 2][6 2 2]
 [4 2 2][8 2 2][6 2 2][4 2 2]

***PROPOSE

TESTKEYS***

*****[11 MINOR]*

1 'EXTERNAL DIADS'
 0

*****[11 MAJOR]*

1 'EXTERNAL DIADS'
 0

*****[10 MINOR]*

1 'EXTERNAL DIADS'
 0

*****[9 MAJOR]*

1 'EXTERNAL DIADS'
 0

*****[8 MINOR]*

1 'EXTERNAL DIADS'
 0

*****[8 MAJOR]*

1 'EXTERNAL DIADS'

J

*****[6 MINOR]*

1 'EXTERNAL DIADS'
0

*****[6 MAJOR]*

1 'EXTERNAL DIADS'
0

*****[5 MINOR]*

1 'EXTERNAL DIADS'
0

*****[4 MAJOR]*

1 'EXTERNAL DIADS'
0

*****[2 MINOR]*

1 'EXTERNAL DIADS'
0

*****[2 MAJOR]*

1 'EXTERNAL DIADS'
0

*****[1 MINOR]*

1 'EXTERNAL DIADS'

'DIAD' (-1,-1) TO (0,-1)
MAINIAIN

*****[1 MAJOR]*

1 'EXTERNAL DIADS'

'DIAD' (-1,-1) TO (0,-1)

MAINTAIN

CONFIRM**

TESTKEYS***

*****[1 MAJOR]*

VIRTUAL [-1 . -1][0 . -1]

'TRANS , TYPE 'JUMP',FROM ' [1 . -1] TO [0 . -1]
1 'EXTERNAL DIADS'

'DIAD' (0,-1) TO (0,-1)
MAINTAIN

*****[1 MINOR]*

VIRTUAL [-1 . -1][0 . -1]

'TRANS , TYPE 'JUMP',FROM ' [1 . -1] TO [0 . -1]
1 'EXTERNAL DIADS'

'DIAD' (0,-1) TO (0,-1)
MAINTAIN

TESTKEYS***

*****[1 MINOR]*

VIRTUAL [-1 . -1][0 . -1][0 . -1]

'TRANS , TYPE 'RUN',FROM ' [0 . -1] TO [-2 . -1]
1 'EXTERNAL DIADS',
' FOLLOWED BY ' 1' INTERVAL DIADS FROM 1ST INT. OF RUN'

'DIAD' (0,-1) TO (-1,-1)
MAINTAIN

'DIAD' (0,-1) TO (-1,-1)
MAINTAIN

*****[1 MAJOR]*

VIRTUAL [-1 . -1][0 . -1][0 . -1]

'TRANS , TYPE 'RUN',FROM ' [0 . -1] TO [-2 . -1]

'TRANSITION CONTAINS ACCIDENTAL'
FAILS

[1 MINOR]
SIDES [1]
VIRTUAL [-1 . -1][0 . -1][0 . -1][-1 . -1]

'DECISION AT NOTE NUMBER' 15
'LENGTH OF TUNE' 22
3 'TRANSITIONS EXAMINED'

PARSE

JUMP

IF

[INT 1 2 [[-1 . -1] [0 . 0]] 1 -1]
[INT 0 2 [[0 . 0] [-1 . -1]] 1 1]
[INT 1 2 [[-1 . -1] [1 . -1]] 1 1]

JUMP

IF

[INT 3 2 [[1 . -1] [-1 . -1]] 1 -1]
[INT 1 2 [[-1 . -1] [1 . -1]] 1 1]
[INT 3 2 [[1 . -1] [0 . -1]] 4 -1]

RJN

NOTE

[INT 8 2 [[0 . -1] [-2 . 0]] 1 1]
[INT 10 2 [[-2 . 0] [0 . 0]] 1 1]
[INT 0 2 [[0 . 0] [-1 . -1]] 1 1]
[INT 1 2 [[-1 . -1] [1 . -1]] 1 1]
[INT 3 2 [[1 . -1] [0 . -2]] 1 1]
[INT 4 2 [[0 . -2] [-2 . -1]] 1 1]

IF

[INT 6 2 [[-2 . -1] [0 . -2]] 1 -1]
[INT 4 2 [[0 . -2] [-2 . -1]] 1 1]
[INT 6 2 [[-2 . -1] [1 . -1]] 2 -1]

I, 12

[0 8 3][1 8 3][0 8 3][11 8 2][4 8 3][5 8 3]
 [10 8 2][9 8 2][8 8 2][7 16 2][5 4 2][5 2 2]
 [7 2 2][8 4 2]

***PROPOSE

TESTKEYS**

*****[10 MINOR]*

4 'EXTERNAL DIADS'

0

'DIAD' (-1, 0) TO (-1,-1)

MAINTAIN

0

0

*****[10 MAJOR]*

4 'EXTERNAL DIADS'

0

0

0

0

*****[9 MINOR]*

0 'EXTERNAL DIADS'

*****[8 MAJOR]*

4 'EXTERNAL DIADS'

0

0

'DIAD' (0, 0) TO (0,-1)

MAINTAIN

0

*****[7 MINOR]*

0 'EXTERNAL DIADS'

*****[7 MAJOR]*

0 'EXTERNAL DIADS'

*****[5 MINOR]*

4 'EXTERNAL DIADS'

0

0

'DIAD' (0, 0) TO (0,-1)
MAINTAIN

'DIAD' (0, 0) TO (-1, 0)
MAINTAIN

*****[5 MAJOR]*

4 'EXTERNAL DIADS'

0.

0

0

0

*****[4 MINOR]*

0 'EXTERNAL DIADS'

*****[3 MAJOR]*

4 'EXTERNAL DIADS'

0

0

0

0

*****[1 MINOR]*

4 'EXTERNAL DIADS'

'DIAD' (0,-1) TO (-1,-1)
MAINTAIN

0

0
0

*****[1 MAJOR]*

4 'EXTERNAL DIADS'

'DIAD' (0,-1) TO (-1,-1)
MAINTAIN

'DIAD' (-1, 0) TO (-1,-1)
MAINTAIN

0
0

*****[0 MINOR]*

4 'EXTERNAL DIADS'

0
0
0
0

*****[0 MAJOR]*

4 'EXTERNAL DIADS'

0
0
0
0

CONFIRM**

TESTKEYS**

*****[1 MAJOR]*

VIRTUAL [-1 . 0][-1 . -1]

'TRANS , TYPE 'RUN',FROM ' [-1 . -1] TO [-3 . -1]
'TRANSITION CONTAINS ACCIDENTAL'
FAILS

*****[1 MAJOR]*

VIRTUAL [0 . -1][-1 . -1]

'TRANS , TYPE 'RUN',FROM ' [-1 . -1] TO [-3 . -1]
 'TRANSITION CONTAINS ACCIDENTAL'
 FAILS

*****[1 MINOR]*

VIRTUAL [0 . -1][-1 . -1]

'TRANS , TYPE 'RUN',FROM ' [-1 . -1] TO [-3 . -1]
 'TRANSITION CONTAINS ACCIDENTAL'
 FAILS

*****[5 MINOR]*

VIRTUAL [0 . 0][-1 . 0]

'TRANS , TYPE 'RUN',FROM ' [-1 . -1] TO [1 . 1]
 'TRANSITION CONTAINS ACCIDENTAL'
 FAILS

*****[5 MINOR]*

VIRTUAL [0 . 0][0 . -1]

'TRANS , TYPE 'RUN',FROM ' [-1 . -1] TO [1 . 1]
 'TRANSITION CONTAINS ACCIDENTAL'
 FAILS

*****[8 MAJOR]*

VIRTUAL [0 . 0][0 . -1]

'TRANS , TYPE 'RUN',FROM ' [-1 . -1] TO [1 . -2]
 'TRANSITION CONTAINS ACCIDENTAL'
 FAILS

*****[10 MINOR]*

VIRTUAL [-1 . 0][-1 . -1]

'TRANS , TYPE 'RUN',FROM ' [-1 . -1] TO [-3 . -1]
 'TRANSITION CONTAINS ACCIDENTAL'
 FAILS
 'NOTE CONFIRMED. ALL KEYS ACC. ALL RESTORED.'

TESTKEYS***

*****[1 MAJOR]*

VIRTUAL [-1 . 0][-1 . -1]UNDEF

'TRANS , TYPE 'JUMP',FROM ' [-3 . -1] TO [-4 . -1]
'TRANSITION CONTAINS ACCIDENTAL'
FAILS

*****[1 MAJOR]*

VIRTUAL [0 . -1][-1 . -1]UNDEF

'TRANS , TYPE 'JUMP',FROM ' [-3 . -1] TO [-4 . -1]
'TRANSITION CONTAINS ACCIDENTAL'
FAILS

*****[1 MINOR]*

VIRTUAL [0 . -1][-1 . -1]UNDEF

'TRANS , TYPE 'JUMP',FROM ' [-3 . -1] TO [0 . -2]
'TRANSITION CONTAINS ACCIDENTAL'
FAILS

*****[5 MINOR]*

VIRTUAL [0 . 0][-1 . 0]UNDEF

'TRANS , TYPE 'JUMP',FROM ' [1 . 1] TO [0 . 1]
'TRANSITION CONTAINS ACCIDENTAL'
FAILS

*****[5 MINOR]*

VIRTUAL [0 . 0][0 . -1]UNDEF

'TRANS , TYPE 'JUMP',FROM ' [1 . 1] TO [0 . 1]
'TRANSITION CONTAINS ACCIDENTAL'
FAILS

*****[8 MAJOR]*

VIRTUAL [0 . 0][0 . -1]UNDEF

'TRANS , TYPE 'JUMP',FROM ' [1 . -2] TO [0 . -2]
'TRANSITION CONTAINS ACCIDENTAL'
FAILS

*****[10 MINOR]*

VIRTUAL [-1 . 0][[-1 . -1]UNDEF

'TRANS , TYPE 'JUMP',FROM ' [-3 . -1] TO [-4 . -1]
'TRANSITION CONTAINS ACCIDENTAL'

FAILS

'NONE CONFIRMD. ALL KEYS ACC. ALL RESTORED.'

TESTKEYS***

*****[1 MAJOR]*

VIRTUAL [-1 . 0][[-1 . -1]UNDEFUNDEF

'TRANS , TYPE 'JUMP',FROM ' [-4 . -1] TO [-1 . 0]
'TRANSITION CONTAINS ACCIDENTAL'

FAILS

*****[1 MAJOR]*

VIRTUAL [0 . -1][[-1 . -1]UNDEFUNDEF

'TRANS , TYPE 'JUMP',FROM ' [-4 . -1] TO [-1 . 0]
'TRANSITION CONTAINS ACCIDENTAL'

FAILS

*****[1 MINOR]*

VIRTUAL [0 . -1][[-1 . -1]UNDEFUNDEF

'TRANS , TYPE 'JUMP',FROM ' [0 . -2] TO [-1 . 0]
'TRANSITION CONTAINS ACCIDENTAL'

FAILS

*****[5 MINOR]*

VIRTUAL [0 . 0][[-1 . 0]UNDEFUNDEF

'TRANS , TYPE 'JUMP',FROM ' [0 . 1] TO [-1 . 0]
4 'EXTERNAL DIADS'

'DIAD' (0, 0) TO (-1, 0)
MAINTAIN

'DIAD' (-1, 1) TO (-1, 0)
FAILS

'DIAD' (0, 1) TO (0, 0)

FAILS

'DIAD' (0, 1) TO (-1, 1)
FAILS

*****[5 MINOR]*

VIRTUAL [0 . 0][0 . -1]UNDEFUNDEF

'TRANS , TYPE 'JUMP',FROM ' [0 . 1] TO [-1 . 0]
4 'EXTERNAL DIADS'

'DIAD' (0, 0) TO (-1, 0)
MAINTAIN

'DIAD' (-1, 1) TO (-1, 0)
FAILS

'DIAD' (0, 1) TO (0, 0)
FAILS

'DIAD' (0, 1) TO (-1, 1)
FAILS

*****[8 MAJOR]*

VIRTUAL [0 . 0][0 . -1]UNDEFUNDEF

'TRANS , TYPE 'JUMP',FROM ' [0 . -2] TO [-1 . 0]
'TRANSITION CONTAINS ACCIDENTAL'
FAILS

*****[10 MINOR]*

VIRTUAL [-1 . 0][-1 . -1]UNDEFUNDEF

'TRANS , TYPE 'JUMP',FROM ' [-4 . -1] TO [-1 . 0]
'TRANSITION CONTAINS ACCIDENTAL'
FAILS

[5 MINOR]

SIDES [1 2]

VIRTUAL [0 . 0][0 . -1]UNDEF[0 . 0][-1 . 0]

'DECISION AT NOTE NUMBER ' 6
'LENGTH OF TUNE' 14
4 'TRANSITIONS EXAMINED'

PARSE

JUMP

NOTE
[INT 0 3 [[0 . 0] [-1 . -1]] 1 1]
RUN
NOTE
[INT 1 3 [[-1 . -1] [0 . 0]] 1 -1]
[INT 0 8 [[0 . 0] [1 . 1]] 1 -1]
JUMP
NOTE
[INT 11 8 [[1 . 1] [0 . 1]] 3 1]
JUMP
NOTE
[INT 4 8 [[0 . 1] [-1 . 0]] 1 1]
NOTE
[INT 5 8 [[-1 . 0] [-2 . 0]] 4 -1]

C minor Organ Fugue 25

[8 2 1][7 2 1][5 2 1][3 2 1][5 2 1][3 2 1]
 [2 2 1][0 2 1][8 2 1][0 2 1][8 2 1][0 2 1]

***PROPOSE

TESTKEYS***

*****[11 MAJOR]*

0 'EXTERNAL DIADS'

*****[9 MINOR]*

0 'EXTERNAL DIADS'

*****[9 MAJOR]*

0 'EXTERNAL DIADS'

*****[8 MINOR]*

1 'EXTERNAL DIADS',
 ' FOLLOWED BY ' 4' INTERNAL DIADS FROM 1ST INT. OF RUN'
 0
 0
 0
 0
 0

*****[8 MAJOR]*

1 'EXTERNAL DIADS',
 ' FOLLOWED BY ' 4' INTERNAL DIADS FROM 1ST INT. OF RUN'
 'DIAD' (0,-1) TO (1,-1)
 MAINTAIN
 0
 0

'DIAD' (0,-1) TO (0, 0)
 MAINTAIN

'DIAD' (0,-1) TO (1,-1)
 MAINTAIN

*****[6 MINOR]*

0 'EXTERNAL DIADS'

*****[6 MAJOR]*

0 'EXTERNAL DIADS'

*****[5 MINOR]*

0 'EXTERNAL DIADS'

*****[4 MAJOR]*

0 'EXTERNAL DIADS'

*****[3 MINOR]*

1 'EXTERNAL DIADS',
 ' FOLLOWED BY ' 4' INTERNAL DIADS FROM 1ST INT. OF RUN'
 0
 0
 0
 0
 0

*****[3 MAJOR]*

1 'EXTERNAL DIADS',
 ' FOLLOWED BY ' 4' INTERNAL DIADS FROM 1ST INT. OF RUN'
 0
 0

'DIAD' (1,-1) TO (1, 0)
 MAINTAIN
 0
 0

*****[1 MINOR]*

```

1 'EXTERNAL DIADS',
' FOLLOWED BY ' 4' INTERNAL DIADS FROM 1ST INT. OF RUN'
0
0
0
0
0

```

```
*****[ 1 MAJOR]*
```

```

1 'EXTERNAL DIADS',
' FOLLOWED BY ' 4' INTERNAL DIADS FROM 1ST INT. OF RUN'
0
0
0
0
0

```

```
*****[ 0 MINOR]*
```

```

1 'EXTERNAL DIADS',
' FOLLOWED BY ' 4' INTERNAL DIADS FROM 1ST INT. OF RUN'
0

```

```

'DIAD' ( 0, 0) TO ( 1, 0)
MAINTAIN

```

```

'DIAD' ( 1,-1) TO ( 1, 0)
MAINTAIN
0
0

```

```
***CONFIRM*****
```

```
***TESTKEYS*****
```

```
*****[ 0 MINOR]*
```

```
VIRTUAL [ 1 . -1][ 1 . 0]
```

```

'TRANS , TYPE 'JUMP',FROM ' [ 1 . -1] TO [-1 . 0]
1 'EXTERNAL DIADS'

```

```

'DIAD' ( 1, 0) TO ( 0, 0)
MAINTAIN

```

```
*****[ 0 MINOR]*
```

```
VIRTUAL [ 0 . 0][ 1 . 0]
```

'TRANS , TYPE 'JUMP',FROM ' [1 . -1] TO [-1 . 0]
1 'EXTERNAL DIADS'

'DIAD' (1, 0) TO (0, 0)
MAINTAIN

*****[3 MAJOR]*

VIRTUAL [1 . -1][1 . 0]

'TRANS , TYPE 'JUMP',FROM ' [1 . -1] TO [3 . -1]
3 'EXTERNAL DIADS'

'DIAD' (1,-1) TO (2,-1)
MAINTAIN

'DIAD' (1, 0) TO (2, 0)
FAILS

'DIAD' (2,-1) TO (3,-1)
FAILS

*****[8 MAJOR]*

VIRTUAL [0 . -1][1 . -1]

'TRANS , TYPE 'JUMP',FROM ' [1 . -1] TO [-1 . 0]
1 'EXTERNAL DIADS'

'DIAD' (1,-1) TO (0,-1)
MAINTAIN

*****[8 MAJOR]*

VIRTUAL [0 . -1][0 . 0]

'TRANS , TYPE 'JUMP',FROM ' [1 . -1] TO [-1 . 0]
1 'EXTERNAL DIADS'

'DIAD' (0, 0) TO (-1, 0)
FAILS

*****[8 MAJOR]*

VIRTUAL [0 . -1][1 . -1]

'TRANS , TYPE 'JUMP',FROM ' [1 . -1] TO [-1 . 0]
1 'EXTERNAL DIADS'

'DIAD' (1,-1) TO (0,-1)
MAINTAIN

TESTKEYS**

*****[8 MAJOR]*

VIRTUAL [0 . -1][1 . -1][0 . -1]

'TRANS , TYPE 'RUN',FROM ' [-1 . 0] TO [0 . 0]
'TRANSITION CONTAINS ACCIDENTAL'
FAILS

*****[8 MAJOR]*

VIRTUAL [0 . -1][1 . -1][0 . -1]

'TRANS , TYPE 'RUN',FROM ' [-1 . 0] TO [0 . 0]
'TRANSITION CONTAINS ACCIDENTAL'
FAILS

*****[3 MAJOR]*

VIRTUAL [1 . -1][1 . -1][2 . -1]

'TRANS , TYPE 'RUN',FROM ' [3 . -1] TO [0 . 0]

'MAINTAIN BY INTERNAL INVERSE IN RUN'
1 'EXTERNAL DIADS'

'DIAD' (2,-1) TO (1,-1)
MAINTAIN

*****[0 MINOR]*

VIRTUAL [0 . 0][1 . 0][0 . 0]

'TRANS , TYPE 'RUN',FROM ' [-1 . 0] TO [0 . 0]

'MAINTAIN BY INTERNAL INVERSE IN RUN'
1 'EXTERNAL DIADS'

'DIAD' (0, 0) TO (0, 0)
MAINTAIN

*****[0 MINOR]*

VIRTUAL [1 . -1][1 . 0][0 . 0]

'TRANS , TYPE 'RUN',FROM ' [-1 . 0] TO [0 . 0]

'MAINTAIN BY INTERNAL INVERSE IN RUN'
1 'EXTERNAL DIADS'

'DIAD' (0, 0) TO (0, 0)
MAINTAIN

TESTKEYS***

*****[0 MINOR]*

VIRTUAL [1 . -1][1 . 0][0 . 0][0 . 0]

'TRANS , TYPE 'JUMP',FROM ' [0 . 0] TO [0 . -1]
1 'EXTERNAL DIADS'

'DIAD' (0, 0) TO (0,-1)
FAILS

*****[0 MINOR]*

VIRTUAL [1 . -1][1 . 0][0 . 0][1 . 0]

'TRANS , TYPE 'JUMP',FROM ' [0 . 0] TO [0 . -1]
4 'EXTERNAL DIADS'

'DIAD' (1,-1) TO (0,-1)
FAILS

'DIAD' (0, 0) TO (0,-1)
FAILS

'DIAD' (1, 0) TO (1,-1)
MAINTAIN

'DIAD' (1, 0) TO (0, 0)
MAINTAIN

*****[0 MINOR]*

VIRTUAL [0 . 0][1 . 0][0 . 0][0 . 0]

'TRANS , TYPE 'JUMP',FROM ' [0 . 0] TO [0 . -1]
1 'EXTERNAL DIADS'

'DIAD' (0, 0) TO (0,-1)
FAILS

*****[0 MINOR]*

VIRTUAL [0 . 0][1 . 0][0 . 0][1 . 0]

'TRANS , TYPE 'JUMP',FROM ' [0 . 0] TO [0 . -1]
4 'EXTERNAL DIADS'

'DIAD' (1,-1) TO (0,-1)
FAILS

'DIAD' (0, 0) TO (0,-1)
FAILS

'DIAD' (1, 0) TO (1,-1)
MAINTAIN

'DIAD' (1, 0) TO (0,0)
MAINTAIN

*****[3 MAJOR]*

VIRTUAL [1 . -1][1 . -1][2 . -1][1 . -1]

'TRANS , TYPE 'JUMP',FROM ' [0 . 0] TO [0 . -1]
1 'EXTERNAL DIADS'

'DIAD' (1,-1) TO (0,-1)
FAILS

*****[3 MAJOR]*

VIRTUAL [1 . -1][1 . -1][2 . -1][1 . -1]

'TRANS , TYPE 'JUMP',FROM ' [0 . 0] TO [0 . -1]
1 'EXTERNAL DIADS'

'DIAD' (1,-1) TO (0,-1)
FAILS

[0 MINOR]

SIDES [1]

VIRTUAL [0 . 0][1 . 0][0 . 0][1 . 0][0 . 0]

'DECISION AT NOTE NUMBER ' 9
'LENGTH OF TUNE' 12
4 'TRANSITIONS EXAMINED'

PARSE

RUN

NOTE

[INT 8 2 [[0 . -1] [1 . 0]] 1 -1]

[INT 7 2 [[1 . 0] [-1 . 0]] 1 -1]

[INT 5 2 [[-1 . 0] [1 . -1]] 1 -1]

JUMP

```
NOTE
  [INT 3 2 [[ 1 . -1] [-1 . 0]] 1 1]
RUN
NOTE
  [INT 5 2 [[-1 . 0] [ 1 . -1]] 1 -1]
  [INT 3 2 [[ 1 . -1] [ 2 . 0]] 1 -1]
  [INT 2 2 [[ 2 . 0] [ 0 . 0]] 1 -1]
JUMP
NOTE
  [INT 0 2 [[ 0 . 0] [ 0 . -1]] 5 1]
NOTE
  [INT 8 2 [[ 0 . -1] [ 0 . 0]] 5 -1]
```

St Anne Organ Fugues
(transcribed in C)

[7 32 1][4 16 1][9 16 1][7 16 1][0 32 2][11 16 1]
[0 16 2]

***PROPOSE

TESTKEYS***

*****[11 MINOR]*

1 'EXTERNAL DIADS'
0

*****[10 MAJOR]*

1 'EXTERNAL DIADS'
0

*****[8 MINOR]*

0 'EXTERNAL DIADS'

*****[8 MAJOR]*

1 'EXTERNAL DIADS'
0

*****[7 MINOR]*

1 'EXTERNAL DIADS'
0

*****[7 MAJOR]*

1 'EXTERNAL DIADS'
0

*****[5 MINOR]*

1 'EXTERNAL DIADS'
0

*****[5 MAJOR]*

1 'EXTERNAL DIADS'
0

*****[4 MINOR]*

1 'EXTERNAL DIADS'
'DIAD' (1, 0) TO (0, 1)
MAINTAIN

*****[3 MAJOR]*

1 'EXTERNAL DIADS'
0

*****[2 MINOR]*

0 'EXTERNAL DIADS'

*****[2 MAJOR]*

0 'EXTERNAL DIADS'

*****[0 MINOR]*

1 'EXTERNAL DIADS'
0

*****[0 MAJOR]*

1 'EXTERNAL DIADS'
'DIAD' (1, 0) TO (0, 1)
MAINTAIN

CONFIRM**

TESTKEYS***

*****[0 MAJOR]*

VIRTUAL [1 . 0][0 . 1]

'TRANS , TYPE 'JUMP',FROM ' [0 . 1] TO [-1 . 1]
1 'EXTERNAL DIADS'

'DIAD' (0, 1) TO (-1, 1)
FAILS

*****[4 MINOR]*

VIRTUAL [1 . 0][0 . 1]

'TRANS , TYPE 'JUMP',FROM ' [0 . 1] TO [-1 . 1]
1 'EXTERNAL DIADS'

'DIAD' (0, 1) TO (-1, 1)
FAILS

'NONE CONFIRMD. NON-ACC. KEYS RESTORED.

TESTKEYS***

*****[4 MINOR]*

VIRTUAL [1 . 0][0 . 1]UNDEF

'TRANS , TYPE 'JUMP',FROM ' [-1 . 1] TO [1 . 0]
3 'EXTERNAL DIADS'

'DIAD' (-1, 0) TO (0, 0)
FAILS

'DIAD' (-1, 1) TO (0, 1)
FAILS

'DIAD' (0, 0) TO (1, 0)
FAILS

*****[0 MAJOR]*

VIRTUAL [1 . 0][0 . 1]UNDEF

'TRANS , TYPE 'JUMP',FROM ' [-1 . 1] TO [1 . 0]
3 'EXTERNAL DIADS'

'DIAD' (-1, 0) TO (0, 0)
FAILS

'DIAD' (-1, 1) TO (0, 1)
FAILS

'DIAD' (0, 0) TO (1, 0)
MAINTAIN

[0 MAJOR]

SIDES [1 3]

VIRTUAL [1 . 0][0 . 1][0 . 0][1 . 0]

'DECISION AT NOTE NUMBER ' 4

'LENGTH OF TUNE' 7

3 'TRANSITIONS EXAMINED'

PARSE

JUMP

NOTE

[INT 7 32 [[1 . 0][0 . 1]] 2 -1]

JUMP

NOTE

[INT 4 16 [[0 . 1][-1 . 1]] 3 1]

JUMP

NOTE

[INT 9 16 [[-1 . 1][1 . 0]] 1 -1]

NOTE

[INT 7 16 [[1 . 0][0 . 0]] 3 1]

Appendix V

Musical Terms

The musical terms that are used in the body of the work without explicit definition are here described, and a couple of very minor generalisations to orthodox use pointed out.

The precise nature of the concept of key is one of the main concerns of Part III. However, as far as mere terminology goes, a key is associated with a scale, or ordered set of seven pitches, or degrees that are of special significance for pieces of music which are in that key. These seven pitches are in a determined relation to each other, and in particular to one of their number, the key-note, or tonic. The scale divides the interval between the tonic and its octave, the note of twice its frequency, into seven non-equal parts, the intervals of a tone or a semitone between each note of the scale and its successor. We show in the text of Part III that not all intervals of a tone are the same, but for the moment it will do to think of tones as approximately equal to two semitones, as the names suggest. In particular, in Bach's time there were two such orderings of tones and semitones used, called the major and minor scales. The major scale is the sequence of pitches defined by the sequence tone, tone, semitone, tone, tone, tone, semitone, while the minor scale is defined by the sequence tone, semitone, tone, tone, semitone, tone, semitone, although in the text it is shown how the minor scale is a little more complicated than this.

The notes of a scale can be referred to in several ways. One way is by ordinal position in the scale, where the tonic is the first of the scale, and the others are the second, and so on up to the seventh and eighth, or octave. By extension, the notes in the version of the same scale beginning on the octave can either be referred to as the ninth tenth and so on, or as the octave second, octave third and so on. Because this set of names would allow of such confusing phrases as "the fifth fifth in the piece", there is an alternative set of names for the notes of a scale in relation to the first. One of this set has already been used, in referring to the first, or key-note as the tonic. There are similarly names for each one of the seven degrees of/
of/

/of the scale. They are, in order, after the tonic, the supertonic, mediant, subdominant, dominant, submediant, and leading-note, respectively, and we might add the octave in as the eighth. In particular the dominant, subdominant, and tonic, are frequently referred to by these names in the text.

The degrees of a scale may also be referred to by note-names. In this case the nomenclature is not relative to the tonic, but is in some sense absolute. The system is based on the assignment of the name C to a certain frequency. The other notes of the major scale which has C as tonic are then given the names D, E, F, G, A, and B. Each of these can be a tonic of its own scale. However, on occasion, these will include notes which have not so far been named. For example, the major scale of D starts with a tone, which takes us to the already-named E. The next interval is also a tone, which takes us to a note intermediate between F, a semitone away, and G, a semitone plus a tone away, for which as yet we have no name. Such notes are named in relation to the degrees of C major. They may be referred to either as the flattened degree above, or as the sharpened degree below. This ambiguity of nomenclature is only apparent, but for present purposes it can be resolved by the following rule of thumb, which will be explained later. For a given key-note name, each successive degree is named with the succeeding note-letter in the alphabet of C major, together with an indication of sharpening or flattening sufficient to satisfy the demands of the earlier definition of a scale as a sequence of tones and semitones. In the earlier example of D major, the third note by this rule must have to be some kind of F, and to be a tone above E, it must be a sharpened F, written F $\sharp$. The alternative "spelling" of the note as a flattened G, written Gb, would be wrong in this case. However, in the case of the key of Eb minor, as a spelling of the third degree of the scale, Gb would be correct, and F $\sharp$ would be wrong. The significance of these distinctions is dealt with in the body of the work, in Part III. For the present purpose the distinction must just be asserted.

Any note defined in this way may be the tonic of its own scale. On occasion, for a rather exotic key, or for an accidental, (a note of a melody/

/melody not in the scale of its key) it may be necessary to name a note by more than one flattening or sharpening step of a semitone. The third degree or mediant of the scale of ~~A#~~-major must be some kind of C, but to be a tone above the supertone B#, it is not enough for it to be raised by one sharp. It is in fact raised by two semitones, and referred to as C double-sharp, written C^X. (The corresponding double-flat is simply written as Ebb.) Accidentals may receive any one of these spellings, which again are seen in the text to reflect real musical distinctions.

Traditional musical notation is a means of writing down pieces of music in a key, made up to notes identified in this system. The conventions of this notation are assumed. In the body of the work the term notation is used to refer to the drawing of these distinctions as to the name and function of notes. However, one concern of Part III is to show that there are similar distinctions of function that are ignored in the traditional notation. A slightly non-standard use of musicological terminology involved in the text is that the term notation is generalised to include the drawing of these further distinctions

The modern keyboard, by blurring the distinctions that have been hinted at, equated all of these "spellings", and this is the last system by which notes can be referred to, by identifying them with the notes of the keyboard that would be used to play them. The transformation by which the distinctions are compressed into the twelve semitone degrees of the piano keyboard is called the system of Equal Temperament, and its precise nature is also the concern of Part III. When people hear music, and perform the tasks that are modelled in the body of this work, the music is generally played in the intonation of Equal Temperament. When therefore, it is necessary to make it clear that some group of notes is to be understood as only defined to within the ambiguity of Equal Temperament, then those pitches will be identified with the numbers from 0 (for the keyboard note C, B#, etc.) to ii, (for B, Cb etc.), since the usual notation would imply a false definiteness.

The intervals between any two notes in a scale are identified by the number of scale steps that must be taken to get from one to the other, plus one. Thus, the interval from a tonic to its supertonic or the version of the leading note just below it is a second, descending in the latter case. There is one exception to this rule. The interval between a pitch and itself is not called a first, but a unison. As with the degrees of the scale, the rule is not limited to the compass of a single octave: such intervals as the ninth will occur.

These interval-names are further qualified by adjectives identifying both their actual size in semitones, (e.g. the minor and major thirds, or the tone and semitone already mentioned,) and their function. (So incidentally are the note-names above.) The same applies when they involve accidentals. However the question of function does not belong in a section on terminology, and the traditional nomenclature is very unsystematic. It will not in general be used without explanation, except in the cases of the tone and semitone and of the prefixes minor and major. These will be used to refer, as in the example above, to the intervals of the same size and nature as those made by a tonic and the degree in question in a minor and major scale respectively. Thus the interval between E and G in the scale of C major is a minor third, because it is of the same size and nature as the interval from C to E $\flat$ in the scale of C minor.

When several notes occur simultaneously, the ensemble is referred to as a chord. When the notes of a chord follow one another in sequence it is called an arpeggio. The nomenclature of chords in general is a subject which is, again, far beyond the scope of an appendix on terminology, and again is taken up in Part III. However certain terms can be defined here.

One note of a chord is of paramount importance in the definition of the function of the chord. This is called the root. The complex nomenclature for chords involves qualifying adjectives applied to the root of the chord, as for example in the chords of C major and C minor.
The/

The definitions of these two chords, which consist of the tonic, mediant, and dominant of the relevant scales, will be assumed. The convention that a chord which includes all of these, together with some other degree of the scale, can be referred to as e.g. C major ninth, will also occasionally be used.

The notes of, for example, a major chord, may occur at any octave, or even in several, and in any order. The chord is still referred to as the major chord. These different forms are termed the inversions of the chord. This is a slightly more general use of the term than is orthodox: usually the process of inversion of a chord would not include the addition of extra notes in additional octaves but only the rearrangement of the same number of notes. Since we shall not in general be concerned with the differences between various inversions, nor very often with the difference between octaves, we can use the term in this way. For similar reasons the term triad which traditionally means a chord made up of a single octave occurrence of each of three notes of a chord, will be used for any chord which contains only three note-names, even if these occur in many octaves.

Chords, and single intervals, and by extension larger fragments of music, may receive the same descriptors although they involve quite different pitches, as we saw for the minor third, and the major and minor chords. Such groups of notes are said to be related by a transposition of some interval. This is to say that the one could be produced by moving every note of the other by the same interval. On occasion the term "the same interval" will be taken to mean the same kind of interval, i.e. some notes may be transposed by the minor, and some by the major, version of an interval.

The term metre is used in the sense in which Cooper and Meyer (1960), among others, use it, to mean the regular framework of accent and unaccent against which the rhythm or local relations of note-durations are to be understood, as for example, being syncopations (non-coincidence of rhythmic and metrical accent) or not. The metre of a piece is written in the traditional/

traditional notation with a time signature, bar-lines and groupings of notes within bars. As before, familiarity with this notation is assumed.

Within a metric group, the first beat is assumed to be accented, and is referred to as the downbeat. All other beats are referred to as upbeats.

There are slightly different conventions for naming note-durations between European and American musicology. The note durations which to an American musician are known as a whole-note, a half-note, a quarter-note, an eighth-note, a sixteenth-note, and a thirty-second-note, are to an English musician a semibreve, a minim, a crotchet, a quaver, a semiquaver, and a demisemiquaver, respectively.

References

- BURSTALL, R.M. et al. (1971) Programming in POP-2. Edinburgh:
Edinburgh Univ. Press.
- CHOMSKY, N. and HALLE, M. (1968) Sound Patterns of English. New York:
Harper and Row.
- COOKE, D. (1959) The Language of Music. London: Oxford Univ. Press.
- COOPER, G. and MEYER, L.B. (1960) The Rhythmic Structure of Music.
Chicago: Univ. Chicago Press.
- ELLIS, A.J. (1874) On Musical Duodenes. Proc. Roy. Soc. 23, 3.
- ELLIS, A.J. (1875) Appendix XX to 1st ed. trans. of Helmholtz (1862)
q.v. London: Longmans.
- ELLIS, A.J. (1885) Ibid. 2nd ed. expanded version.
- FORTE, A. (1967). Syntax-Based Analytic Reading of Musical Scores.
Mac-Tr-39, M.I.T.
- GOMBRICH, E.H. (1960). Art and Illusion. London: Phaedon.
- GOMBRICH, E.H. (1963) Meditations on a Hobby-horse. London: Phaedon.
- GOMBRICH, E.H. (1972) The Visual Image. Scientific American, 227,
September, 82.
- HELMHOLTZ, H.L.F. (1862) Die Lehre Von Den Tonempfindungen. Brunswick:
Vieweg. Trans. Ellis, A.J. (1875) (q.v.) as The Sensations of Tone.
London: Longmans.
- GROVE, Sir C. (1940) Grove's Musical Dictionary (4th ed.) London:
Macmillan.

HIGGS, J. () Fugue. London: Novello.

HINDEMITH, P. (1937) (Trans. 1942). The Craft of Musical Composition.
Part I. London: Schott.

LINDBLOM, B. and SUNDBERG, J. (1970) Towards a Generative Theory of
Melody. Quart. Progress Report, January. Speech Transmission
Lab., K.T.H. Stockholm.

LONGUET-HIGGINS, H.C. (1962a) Letter to a Musical Friend. Mus. Rev.
23, August, 244.

LONGUET-HIGGINS, H.C. (1962b) Second letter to a Musical Friend.
Mus. Rev. 23, Nov., 280.

LONGUET-HIGGINS, H.C. (1965) The Three Dimensions of Harmony.
Cambridge Research, Oct.

LONGUET-HIGGINS, H.C. and STEEDMAN, M.J. (1971) On Interpreting Bach.
in Machine Intelligence 6 (eds. Meltzer, B. and Michie, D.)
Edinburgh: Edinburgh Univ. Press.

LONGUET-HIGGINS, H.C. (1972) Making Sense of Music. Proc. Roy. Inst.
(in press).

MEYER, L.B. (1956) Emotion and Meaning in Music. Chicago: Univ. of
Chicago Press.

MORLEY, T. (1957). A Plain and Easy Introduction to Practical Music.
(Ed. Harman, R.A. (1952) London: Dent.

SIMON, H.A. and SUMNER, R.K. (1968) Pattern in Music. in Formal
Representation of Human Judgement. (Ed. Kleinmuntz, B.)
New York: Wiley.

SMOLIAR, S. (1971) A Parallel Processing Model of Musical Structure.

Mac-Tr-91. M.I.T.

THORNE, J.P. (1970) Generative Grammar and Stylistic Analysis. In

New Horizons in Linguistics (Ed. Lyons, J.) London: Penguin.

TOVEY, D.F. and SAMUEL, H. (Eds.) (1924) Bach Forty-Eight Preludes

and Fugues. London: Associated Board.

WEBER, G. (1817) (Trans. 1851) Theory of Musical Composition.

London: Cocks.

WINOGRAD, T. (1968) Linguistics and the Computer Analysis of Harmony.

Jou. Mus. Theory, 12, 2.

Addenda.

BAMBERGER, J. (1972) Untitled Report on Teaching Music. A.I. Lab., M.I.T.

SCHOENBERG, A. (1954) Structural Functions of Harmony. London :

Williams & Norgate.
