Bernd Härpfer

The Metric Malleability and Ambiguity of Cyclic Rhythms
A Quantitative Heuristic Approach

wolke verlag, 2023
www.wolke-verlag.de

UNIVERSITY OF MUSIC KARLSRUHE

DOCTORAL THESIS

The Metric Malleability and Ambiguity of Cyclic Rhythms
A Quantitative Heuristic Approach

Author:
Bernd HÄRPFER

Supervisors:
Prof. Thomas TROGE
Prof. Denis LORRAIN

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Date

September 2018, revised in October 2020

UNIVERSITY OF MUSIC KARLSRUHE

Abstract

Institute for Musicology and Music Informatics (IMWI)

Doctor of Philosophy

The Metric Malleability and Ambiguity of Cyclic Rhythms. A Quantitative Heuristic Approach

Bernd HÄRPFER

Metric malleability relates to the cognitive flexibility which allows for the construction of different metric frameworks for a particular rhythm. It is also closely related to metric ambiguity, indicating the space of plausible metric interpretations for a given rhythm. Different listeners may construct individual metric frameworks for the same rhythm (metric ambiguity), and different interpretations may compete within individual cognitive processes (metric dissonance or metric conflict). Such conflicts may be purposefully provoked by appropriate musical designs.

The present study examines the perceptual and cognitive foundations of these phenomena, and establishes heuristic and formal categories, concerning metric ambiguity, as a basis for a model of metric malleability. Given the complexity of the matter, features of rhythm and meter are reduced to a simple formal level, using a pulse- and cycle-based approach. The model proposes a quantitative heuristics for an individual characterization of the metric malleability of cyclic rhythms, which is exemplarily demonstrated. Thereby, it provides a conceptual basis, as well as implementable tools for generative and analytical applications, intending to address a broader artistic and scientific audience.

Meanwhile the model has become part of an ongoing software project, the development of a creative tool for rhythm design focusing on metric malleability. For updated information on this project and for all other issues, please contact bernd[at]haerpfer.net or check www.haerpfer.net.

Acknowledgements

I would like to thank my supervisors Thomas Troge and Denis Lorrain for their guidance and generous help over the years of my doctoral research.

This work would not have been possible without the support and encouragement of all my teachers, colleagues and friends, who conveyed to me their knowledge and their skills. I thank Clarence Barlow for his inspiring teaching on algorithmic composition, and for encouraging me to do quantitative research on related issues. Many people helped me with their input in discussions about my topic and more generally, about researching and writing a thesis. In this respect, I specially thank Yasmine Chehata, Siegfried Koepf and Christoph Seibert.

Some final improvements were supported by helpful comments from Justin London who also provided me some of his original graphics, and by Kai Werth and Albert Gräf from University of Mainz, who helped me to ensure consistency in mathematical terms.

I am grateful about the support of my friends and relatives, making it possible to keep to the work. I thank Don Childs and Michael Fuchs for special courtesy. Most fundamentally, my partner Sonja Werner assisted me in manifold ways, and with enormous patience. All my thank and love to her, to my parents, and to my daughters Kyra and Thea.

Contents

To whom it may concern[1]

[1]Tomkins, 2013

Chapter 1

Introduction

> Music is a creative enterprise whose goals and methods are not fixed.[1]

The statement above opens HURON's discussion of musical strategies playing on perceptual and cognitive preconditions of listeners. From this perspective, emotional and cognitive schemes and expectations can be enhanced or challenged by musical means to create predictability, surprise, tension and so forth. The systematic approach pursued in this thesis is similarly motivated by artistic curiosity: how to creatively exploit ambiguity in rhythmic structures concerning listener's expectations and interpretation practices. From such an artistic research perspective, appropriate concepts and models developed in the context of music theory and cognitive sciences can be valuably implemented in artistic methods. Hence, science and art do inform each other (and have always done), as

> artistic methods can be found ubiquitously within science, particularly where surprising new discoveries are made. [...] All scientific disciplines stem from the "arts" (*technai* in ancient Greece).[2]

The word "music" in the entrance statement could thus be exchanged with "art" and just as well with "science". FEYERABEND amplifies that in the ancient arts, the development of conceptions and the recognition of technical problems were not separable from practical learning processes and specific problem solving activities. The claim of generalization was therefore not the aim of the arts. It was developed later in the sciences as a theory and a demand for a generic logic in scientific research.[3] I think that from a current point of view, artistic research can be specifically effective but can also lead to broader applicability and generalization in a scientific sense if the artistic method is formalized.

[1]Huron, 2006, p. 239
[2]Feyerabend, 1984, p. 8 (my translation)
[3]ibid.

The research method of this study correspondingly aims at a quantified axiomatic system, a model which is likewise theoretically valid and implementable.[4] Due to its essential complexity, investigating musical ambiguity[5] demands a thorough definition of its conditions. Nevertheless, we only can expect a heuristic approach to be adequate. In other words, the model should be kept open to be adaptable to a variety of applications.

1.1 Main topics and hypotheses

Figure 1.1 illustrates the type of musical ambiguity which is of main interest in this study. A simple rhythmic cycle, supposed to be played repeatedly, is aligned with different metric frameworks. Any of the twelve variants yields a plausible metric interpretation of the rhythm, even if the cycle starts at an arbitrary temporal position (which may lead to *anacruses* with different lengths).

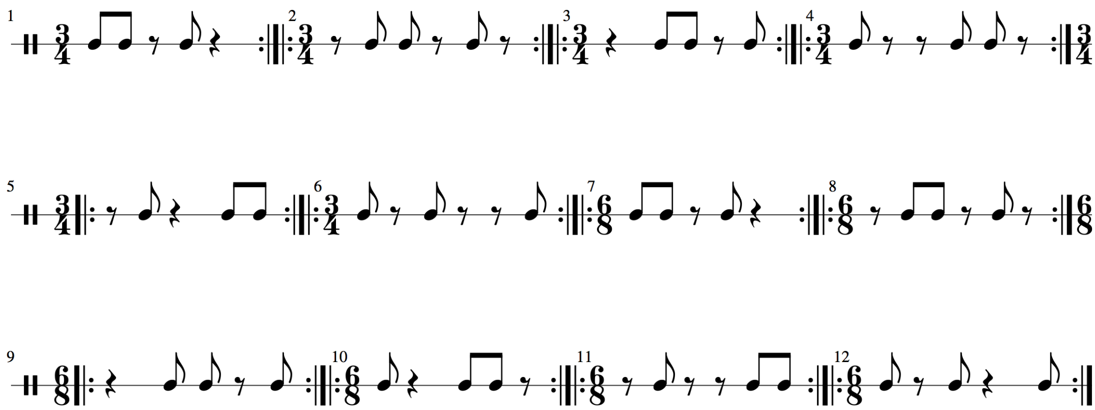


FIGURE 1.1: Many possible metric frameworks for the same rhythmic cycle

The rhythmic cycle itself can be notated without metric context as a cycle of intervals between rhythmic onsets: [1-2-3], or as a regular pattern of onsets and "rests": [xx.x..]. The former displays the inter-onset intervals by integer multiples of a smallest temporal unit (*interval notation*).[6] The latter is known as a variant of *box notation*[7] and suggests a categorical grid of such units, either filled with onsets ("x") or rests, respectively continuations of the onsets (".").[8] During this study, these two notation formats are used

[4] For instance, in applications for algorithmic composition.

[5] Specifically, metric ambiguity is investigated in this study.

[6] In case of figure 1.1 this unit is equivalent to the duration of an eighth-note or -rest.

[7] cf. for instance Hall and Klingsberg, 2004, Sethares, 2007, and Toussaint, 2013

[8] In contrast to traditional notation, these representations do not include information about sound sustain or distinguish note- and rest values. This may be regarded as irrelevant for studying sequences of inter-onset intervals.

along with a third variant, exemplified in figure 1.2. This so-called *necklace notation*[9] pronounces the circular structure of the rhythm, which is indicated by means of square brackets in the other formats. As discussed in chapter 4, it clearly represents metric and geometric features of cyclic rhythms. It is also particularly suitable to illustrate *metric malleability*, the main topic of this thesis, as demonstrated in section 5.3.

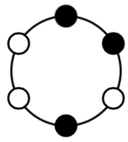


FIGURE 1.2: The pattern [1-2-3] or [xx.x..], displayed as a series of beads arranged "clockwise" on a "necklace" (black beads stand for onsets and white beads represent rests or continuations).

The metric alternatives in figure 1.1 arise from two types of variation. First, the location of the metric *downbeat* (the "one"), so that it coincides with any of the smallest temporal units, creates a *metric rotation* or *metric phase* of the (continuous) rhythmic cycle (see section 4.1.3). Second, the two metric types 3/4 and 6/8 involve different *metrical accent hierarchies* which are imposed on the rhythm.[10] The first aspect implies that six different units of the cycle could be interpreted as the downbeat position at the start of each bar, either coinciding with a note onset or with a rest. The latter is also known as a *loud rest*[11] or a *virtual beat* or *articulation*.[12] Rests on other strong metric positions may also cause similar feelings, which are often described in terms of syncopation (see sections 2.3.3 and 4.3.1). These aspects are complex and there exists a huge amount of analytic literature to describe those kinds of interrelations and interplay between rhythm and meter.[13] This thesis will in contrast focus on the aspect of variability and equivocality, which is exemplified in figure 1.1. The temporal structure of the rhythmic cycle is always the same but as our metric framing can be different, the impression of the whole musical passage, the rhythmic-metric entity, can change a lot.

Western musicians reflect the different notational variants in figure 1.1 in their musical interpretations.[14] Consequently, the rhythmic figure may be less ambiguous when it is performed in a musical context. On the other hand, there still may be situations where different listeners construct different internal images by building up different

[9] cf. for instance Benadon, 2007, Sethares, 2007, and Toussaint, 2013. See section 4.1.3 for an examination of the mathematical and word-theoretic notions of necklaces, clarifying the benefits of their application in the present study.

[10] The cognitive foundations of metrical accent are discussed in chapter 2. They form the basis of hierarchic metrical frameworks, which are further examined in chapters 3 and 4, regarding different features and research perspectives.

[11] cf. London, 2012, pp. 107 f.

[12] cf. Hasty, 1997, pp. 89, 130 f., 152, 154.

[13] Descriptions and investigations on the interaction between rhythm and meter largely depend on the actual definitions of the relative concepts of rhythm and meter (see sections 1.2 and 1.3).

[14] See for instance London, 2012, pp. 99 f.

metric frameworks. I will call that space of tolerance the space of *metric ambiguity*. Accordingly, a rhythmic figure or cycle has a certain flexibility allowing it to be molded into different metric frameworks. In other words, it has the potential to elicit different metric interpretations. I will call this the *metric malleability* of that rhythm. This term and the precise notions of metric ambiguity and metric malleability I adapt from LONDON.[15] Metric malleability also relates to "the listeners' ability to change their metric construal"[16] of a continuously repeating rhythmic figure. LONDON reports a correspondent experiment conducted by VAZAN and SCHOBER. A "set of computer-generated stimuli"[17] based the rhythmic cycle [x.xxx.] or [2-1-1-2], which is similar to the cycle shown in the figures above, was presented in different metric contexts, corresponding to figure 1.1.

> In one part of the experiment, subjects were given a 3/4 prompt, and then they heard a particular rotation of the pattern; subjects were asked to tap out the beats. They were then given a visual presentation of a metric alternative (either a shift of the downbeat or shift to 6/8) while the aural presentation of the rhythmic figure continued; their task was to tap the beat, which required them to suppress the initial metric prompt. This proved quite difficult, especially if the rhythmic pattern began with a rest.[18]

The reverse task, to shift the rhythmic pattern from one to another metric rotation while the meter is constantly maintained, turned out to be "far easier, even when the rhythmic pattern began with a rest."[19] In short, it is more difficult to recast the meter of a continuously repeating rhythmic cycle than to shift the metric phase of the cycle to a continuous meter. The performance difficulty of reinterpreting meter is well known. This is also reflected by the musical challenge of *metric modulation*.[20] In sum, the interpretive field of metric ambiguity and malleability suggests an interesting scope for musical design, comprising hemiola (see section 3.2.4) and metric conflict (see section 2.3.3).

This short and plain introduction of the topic necessarily lacks of descriptive and analytic precision. Some important aspects are not yet taken into account, as for instance the impact of absolute temporal durations (inter-onset intervals) on metric interpretation (see sections 2.2.2 and 2.2.3). The concept of malleability will be approached from different perspectives. First, metric interpretation will be described as a cognitive activity, exhibiting a remarkable flexibility which particularly facilitates metric recasting to adapt to changing contexts. Secondly, rhythmic structures will be formally

[15] London, 2012, p. 13, uses the terms *metrically malleable* and *metrically neutral* in parallel. He also shows that metric ambiguity and metric malleability are not synonymous (pp. 99 f.) and differentiates these notions compared to Povel, 1984, p. 325 (note 2, p. 199).

[16] London, 2012, p. 65

[17] ibid.

[18] ibid.

[19] ibid.

[20] cf. for instance Benadon, 2004, pp. 563, 566. See also section 4.4

analyzed in terms of possible determinants which may correlate with metric ambiguity and malleability. Finally, musical introspection and artistic research can help to gain more practical insight for the development of applications. From an artistic perspective it is interesting to know how the context may be changed or built up to induce perceptual effects with musical meaning, as mentioned above. A multidisciplinary approach to this topic is anyway appropriate as many researchers are aware of its benefits across the domain of rhythm. LONDON for instance states that "the results of psychological research in rhythmic perception and cognition are old news to musicians"[21] and hopes that "it is obvious how much research in music perception and cognition can add to music theory and musical analysis and vice versa."[22]

Obviously, there exist multiple variants of musical hearing, both among different listeners, as well as for a same listener, who will adapt the mode of hearing to different musical stimuli or to an evolving musical input. Given this variety, perceptual and cognitive modeling conversely depends on the definition of commonalities or invariants to delimit the space of perceptual variation. I will examine relevant models of rhythm perception and metric interpretation, based on empirical research (sections 2.2.3 and 5.1), but it is also helpful to look at musicological theories which rely much more on musical practice and introspection. Conceptions of meter are as heterogeneous as their contexts differ (see section 1.3). As we deal with musical experiences we have to be aware – as HASTY puts it – that

> descriptive categories are by nature abstract. Apart from these problems of abstraction, but at least as problematic for analysis, is the fact that the description of an act of hearing is not that act. At best, analysis can be only a speculation on certain possibilities of experience.[23]

Therefore, we can only try to reconstruct a space of possible experiences although the actual character of an experience is its particularity,[24] its uniqueness and embodied individuality. HASTY continuously gives examples of introspective exercises of listening against unconscious perceptual mechanisms.[25] Conscious exploration of hearing possibilities can give an idea about the amount of voluntary control, we can obtain in listening to a musical passage or a sequence of sounds. Hence, it may not be possible to predict precisely the variance of metric interpretations of a particular rhythm. I thus favor a heuristic approach to characterize spaces of metric ambiguity and malleability, respecting the openness and subjectivity of interpretation and the possibilities of voluntary control.

[21] London, 2012, p. 198
[22] ibid.
[23] Hasty, 1997, p. 154
[24] The particularity of meter as an experience is a central notion in Hasty, 1997.
[25] Hasty, 1997, see for instance pp. 103 f.

The primary goal of this thesis is a better understanding of the subjective character of musical rhythm and a computational approach to metric malleability that reflects this character and that is accessible for diverse applications. To establish an appropriate basis for this project, the relevant fundamental aspects of rhythm perception and cognition are surveyed in chapter 2. Thereafter, metric ambiguity and malleability are systematically described and related to specific rhythmic structures and traditions, reflecting the variety of cultural contexts (chapter 3). Chapter 4 then identifies the formal framework for a model of metric malleability, including extended definitions of hierarchic schemes of metrical accent and their relations. This model is finally developed in chapter 5. First, models of metric interpretation or beat induction are evaluated (section 5.1) and adapted (section 5.2.1) to suit the presented heuristics of metric malleability. Then, specific aspects of this model are defined and examined by means of examples (sections 5.2 and 5.3). The original contributions of the present thesis can be overviewed in the following shorthand list, ordered by their relevance (from central to additional).

1. a quantitative heuristics of the metric malleability of cyclic rhythms (chapter 5), including analyses of exemplary computations (section 5.3)
2. adjustments and extensions on the model of beat induction (*pulse salience*) by PARNCUTT[26] and its adaptation by FLANAGAN,[27] resulting in a formal definition of *meter salience* (section 5.2)
3. a formal description of hierarchic meter, integrating *mixed* beat durations or *non-isochronous* metric pulses (section 4.2)
4. a formal definition of *metric contrast* between two hierarchic meters (section 4.4.2)
5. an extension to the theory of *metric indispensability* by BARLOW[28] and an alternative algorithm (section 4.3.3)

1.2 Perspectives on rhythm

A concise definition of rhythm can hardly, if not impossibly be adequate to indicate the complexity of the matter. There exist many perspectives and domains where rhythm or rather certain aspects of rhythm play different roles, even from a restricted view on the

[26]Parncutt, 1994
[27]Flanagan, 2008
[28]Barlow, 2012

realm of musicology and music research.[29] The diversity of perspectives on rhythm is raised by the various aims of studies on rhythm. In this section, three aspects of rhythm are shortly reviewed which are relevant for the further discussion: the motional qualities of rhythmic performance, the perception and cognition of rhythm and formal properties of rhythmic structure. The latter two are examined further in chapters 2 and 4.

It may be informative to examine roughly the general phenomenon of rhythm, before we focus on the mentioned specific aspects of musical rhythm. In a general sense, rhythm can be realized and interpreted in any accumulation or collection of elements. Think about a succession of events or movements in time or a spatial arrangement of things. The perceptual and formal properties of rhythm emerging from a temporal succession or a spatial arrangement are not mainly dependent on the properties of the elements including their position in time or space. They rather arise from the temporal or spatial relations between different elements and groups of elements. The psychology of representing rhythmic structure is investigated since the 19th century[30] and in the 20th century, rhythmic interpretation and grouping principles were explored especially in the tradition of gestalt psychology.[31] The perception of temporal or spatial relations may have already been the background concept in PLATO's definition of rhythm as the order of movement.[32] According to this view, rhythm is perception in the sense of active ordering of movement.[33] HASTY similarly notes: "What is rhythmic is ordered and therefore comprehensible, but this is an order that cannot be abstracted from the thing or event."[34] His supplement, that order depends on a vehicle, is further discussed in section 2.1. Contemporary notions and definitions of rhythm often follow the perspective on rhythm as perception.

> The concept of rhythm has been less straightforward to define and can be viewed either as a serial pattern of durations or the perceived temporal organization of that pattern. In the second sense of the term, rhythm as a perception has several elements including grouping, beat, and meter.[35]

MCAULEY here classifies grouping and meter as elements or modes of rhythm perception, and we will focus on this aspect throughout this study. Metric ambiguity and

[29] For recent illustrations see e.g. Toussaint, 2013, pp. 2 ff. and Bello et al., 2015. Toussaint lists a collection of definitions and characterizations of rhythm from twenty-two different authors and concludes by quoting C. Egerton Lowe's statement that there is "no other term used in music over which more ambiguity is shown." Bello et al. introduce the reader of the Journal of New Music Research to five articles comprising five current scientific views on rhythm: from cognitive science, performance analysis, ethnomusicology, music information retrieval and signal processing.

[30] e.g. Bolton, 1894

[31] e.g. Deutsch, 1982, Tenney and Polansky, 1980

[32] see Fraisse, 1982, p. 150

[33] see Hasty, 1997, pp. 10 f. for a survey of Plato's contribution to the ancient Greek notion of *rhuthmos*.

[34] Hasty, 1997, p. 6

[35] McAuley, 2010, p. 192

malleability are certainly consequences of the diversity and flexibility of our perception, specifically, to build metric frameworks while perceiving a rhythm. TOUSSAINT characterizes "rhythm as a manifestation of a *process* that emerges from the amalgamation of a physical signal with perceptual and cognitive structures of the mind."[36] In accordance with HASTY, processes of rhythmic activity are regarded as closely related to the properties of a physical vehicle, as complex this relation may ever be.

Human rhythmic behavior in its entirety involves the whole organism. Theories of embodied music cognition[37] and the exploration of neurophysiological aspects of rhythmic behavior[38] contribute to the comprehension of rhythm as a deeply internalized way of human interaction. The theory of *dynamic attending* (DAT), developed by JONES et al.[39] was particularly influential for the development of theories on essential aspects of rhythmic behavior: rhythmic entrainment and synchronization,[40] pulse perception,[41] expectancy and anticipation.[42] These are major factors in the process of metric framing and are extensively discussed in chapter 2.

Rhythmic practices in music are also embedded in broader cultural contexts. Language and speech for example may be determining factors for certain culture-specific rhythmic habits.[43] A particular source of cyclic rhythms or rhythmic repetition are performance practices in ritual traditions. In these contexts, sustained rhythmic entrainment induced by repetitive structures is frequently used as a device to enter states of trance and ecstasy.[44] In sum, rhythmic behaviors and traditions in different musical cultures exhibit multiple and diverging motional, emotional and cognitive qualities.[45] Nevertheless, comparative studies account for certain commonalities across cultural differences. In section 1.3, some of them are surveyed in terms of meter.

[36]Toussaint, 2013, p. 6

[37]cf. Leman and Maes, 2014

[38]See e.g. Brochard et al., 2003, Chen, Zatorre, and Penhune, 2006, Grahn, 2012, Honing, Bouwer, and Háden, 2014, Vuust and Witek, 2014

[39]See e.g. Jones and Boltz, 1989, Barnes and Jones, 2000

[40]A comprehensive survey of entrainment theories is provided by McAuley, 2010. London, 2012 considers entrainment as the behavioral basis for the perception of meter.

[41]cf. Parncutt, 1994 and Parncutt, 1987 for an elaborated model, which is an essential basis for the propositions of this thesis.

[42]Expectancy and anticipation are closely linked aspects, see for instance Desain, 1992, Barnes and Jones, 2000, Huron, 2006

[43]Huron, 2006, pp. 188 ff., Patel, 2008, pp. 96 ff.

[44]cf. McLachlan, 2000, pp. 63 f., referring to Becker, J. (1994). "Music and Trance". In: *Leonardo Music Journal* 4, pp. 41–51: "Rhythmic entrainment of the mind through music may assist performers to enter these states and remain within them. Once performers enter ecstatic states, their brain chemistries alter, for example to inhibit pain and fatigue." (p. 64)

[45]We restrict the discussion to movement aspects (in section 1.2.1) and cognitive aspects (in section 1.2.2).

1.2.1 Motion and performance

> Rhythm, musical rhythm included, essentially may be considered an important subset or aspect of motion.[46]

BENGTSSON here understands motion as "both real, bodily, motoric and kinetic movement as well as virtual and imagined."[47] For sound to be produced, there has to be movement, and hearing involves gathering movement from sound. This ability probably raised from the evolutionary pressure to be able to infer potential threats from the sounding environment.[48] Normally, we unconsciously gather the physical sources of the sound we hear. In contrast, listening to music involves considering sounds as "ends in themselves".[49] A similar distinction holds for musical rhythm, which

> is distilled from its everyday ecological significance in the concert hall. [...] rhythm signifies movement, but musical tones do not move. Rather, we hear a kind of virtual motion in a virtual, acousmatic space.[50]

Hence, the range of rhythmic motion includes the physical space of performance and the virtual space of musical listening. Indeed, rhythm inextricably links action and perception. We embody rhythm through movement, in musical performance, dancing or listening, emphasized by motor action (for instance by tapping the beat or moving the head with the beat).[51] Rhythmic experience is characterized by the integration of motor behaviors with the internal processes of cognitive representation and rhythmic anticipation. This capability forms the basis of voluntary movement coordination, synchronization and dynamic adaptation in rhythmic practice.[52]

The scientific measurement of performance timing in music exhibits certain biases. Such systematic temporal "variations in performance from a mechanical norm"[53] are usually perceived as accurate in regard to this ideal norm.[54] As a musician is also listening while performing, there is intrinsic and immediate perceptual feedback to the performative system. Thus, it is not surprising that performance biases closely relate to perceptional biases.[55] Both are also systematically related to properties of the rhythmic

[46] Bengtsson, 1987, p. 71
[47] ibid.
[48] A related issue, emphasized by Huron, 2006, is the evolutionary development of the ability to anticipate future events by means of accurate expectations. Expectation and anticipation are essential aspects of rhythm perception, and are extensively discussed in section 2.2.1.
[49] London, 2012, p. 6
[50] ibid.
[51] cf. Parncutt, 1994, p. 411, referring to the ecological approach of J.J. Gibson (see also London, 2012, pp. 9 f.).
[52] Fraisse, 1982, pp. 175 f.
[53] Gabrielsson, 1982, p. 165
[54] See also Bengtsson, 1987, pp. 74 f.: "ideal performance is not metrical exactitude".
[55] However, these relations may be complex as differences occur in comparisons of perception and production measurements. Supporting the assumption, that rhythm perception and performance are closely linked, Sadakata, Desain, and Honing, 2006 showed with Bayesian formalization that it is nevertheless

structures to be produced.[56]

Moreover, rhythmic performance in music not only aims at timing accuracy in regard to norms like evenness or regularity. Rhythmic expression is communicated via deliberate but consistent temporal deviations from this norm. *Expressive timing* is found across most musical styles and cultures.[57] For a successful communication of rhythmic expression, two kinds of information – symbolic norm and expressive deviations – have to be decoded and separated in the process of listening. In section 2.3.4, this process is further discussed in terms of *categorical rhythm perception*.[58]

To summarize, rhythm is communicated by motion. The peculiarities of rhythmic performance have impact on the perception of rhythmic properties. Performing a rhythm or moving with it enhances its perception and alters its phenomenal quality. Hence, the motional aspect is relevant for the interpretative behavior of a listener. To induce (real or virtual) motion, a rhythm has to have motional quality. This study focusses on rhythms implying motion of this kind, which is a requirement for the emergence of metric frameworks. Rhythmic motion is also restricted to a certain time scale. This is thoroughly discussed in section 2.2.2. Hence, rhythms obscuring motion by some sort of complexity or irregularity, or by their time scale, can not be metrically malleable. Expressive timing and performance biases are also important for that they often reduct metric malleability. In many musical contexts, these temporal nuances and other cues like dynamic accents are employed to metrically disambiguate the rhythm.

> In counting according to one meter and not another, a musician gives a series of tones a particular rhythmic shape and nuance: their sense of meter leaves a kind of residue in performance, such that the "same" series of notes played under different counting frameworks will have distinctive differences in its expressive timing and dynamics.[59]

There nevertheless exist musical aesthetics and practices where rhythmic performance is oriented to increase malleability and to expand the space of metric ambiguity. For instance, LOCKE describes African rhythmic timing practices which increase metric ambiguity by simultaneous cues for multiple metric interpretations.[60]

---

possible to predict perception data from production data. This "may be interpreted as an optimal adaptation of our perceptual system to the environment in which the produced rhythms occur." (Sadakata, Desain, and Honing, 2006, p. 269)

[56]Repp, London, and Keller, 2011 e.g. examined perception and production of repetitive two-interval rhythms with simple and complex interval ratios and found "similar ratio-dependent biases: rhythms with small ratios were produced with increased ratios, and timing perturbations in these rhythms tended to be harder to detect when they locally increased the ratio than when they reduced it. The opposite held for rhythms with large ratios." (Repp, London, and Keller, 2011, p. 227)

[57]See for instance Polak, 2010 and Polak and London, 2014 for studies on African variants of non-isochronous beat subdivision. Baraldi, Bigand, and Pozzo, 2015 for instance measure expressive timing in an Aksak style of Transylvania.

[58]cf. Clarke, 1987, Desain and Honing, 2003

[59]London, 2012, p. 4

[60]cf. Locke, 2011

1.2.2 Perception and cognition

> current music research does not just borrow from the (cognitive) sciences, it also contributes to it in significant ways. Music, like language, is more and more considered an important domain to study cognition and culture[61]

Cognitive musicologists often describe human interaction with musical rhythm as a special case of more general cognitive behaviors.[62] Thus, to study rhythm is to study cognition at a concrete level. According to cognitive theories, the phenomenal qualities of rhythm are determined by the way we generally process sensuous input. This process is guided by principles of economy and consistency. FRAISSE emphasizes the interdependence of temporal intervals in rhythmic structures which are subjectively perceived according to the "salience of good form".[63] This involves cognitive mechanisms like assimilation and distinction, that is, the rationalization of interval ratios according to processing capacities. These properties of rhythm cognition are also involved in performance timing. As mentioned before, categorical rhythm perception (see section 2.3.4) implies separate processing of a rationalized rhythmic "grammar" and the "phonetic" variants articulated in expressive timing. FRAISSE further notes that rhythm emerges when prediction becomes possible on the basis of what has already been perceived.[64] Accordingly, it was mentioned at the beginning of this section that rhythm cognition theories stress the role of entrainment as a cognitive behavior, involving dynamic attentional patterning, expectancy, and pulse sensation. These notions are central for specifying rhythm perception as an active, temporally structured cognitive engagement and will be differentiated and amplified in section 2.2.

Expectancy and pulse sensation may be specifically distinguished in regard to metric interpretation. PARNCUTT suggests that "expectancies also occur in nonrhythmic sequences" and that it is therefore "inappropriate to define musical rhythm in terms of expectancy."[65] This can be understood from his perspective, assuming that the phenomenon of musical rhythm perception essentially involves pulse sensation.[66] As pulse sensation implies the perception of periodicity, this stands in contrast to cognitive concepts of rhythm claiming that "periodicity is but one type of rhythmic organization".[67] Non-periodic speech rhythms, for instance, are as well perceived in terms of temporal expectations but do not induce pulse sensations which are the basis for metric interpretation. Pulse sensations imply periodic motion which can be bodily

[61] Honing, 2006, p. 4
[62] cf. for instance Huron, 2006, Patel, 2008, Vuust and Witek, 2014
[63] Fraisse, 1982, p. 167
[64] Fraisse, 1982, p. 150
[65] Parncutt, 1994, p. 452
[66] Parncutt, 1994, p. 453
[67] Patel, 2008, p. 96

felt only within a confined temporal window. Hence, the emergent quality and eventually the metric implications of rhythm highly depend on temporal parameters (see section 2.2.2).[68]

1.2.3 Structural and formal analysis

Rhythm can also be studied purely from a formal perspective. For this purpose, it is necessary to determine a level of symbolic abstraction from the complexity of sounding rhythms in real life. In music, any kind of event (e.g. attack, accent, rest) or change (e.g. of pitch, harmony, timbre), as part of the musical flow, may have rhythmic impact because the temporal flow gets rhythmically structured and divided. Section 2.1 provides a discussion of those aspects of a musical surface, which can carry rhythmic meaning in terms of *rhythmic components*. In contrast, to study its formal properties, a rhythm has to be defined as a structure in a metric system in the mathematical sense. On the perceptual level, a similar abstraction, already mentioned above, takes place by means of categorical perception (see section 2.3.4). Hence, both rhythmic perception and analysis are facilitated by a simplification and categorization of a multifaceted acoustic reality. These are similar processes in that they abstract distinct categories from a potential musical continuum.

Music theory in the tradition of the Pythagorean paradigm is also based on abstraction from concrete sound. Studied objects are mainly patterns and relations, like scales and intervals, simplified and rationalized. Structural principles are generalized,[69] for instance, pitch and rhythmic pulse rate are regarded as frequencies or periodicities in different time-dimensions (see section 4.1.1). In the 20th century, "the emergence of algebraic methods in music has been a long process that has occurred, surprisingly, independently from stylistic considerations and geographical contexts."[70] The fact that rhythmic cycles and musical scales are isomorphic,[71] allows to apply algebraic and geometric methods to either domain in an analogous manner.

Computational modeling also has reached an outstanding role in cognitive musicology.[72] Formalized models of rhythm perception have been developed since a number of decades.[73] Both formal investigations in rhythms and studies on rhythm perception

[68] See e.g. Handel and Oshinsky, 1981, p. 9, London, 2012, p. 5

[69] See for instance Lewin, 1980, Xenakis, 1992.

[70] Andreatta et al., 2006, pp. 277 f. The authors specify that "in the *set-theoretical* (Forte) and *transformational* (Lewin) approaches Milton Babbitt's initial formalization of the twelve-tone system has some very deep intersections with the music-theoretical constructions proposed around the 60s by some European theorists/composers, in particular Anatol Vieru and Iannis Xenakis. The system was conceived as a 'collection of elements, relations between them and operations upon them'" (emphases in source, embedded quote from M. Babbitt, The function of Set Structure in the Twelve-Tone System, PhD, Princeton University, 1946.)

[71] See section 4.1.1

[72] cf. Honing, 2006

[73] See e.g. Lerdahl and Jackendoff, 1983, Longuet-Higgins and Lee, 1984, Lee, 1991, Parncutt, 1994

and cognition have widely pursued the paradigm of working solely with representations of temporal interval sequences. Abstracting from properties of sound and from the differences in articulation of these intervals is clearly a reduction from a mass of possible influences on perception. Though, both for investigations on rhythm perception and cognition, as well as for studies on formal properties of rhythm, this approach has turned out fruitful.[74] Numerous studies of the latter type are voluntarily restricted to temporal sequences.[75] This does not mean to ignore the complexity of contextual manifestations of rhythm. Rather, the mathematical complexity inherent in abstract patterns already attracts separate investigations.

TOUSSAINT concludes from his mentioned characterization of rhythm as "a *process* that emerges from the amalgamation of a physical signal with [...] structures of the mind" (see above, page 7), that "rhythm may be studied at any level" to gain knowledge about mathematical, cognitive and cultural aspects which "helps us to understand the totality of of rhythm."[76] Nevertheless, he focuses "predominantly at one extreme of the above panorama: rhythm is considered purely in durational terms as a symbolic binary sequence of isochronous elements representing sounds and silences."[77] Such a "pulse-based", discrete mathematical definition contrasts with approaches to rhythm as a continuum.[78] However, it turns out to be the appropriate form of representation for the studies in this thesis. On the perceptual level, metric interpretation is based on pulse sensation (section 2.2) and categorical rhythm perception (section 2.3.4), suggesting isochronous grids of rhythmic pulses. Several types of corresponding formal representation systems are surveyed in chapter 4.

As mentioned at the beginning of this section, the analytical reduction of rhythm to temporal patterns necessarily lacks in other dimensions which have impact on rhythm perception. Complex musical structures (polyphony, melodic and harmonic progression, etc.) can only be taken into account if many other features like pitch, timbre and sound location are examined. It is a challenge to create a reasonable balance between analytic precision and *ecological validity*[79] of the investigation method. The scope of this study is thus restricted to analytical hypotheses and a computational heuristics of metric malleability. It would be the task of further work to test it on an ecologically valid level.

Properties of rhythmic structure which become obvious when formally analyzed may

[74] Among such studies on rhythm perception, Desain and Honing, 2003, Flanagan, 2008, Fraisse, 1982, Handel and Lawson, 1983, Madison and Merker, 2002, Parncutt, 1994, Povel and Okkerman, 1981, Povel and Essens, 1985, and others form the basis of argument in this thesis (see chapters 2 and 5).

[75] Some of these studies are particularly discussed in chapters 4 and 5.

[76] Toussaint, 2013, p. 6

[77] ibid.

[78] cf. for instance Desain, 1992

[79] cf. Jones, 2010, pp. 6 ff., London, 2012, pp. 171 f.

not necessarily be perceivable.[80] Moreover, there are many ways to define and measure abstract properties of rhythm. For instance, to measure the complexity of a rhythm, many notions of formal complexity could be applied,[81] but these measures may not at all correspond to cognitive estimates of rhythmic complexity.[82] Generally, the applicability of formal measures to rhythm has to be carefully verified or falsified in each case. A systematic connection of formal properties of rhythmic cycles to a cognitive measure of metric malleability thus seems to be unlikely. Nevertheless, it remains an interesting challenge to explore possible correlations in future research. In section 5.3, an exemplary heuristic approach of such associations is sketched.

1.3 Perspectives on meter

Meter inheres in rhythm and rhythm brings about meter. As an aspect of musical – and more specifically, rhythmic – structure, meter is a matter of musicological analysis and of practical musical communication. Metric notation is used to provide a basis for synchronization of different voices and to communicate more subtle instructions for timing and articulation. Thus, there are technical and aesthetic reasons to employ meter as a temporal pattern, based on periodicity, which orientates rhythmic activity. From a cognitive or behavioral point of view, meter is regarded as emerging from a dynamic process of engagement with rhythm. It can elicit body movements[83] and influences the mode of listening. Attentional energy gets temporally shaped in a periodic fashion (see section 2.2.1) and expectations about the proximate progression arise.[84]

Thus, two main concepts are preliminarily distinguished: meter as a practical tool for musical communication and meter as a dynamic cognitive process. For the purpose of this study, I suggest to regard both perspectives as aspects or phases of a single communicative process. On an inter-individual level, meter is externally communicated by conventional means: musical notation, conducting and interpretation. On an individual level, meter is internally communicated by perceiving and cognitively processing rhythm. Both processes intertwine: we may notate a rhythm in a certain meter as a result of internal reflection and interpretation, or we internally process a rhythm according to a meter, strongly suggested by a certain musical interpretation. On both levels, meter informs rhythm. In specific musical languages, particular senses of meter developed as generally admitted grammars, according to which rhythm may be

[80] For instance, an often discussed feature of pattern structure which is hard to recognize aurally, is ("mirror") symmetry, that is, a rhythm and its inversion (or *retrogradation*) are identical (cf. sections 4.1.3 and 4.2.2).

[81] cf. for instance Polansky, 1996, Thul and Toussaint, 2008

[82] Pressing, 1997, Shmulevich and Povel, 2000, Toussaint, 2013, pp. 107 ff.

[83] As it was mentioned in section 1.2.1, rhythm inextricably links action and perception. Meter can provide a scheme or a temporal orientation for the coordination of these aspects.

[84] Hasty, 1997 characterizes meter as *projection*. This is in line with cognitive theories of expectancy, e.g. Huron, 2006, Desain, 1992.

parsed. At the cognitive level, rhythm emerges as a *gestalt*, shaped by that sense of meter.

As mentioned earlier, cognitive models of meter are based on the idea that attentional energy gets temporally organized in regular patterns. Peaks of attentional energy are perceived as accented or stressed. Those patterns can be simple or complex. At this point, two aspects of complexity may be mentioned, before examining them in depth in the course of this thesis.

The first aspect reflects that meter can be organized hierarchically as stratified levels of motion. In his classical work on the stratification of musical rhythm, YESTON describes meter as "an outgrowth of the interaction of two levels - two differently-rated strata, the faster of which provides the elements and the slower of which groups them."[85] A metric level of motion can be established by a regular succession of similar musical events, either of elementary or of more complex events like harmonic changes. Rhythmic stratification can be more complex, for instance when melodic and harmonic rhythms differ. In contrast, a stratified meter consists of embedded levels of regular motion. YESTON also points to the effect of metric grouping, that is, a slower metric level groups the elements of the faster one. A potential metric hierarchy may be embodied by concrete rhythmic activity to a certain extend. In a particular musical situation, the depth of a metric hierarchy can dynamically be built up and dismantled again, or there may be abrupt changes (see section 2.2.1, particularly figure 2.7).

A second aspect refers to the type of regularity involved in metric movement. A level of plain regularity consists of isochronous temporal intervals. If this isochrony is broken on a certain metric level, another type of meter may result. Nomenclature for this phenomenon is varied: *additive* meter,[86] *non-isochronous* meter[87] and *mixed* meter[88] are terms in use (see sections 3.2.1 and 4.2).

The prevalences of metric types vary in different musical cultures. Mixed meters are used more frequently in Eastern traditions than in Western music, from South-East-Europe to India and further. Whether the concept of Western stratified meter can appropriately apply to sub-Saharan African music is subject to debate.[89] Comparative studies explore if cognitive models of meter allow for culturally independent generalization. For instance, some authors asked to what extend the classical formal model of meter, proposed by LERDAHL and JACKENDOFF in the *Generative Theory of Tonal Music* (GTTM)[90] would be applicable to specific cultural contexts. CLAYTON[91] studied

[85] Yeston, 1976, p. 66
[86] As a complement to "divisive meter", this term has been introduced by Sachs, Curt (1953). *Rhythm and Tempo: A Study in Music History*. New York. (cf. Wright, 2008, pp. 25 f.)
[87] cf. London, 2012
[88] cf. Gotham, 2015b
[89] cf. Locke, 2011, Temperley, 2000, Toussaint, 2015
[90] Lerdahl and Jackendoff, 1983
[91] Clayton, 1997

the compatibility with the Indian *tal* system and TEMPERLEY[92] examined sub-Saharan African music with the result that the theory can widely account for it with minor extensions and assimilations. Nevertheless, African and Indian traditions developed very different aesthetics in respect to metric ambiguity. While in African music metric ambiguity (multiple possibilities of metric interpretation) is part of the aesthetics, it is a foreign concept to Indian musicians (see section 3.1.2).[93] More generally, metric systems and aesthetics evolve within cultural contexts which enable certain features and hinder other ones. Individual learning and enculturation processes are embedded in those collective developments. LONDON characterizes meter as a "kind of behavior"[94], which is acquired and trained during a long-term learning process.

> Listening metrically involves our musical habits, and not just a few generic habits but a rich repertoire of metric responses to rhythmic patterns and processes.[95]

Indeed, many studies focus on the issue of how we implement our long-term experience with music as top-down strategies in momentary listening behaviors. In respect of meter, as well as of other musical aspects, it is shown that experience and sustained exposure to specific music and its cultural background have impact on music perception and recognition.[96]

Ideally, the outcomes of the present study should be independent from any cultural predetermination. Moreover, a certain amount of inter-cultural dispersion of metric interpretation is another factor which has to be taken into account, in addition to inter-individual differences in metric interpretation behaviors, which we described as metrical ambiguity. Nevertheless, the aim – or rather, the utopia? – of generality beyond and independent of cultural context or musical habits ("styles") can serve as an orientation for developing generic computational models for perceptual phenomena, with all their inherent ambiguities.

The relation of meter and rhythm is a contentious issue. The citation[97] on page 7 suggests that meter, as an aspect of rhythm perception, is a cognitive activity. As mentioned, meter can only be communicated and embodied by concrete rhythmic activity. HASTY describes the emergence of meter as a process of temporal projection, leading to a unique interpretation and denies an opposition of meter and rhythm.[98] Meter evolves

[92] Temperley, 2000
[93] cf. Clayton, 1997, Locke, 2011
[94] London, 2012, p. 8
[95] ibid.
[96] For an overview, see Huron, 2006. Palmer and Krumhansl, 1990 is an example of a classical study in regard to meter. Hannon and Trehub, 2005 focus on the effect of enculturation processes on meter perception.
[97] McAuley, 2010, "In the second sense of the term, rhythm as a perception has several elements including grouping, beat, and meter."
[98] Hasty, 1997, p. xi: "a creative process in which the emerging definiteness or particularity of duration is shaped by a great range of qualitative and quantitative distinctions, we will have no reason to oppose meter to other domains or to rhythm"

as a particular experience. On the other hand, formal or generic aspects of meter can be abstracted from this integral experience as a reduced template or pattern. We could call this pattern a class of potential experiences of concrete projections in the sense of HASTY. On this abstracted level, the connotations of rhythmic and metric patterns run the risk of getting blurred, maybe for historical reasons.

> In the Renaissance, theorists spoke of "rhythmic modes" – by analogy to pitch-related modes – instead of meters. From a psychological perspective, this analogy seems warranted. Well-known rhythms (such as dance patterns) act in a manner very similar to meters.[99]

HURON calls those *metric* rhythms "rhythmic schemas"[100] to emphasize that they represent a metric repertoire in musical memory. They are rhythmic instances of metric schemes. In section 3.2.2, the problems arising from the opposition of rhythm and meter and the abstraction of metric schemes are discussed in terms of *rhythm-meter ambiguity*. Nevertheless, this study suggests that metric interpretation can be regarded as the emergence of rhythmic gestalt from the interaction of rhythmic and metric aspects. Hence, their differentiation and the discussion of metric schemes seem to be valid tools to study metric ambiguity and malleability.

Meter (regarded as a communicative process or a grammatical system) allows for interpretative variance and is prone to interferences. As demonstrated above, the gestalt of the same rhythm may dramatically change according to different metric interpretations (see figure 1.1). The perspective on metric interpretation is different for the musician and for the listener. Performers usually seek to effectively communicate metric structure by certain cues such as timing and accentuation. They may thus disambiguate potentially malleable rhythms[101] as in figure 1.1. This communication can also fail, causing accidental metric ambiguity. In a more sophisticated manner, a performance can even deliberately throw somebody off the scent and construct musically challenging traps, such as metrical "garden path surprises".[102]

A model of pulse salience and metrical accent by PARNCUTT (see sections 2.2.3 and 5.2) is explicitly based on the premise that rhythmic patterns can evoke several pulse sensations at the same time.[103] To concentrate on pulse sensation rather than on the more complex issue of cognitive emergence of meter is advantageous, in that a concrete rhythm may embody more or less complex metric depths and has more or less

[99] Huron, 2006, p. 202
[100] ibid.
[101] London, 2012, p. 22
[102] Huron, 2006, pp. 279 ff. Referring to Temperley, D. (2001). *The Cognition of Basic Musical Structures*. Cambridge, Mass.: MIT Press., Huron, 2006 describes an instance of a metrical "garden path" at the beginning of the third movement of Ludwig van Beethoven, *Piano Sonata*, op. 14, no. 2: "Ostensibly, the movement is written in 3/8 meter. However, the opening rhythm leads unsuspecting listeners down a different metrical path. [...] Most listeners are apt to hear the passage as beginning with a 2/8 meter and switching to 3/8 at the third measure. A metric 'glitch' is thus introduced." (pp. 280 f.)
[103] Parncutt, 1994

potential to be metrically ambiguous. Pulse sensation, considered as an experience of a listener, can be determined experimentally by recording motor reactions of listeners to a perceived rhythm, according to their principal pulse sensations. Moreover, a listener may sense several temporally embedded pulses at once, or different listeners may have different principal pulse sensations. The former can be interpreted as an emergence of a metric hierarchy and the latter can be regarded as metric ambiguity in the mentioned sense of inter-individual differences in conceptualizing rhythm. PARNCUTT's model is an essential background of this study, as the phenomenal experience of certain peculiarities of meter is emphasized, specifically, as possible phenomenal consequences of metric malleability are examined. According to this model, a rhythm is metrically ambiguous when saliences of incommensurable or metrically *dissonant*[104] pulses are similar, that is, when they receive a similar amount of response on an inter-individual level.

[104]cf. section 2.3.3

Chapter 2

Relevant aspects of rhythm perception and cognition

Two aspects govern recent research on rhythm perception and related cognitive mechanisms. One is the way rhythm is processed by our perceptual and cognitive system. Cognitive coding mechanisms may underlie perceptual grouping phenomena. They seem to be guided by economic principles, already proposed by gestalt theory. The other aspect is our response to rhythm, the behavior we adopt in rhythmic interaction. Rhythmic synchronization and anticipation are interpreted as specific capabilities of embodied music cognition.[1] This chapter discusses the essential features and dynamics of rhythm perception and cognition which account for the phenomena of metric ambiguity and malleability. In musical contexts we act within our perceptual and motional characteristics and constraints. Hence, the goal of this chapter is the description of the mentioned phenomena as consequences of these constraints and features of rhythm perception. Insight of studies in empirical and cognitive musicology will be contrasted with introspective analysis of rhythmic interpretation, that is, analytic methods involving subjective listening. It is a challenge to combine these different types of knowledge, although, ideally they may complement each other in a synergic way.

The sections of this chapter divide the topic into four aspects which have many mutual relations. Unfortunately, the linear progress of discussion can only successively shed light on these relations. They are essential in approaching higher-level cognitive processes and to gain insight in metric ambiguity and malleability.

In section 2.1, temporal intervals are distinguished from their musical articulation. Classes of articulation and their perceptual properties will be conceptualized as *components*[2] of rhythm. Both temporal intervals and rhythmic components can function as perceptual cues for rhythm.

Section 2.2 is dedicated to pulse sensation, a typical human response to many musically rendered rhythms, accompanied by embodied activities like entrainment and

[1] cf. Leman and Maes, 2014, Honing, 2006

[2] I have adopted this term from Petersen, 2010.

anticipation. Rhythm is molded into a cognitive scheme, retroacting on the subjective impression of rhythmic components.

In section 2.3, pulse perception is reconsidered and differentiated in terms of periodic perceptual grouping on different levels. Periodic grouping can be distinguished from figural grouping. Both emergent phenomena stem from different cognitive mechanisms and have a backward influence on the sensuous stream of components. Literature on phenomenal accents, metric priming, and categorical rhythm perception is discussed, providing additional information on cognitive issues relating to memory and attention. Metrical accent, metric priming, and temporal categorization are interpreted as consequences of pulse sensation, grouping, and phenomenal accent.

Finally, section 2.4 examines some examples of musical complexity arising from the interaction of the perceptual and cognitive features discussed in the previous sections. It is indicated that these processes account for musical effects of metric ambiguity, malleability, and conflict, a topic which is further discussed and differentiated in chapter 3.

2.1 Time and rhythmic components

Properties found in an acoustic signal, can be subdivided into two types of determinants for rhythm perception. One type summarizes all aspects of perceptible cues and onsets, which act as markers for the other type, comprising temporal intervals and relations. Both have influence on the perceptual gestalt of rhythmic movement.

Rhythmic cues are produced by rapid transformations within the signal, articulations like tone onsets, attacks, or rapid changes of pitch, timbre, and so forth. Those attention drawing moments are associated with the perceived properties of their articulation, which tend to be categorized on the basis of similarity. Timbres can be grouped according to intuitive features like “bright”, “sharp”, or “dull”, or into more reflected categories like instrumental color. Pitches can be associated to pitch classes or by proximity, and so on. The other type of information comprises the temporal intervals, magnitudes, and relations which are generated by the temporal marks of the first type. This study primarily explores this latter aspect of rhythm by asking how temporal information is processed, and how it contributes to metric ambiguity, tension and malleability. Nevertheless, it is necessary to consider the interaction of both introduced aspects, as they always jointly occur. As mentioned, their interaction can cause complex perceptual effects.

2.1.1 Perceptual cues for rhythm and temporal segmentation

For rhythm to be gathered from a musical signal, it has to change and evolve over time in certain ways. For listeners, time is marked and segmented only when perceptible differences occur in the signal. To be perceived as rhythmic, successive or recurring distinctions have to be made in a certain range of temporal intervals (see section 2.2.2). Generally we talk about a rhythmic succession of events. Consequently, the notion of components of rhythm is based on the general notion of musical events. Every distinction which can be made, every musical event or phenomenal component of a musical signal can carry rhythmic meaning.[3]

Perceiving rhythmic order or character implies distinguishing properties of events or groups of events – functioning as differentiated and organized cues for rhythm – and their temporal alignment. FRAISSE discerns "patterns of time" and "patterns in time" to refer to this distinction.[4] PETERSEN claims that an analysis of the rhythmic meaning of any musical component has to focus on "duration".[5] The mentioned properties of events – like a bright or dark timbre, a bowed, blown, or sung tone – only get rhythmic in a temporal relation to other events with correspondent properties. Only the temporal intervals between events constitute the rhythmic meaning and interpretation of these events. Consequently, PETERSEN's idea of rhythmic analysis into "component rhythms" reflects that a rhythmic ensemble may be differentiated into rhythms of pitch, of timbre, of harmony, and so on (see figure 2.1). Successions of differences and changes in such components artistically embody successions of temporal intervals and therefore contribute to *the* rhythm of the music.

As shown in figure 2.1, the rhythmic analysis of music usually has to take several simultaneous layers of component rhythms into account. Often, different simultaneous patterns of temporal intervals are superposed, made up by different components, both in multi-voice and polyphonic music (as rhythmic contrast between voices) as well as within one voice. This resembles the notion of rhythmic stratification by YESTON.[6] As PETERSEN notes while discussing YESTON, the durations between recurring qualities[7] constitute rhythmic strata,[8] while metric strata are established as a byproduct of rhythmic movement.[9]

On the perceptual level, durations and temporal intervals, embodied and marked by components, induce qualitative feelings. Thus, they are not just quantities with gradual differences. They rather divide into categories with different perceptual qualities. In terms of rhythm perception, FRAISSE suggests two basic temporal categories: short

[3] cf. Petersen, 2010, p. 233, interpreting Messiaen's sentence "au Commencement était le Rhythme" (*Traité de Rhythme, de Couleur, et d'Ornithologie*, Tome I, p. VIII)

[4] Fraisse, 1982

[5] Petersen, 2010, p.234

[6] Yeston, 1976

[7] e.g. dynamic accents, cf. section 2.1.2 (second-order components)

[8] Petersen, 2010, p. 270

[9] Petersen, 2010, p. 271. Meter as a cognitive activity can indeed be characterized as induced by rhythm (cf. section 2.2.1). Parncutt, 1994, p. 410, correspondingly notes: "Integral to the theory [of Yeston] is the idea that a single rhythmic sequence can evoke several metric levels at the same time."

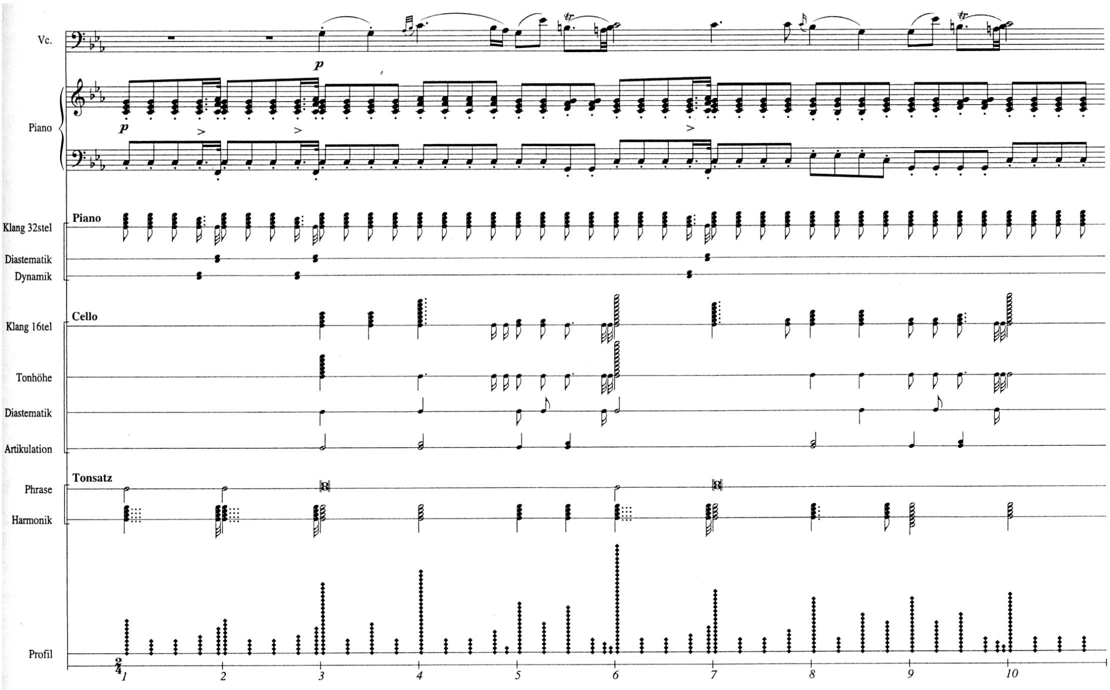


FIGURE 2.1: Example of a component analysis by Petersen, 2010, p. 253. Schubert, *Klaviertrio* E-flat major op. 100, beginning of the second movement.

times (200-300 ms) and long times (450-900 ms).[10] CLARKE and LONDON discuss the qualitative distinction implied in that categorization. Intervals of the long-times category induce a feeling of duration between marked time-points, whereas shorter intervals cluster together their component events, which are perceived as a collection.[11] Moreover, judgements about the quality of component events are related to their temporal relations. HURON refers to the psychological effect of "event-related binding" to reflect

> the tendency for a relational attribute to be mentally assigned to a single event. [...] The distance between two sound events is not some disembodied property, but a property of some object—the second tone.[12]

This stands in contrast to FRAISSE and other authors, who claim that temporal intervals between events are the crucial type of information for the cognitive processing of rhythm.[13]

To be clear about the notion of musical events and their role as markers for temporal

[10] Fraisse, 1982, p. 167. See also section 2.3.4 about the related cognitive principles of *assimilation* and *distinction*.

[11] See London, 2012, p. 34, emphasizing that "one of the functions of musical meter is to mediate among these qualitatively different durations by integrating them into a coherent attentional framework."

[12] Huron, 2006, p. 200

[13] Huron, 2006, p. 199

segmentation, note that a musical stream can be continuous and consist of only gradual changes in sonic quality. In other words, its rhythm can be too smooth to define concise temporal intervals. The traditional musicological discourse about rhythm nevertheless depends on discrete symbolic abstractions like notes, chords, etc. Some authors have thus proposed to describe rhythm as something continuous on a *subsymbolic* level.[14] Between the acoustic musical signal and its symbolic abstraction, a rhythmic microstructure is supposed, where nuances like smooth or pointed articulation and *expressive timing* (see section 2.3.4) refine or complicate the temporal structure.[15]

> Any factors that can create auditory boundaries can be used to create patterns in time that can be perceived as rhythmic.[16]

To support this statement, SETHARES gives a list of possible perceptual cues for rhythms "without notes".[17] He demonstrates various perceptual clues of auditory boundaries, keeping the loudness of the signal constant, including pitch-change, change of bandwidth while band-filtering noise, and relatively slow amplitude and frequency modulation. Hence, all these types of changes can have a rhythmic function and can separate musical "events".

As suggested, musical events are abstractions at a symbolic level and imply discrete elements with temporally precise beginnings and ends. On the acoustic or subsymbolic level, the matter is less clear. Research on the relation of acoustic features and perceptual boundaries address the notion of the perceptual onset of an event.[18] In the field of *Music Information Retrieval* (MIR), the development of computational methods for the detection of perceptual onsets in an acoustic signal is a complex and vast topic.[19]

The onset time of an event may not be the perceived moment of its rhythmic placement.[20] As the perceptual properties of sound-events are unique and complex, there may be sounds, for instance, with very sharp attacks where it seems to be easy to measure their perceptual onsets. Others may be more fluid and onsets would be more obscure or smeared in time, which may cause a stochastic distribution of perceived onset times. WRIGHT distinguishes between perceptual onset, perceptual attack and perceptual center, which usually occur successively in time while an event is passing.[21]

---

[14] See e.g. Desain, 1992, Moelants, 1997, Wright, 2008, p. 14, mentions "Eric Scheirer's influential paper [which] critiques models of music perception that proceed bottom-up via a stage that represents music entirely in terms of notes, which he terms the 'transcriptive metaphor'". (Scheirer, Eric David (1996). "Bregman's Chimerae: Music Perception as Auditory Scene Analysis". In: *Proceedings of the International Conference on Music Perception and Cognition*, pp. 317-322.)

[15] See Moelants, 1997, p. 270: variations in timing, amplitude and timbre contribute vitally to the artistic realization of a temporal structure, but they are difficult to investigate, as they are factors in a complex interplay.

[16] Sethares, 2007, p. 97

[17] ibid.

[18] See e.g. Vos and Rasch, 1981

[19] See e.g. Bello et al., 2005

[20] Wright, 2008, pp. 19 f.

[21] ibid.

While the *perceptual attack time* (PAT) is the perceived moment of rhythmic placement, the perceptual center plays a role in terms of stress, e.g. in speech perception and is more closely related to amplitude or energy maxima. The PAT of an event is thus the crucial cue for perceiving the rhythmic structure we are interested in. If we assume that the PATs are more or less precisely measurable, we can derive more or less precise temporal intervals between rhythmic events from their PATs. These may be called inter-attack-intervals. However, most research about temporal intervals relies on the concept of the *inter-onset interval* (IOI). As IOI is a common expression to refer to the relevant temporal intervals for rhythm perception, I will use it throughout this thesis as synonymous to "temporal interval", bearing in mind the outlined difference between perceptual onset and attack.

Obviously, a continuous musical stream is not always perceptually segmented in a way that implies rhythmic movement. "Onset and attack times are not meaningful for musical sound that is not perceived in terms of discrete events."[22] Also, there may be many cases where it is not clear if or when event attacks occur.[23] Thus, the abstracted representation of rhythmic structure on the symbolic level may not reflect many details of the particular rhythmic character we perceive.

The main concepts developed in this thesis are formulated at a symbolic level, and the introduced phenomena of metric ambiguity and malleability are discussed from an event-related perspective. Therefore, it is necessary to bear in mind the outlined downsides and limits entailed in a symbolic approach: it is abstracted from the aural signal and provides only limited access to, and restricted representation of precise temporal structures. On the other hand, symbolic abstraction corresponds to categorical perception, as discussed in section 2.3.4. Thus, the cognitive processing of rhythm can be regarded as a form of abstraction as well. Furthermore, notation and other musical communication methods reflect the cognitive dimensions of perception and memory on a discrete, symbolic level.[24] It may be nevertheless desirable to make findings – though developed on the symbolic level – accessible for research on other levels. On the one hand, the outlined perspectives and goals of MIR (for instance, onset detection for automated audio segmentation and for audio-to-score transcription of rhythm) are

[22] Wright, 2008, p. 21

[23] See e.g. Wright, 2008, pp. 14 f. andTanghe et al., 2005, p. 54, for comments on annotation problems of percussion performances: "some drum types and articulations are very difficult to annotate. Brushes for example have a typical 'dragged' sound which is hard to annotate as a single percussive event. In this case most annotators chose to register the accents of the brush sounds. Snare rolls do consist of a series of discernable percussive onsets, but it's very hard to annotate the many fast strokes accurately. The same is true for 'flammed' drums (typically the snare drum) where two hits of the same drum type are deliberately played almost (but not quite) at the same time, leading to the sensation of a ghost note occurring slightly before a main note."

[24] cf. section 4.1.2, and Toussaint, 2013, pp. 293 ff., who provides an illuminating description of principles of musical communication, focusing on the evolution, migration, and long-term tradition of rhythm. He argues that rhythmic communication gets more accurate and successful, when "the information transmitted across generations is not the rhythm itself, but rather the *instructions* for *creating* the rhythm." (p. 295, emphases in source)

aimed at transferring rhythmic information from the one level to another, towards an increasing degree of abstraction. On the other hand, it seems to be an open question whether knowledge, generated by the retrieval of musical information from symbolic structures, is useful for applications on less abstract levels.

2.1.2 Temporal intervals, components, and cognitive interaction

Duration, as a property of sound can only be perceived retrospectively: "*what cannot remain fixed and what cannot be determinate while the sound is going on is its duration.*"[25] The same holds for the durations of temporal intervals in a rhythmic context, which are not properties of sounds or of their symbolic event-abstractions. They correspond to the time elapsed between any musical articulation, repetition, or acoustic change of a component. Thus, in this study durations are treated as generalized temporal intervals, meaning that the acoustic and musical articulation of those intervals remains open for any component, such as dynamic accents, pitch change, and so on. As mentioned in the last section 2.1.1, temporal structure can be inferred from many possible musical movements, that is, components or changes thereof. Examples of different musical articulation of the same temporal interval can help to clarify the rhythmic impact of musical events (see figure 2.2).

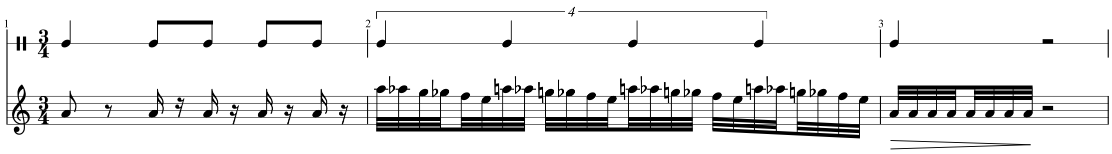


FIGURE 2.2: The same temporal structure can be articulated by different component structures.

The two voices in figure 2.2 basically represent the same temporal structure. In spite of different note durations, the rhythms of the two voices would usually be interpreted as identical in bar 1, differing only in articulation, as we primarily regard note onsets or attacks as rhythmic cues. Thus, rhythmic structure is mainly conveyed by temporal intervals between musical articulations. As indicated, the quality of the articulations refer to components such as pitch, timbre, or articulation types like staccato, legato, et cetera. Accordingly, the schematic rhythm of the upper voice is paralleled in the lower voice also in bar 2, though differently articulated by means of pitch structure. A chromatic scale runs downward in six thirty-second steps, before it is repeated three times. The quarter-quadruplet (or dotted-eighth) intervals are clearly embodied by the scale repetitions, and the intervallic jumps in between function like the simple rhythmic onsets in the upper voice. These jumps can also be regarded as *melodic accents* which take

[25] Hasty, 1997, p. 93 (italics in source)

rhythmic function, as further discussed below and in section 2.3. Therefore, changes in pitch, timbre, or loudness can be perceived as rhythmic events, which are represented by relations between different events. I will refer to them as rhythmic components of higher order. Musical works are full of higher-order components, forms like figuration or sequences, which generate rhythm on different levels. They form rhythmic shapes and set out temporal boundaries and marks, which can be regarded as temporal or rhythmic information. The notion of rhythmic duration should thus be detached from the duration of single events.

> The detachment of the concept of duration from single sound sources is required for a consistent description of all temporal structures in the rhythmic dimension, that is, the recognition and analysis of component rhythms.[26]

Hence, temporal relations between different musical events and groups of events specify the rhythmic shape of music. The properties of those events constitute further dimensions, which contribute to the overall musical gestalt.

To examine only temporal interval patterns as the crucial rhythmic information, is a common means to explore rhythm perception abstracted from the multidimensional perceptual reality of musical rhythm. The analytic necessity to abstract from the complexity of component interplay is problematic, as it is a potential source of the type of ambiguity studied here.[27] Nevertheless, restricting the research perspective to the more abstract level of temporal patterns, still bears sufficient potential to study metric ambiguity and malleability, as extensively investigated in the context of rhythm perception.[28] This area still invites further inquiry, despite the many limits already mentioned. However, as metric ambiguity is a cognitive phenomenon, we have to be aware of the caveats of a symbolic approach.

> Real music is complex. One musical component can change over time within a single piece in ways that highlight certain components but not others. Rhythm can, for example, enhance a melodic line in some phrases and obscure it in others.[29]

JONES gauges the area of conflict between the need for ecologically valid research to use "real music as stimuli for experimental investigation"[30] and the limitations on precise understanding of determinants for certain listener responses. Though we are interested

[26] Petersen, 2010, pp. 251 f. (my translation)

[27] See, for instance, the related discussions of *joint accent structure* in sections 2.3 and 2.4, and of *tonal accents* in section 2.4.

[28] For instance, in Flanagan, 2008, Handel and Oshinsky, 1981, Handel and Lawson, 1983, Handel, 1984, Longuet-Higgins and Lee, 1984, Parncutt, 1987, Parncutt, 1994, Povel, 1984, Vazan and Schober, 2004, Volk, 2004, Vuust and Witek, 2014, Yu, Getz, and Kubovy, 2015

[29] Jones, 2010, p. 7

[30] ibid.

primarily in rhythm perception, we must also keep track of the complex interplay between components. Looking for "a strategy that allows some insight into musical complexities while retaining control over variables that contribute to this complexity",[31] we would need a heuristic to correlate analytical data from simple stimuli in regard to a hypothetical context.

Any analytical research on rhythm has to abstract to a certain degree from the particularity of musical experience. No isolated component can be extracted from that experience, but if we analyze isolated components, they "could be taken as a sensible distinction that contributes to the definiteness of an aural experience."[32] We can test, with HASTY, by "asking if there is some effect or feeling that could be imagined to result from the distinctions an analytic object hypothesizes."[33]

Hence, patterns of IOIs can only be analyzed after they were derived from their materialized components in a particular musical context. On the other hand, they can be synthesized first and then employed as the temporal structure of a musical sequence, or rather, of a component thereof. HASTY claims that "a feeling of duration is always a feeling of particularity."[34] His notion of metric particularity especially characterizes meter as something immediate and sensuous, always renewing itself, contradicting the common notion of meter as an abstract and mechanic grid underlying the rhythmic flow. The recurrence of correspondent metric units in successive metric cycles is not equivalent to the reoccurrence of the same particular feeling. Particularity is also tied to the complex interplay of components: "differentiation provided by tonal quality and contour [...] can create abundant opportunities for durational correspondence (whether in conformance or contrast)".[35] Induced by the interplay of components, rhythmic and metric particularity involves as well "our attentiveness and interest."[36] More generally, we are influenced and constrained by our cognitive system when we perceive rhythm. The interaction between perception and cognition involves top-down processes, in that cognition, fed by perceptual input, retroacts on perception.

Thus, the way components and event properties contribute to rhythm perception, cannot be understood as a one-way causal relation. In fact, a number of specific cognitive peculiarities influence the processing of components. The immediate temporal context has influence on the perception of acoustic properties of components. An event may be distorted by another immediately following event, as the processing of strength or loudness takes time and can be interrupted.[37] An established metric framework, induced by the perception of component rhythms in the first place, produces as well a

[31]ibid.
[32]Hasty, 1997, p. 155
[33]ibid.
[34]Hasty, 1997, p.150
[35]Hasty, 1997, p. 152
[36]Hasty, 1997, p. 150
[37]cf. sections 2.2.2 (footnote 149) and 2.3.2 ("echoic" memory).

backward effect on the perception of the components (section 2.3.3). Such capacities also have profound impact on processing temporal structures, and are consequently further discussed in the subsequent sections of this chapter.

> A range of acoustic cues, including frequency, duration, and amplitude (intensity) have the potential, depending on their patterning, to convey to the listener a sense of inherent sequence organization or structure.[38]

Hence, component patterns and their temporal structures interact with rhythmic *grouping*, the perceptual segmentation of an auditory stream into groups of events. Elements between group boundaries cluster together "to form a temporal unit".[39] The perception of groups intertwines with the perception of accents on particular elements within a group, typically on the first or last event. Grouping and accentuation seem to emerge as perceptive correlates of a cognitive organization process. This will be more extensively discussed in section 2.3. In the current context, it should yet be noted that empirical studies demonstrate the influence of particular components on grouping, such as for "pitch cues, there is a tendency to perceive the rhythmic organization of sequences according to repeated pitch patterning."[40] Obviously, events can as well be perceived as accented because of their properties, or rather, as a result of property relations in component patterns. The type of accent is usually related to the type of event property, for instance, a dynamic accent is caused by a louder event, related to its context. In section 2.3, such accents are classified as subgroups of *phenomenal* accents, defined by LERDAHL and JACKENDOFF.[41] They "may be induced by changes in IOI, loudness, timbre, and pitch and by combinations thereof."[42] From the perspective of the present study, it is most interesting that grouping processes are complex cognitive interactions with component properties and may lead to surprising and ambiguous perceptual effects. DEUTSCH analyzes these effects on the level of single parameters of "tonal percepts" which may "be described as a bundle of attribute values."[43]

> in situations where more than one tone is presented at a time, these bundles of attribute values may fragment and recombine in other ways, so that illusory percepts result. Perceptual grouping in music is therefore not simply a matter of linking different sets of stimuli together; rather it involves a process whereby these stimuli are fragmented into their separate attributes, followed by a process of perceptual synthesis in which the different attribute values are recombined.[44]

---

[38] McAuley, 2010, p. 183
[39] Patel, 2008, p.106
[40] McAuley, 2010, p. 185. See also section 2.4 on repetition and *parallelism*.
[41] Lerdahl and Jackendoff, 1981, p. 485: "By *phenomenal accent* we mean any event at the musical surface that gives emphasis or stress to that moment in the musical flow." (italics in source)
[42] Parncutt, 1994, p. 426 (written in italics in the source)
[43] Deutsch, 1982, p. 101. As attributes of tones, Deutsch lists pitch, loudness, timbre and spatial location.
[44] ibid.

Well-known experimentally proved tonal illusions, for instance, include the scale illusion, exemplified in figure 2.3.[45] Here, frequency proximity is a much more powerful cue for organizing a polyphonic structure than spatial proximity.

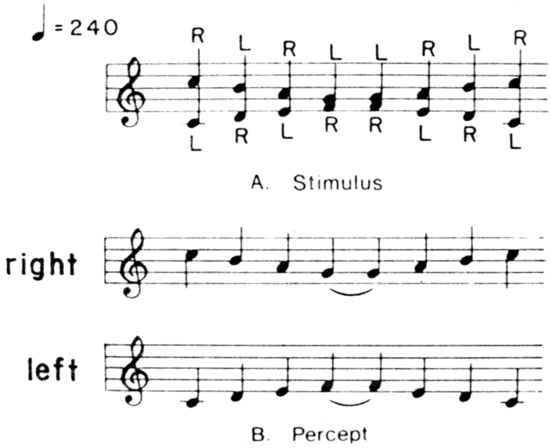


FIGURE 2.3: Deutsch, 1982, p. 102: "(A) Representation of the configuration producing the scale illusion. This basic pattern was repetitively presented 10 times without pause. (B) Representation of the illusory percept most commonly obtained".

Grouping mechanisms occur – and interact with component patterns – in diverse areas. For instance, music and speech "share a number of acoustic cues for marking phrase boundaries."[46] PATEL therefore assumes shared cognitive processes for grouping across domains and discusses empirical evidence for his claim.[47] In musical rhythm perception, such general cognitive mechanisms interact with more specific types of grouping. This is further discussed in sections 2.3 and 3.2.1, where the more general principle of *serial grouping* is distinguished from *periodic grouping*. The latter emerges in the context of metric entrainment, featuring other related processes such as *pulse sensation*, *expectation* and *anticipation* (see section 2.2). HURON distinguishes musical expectations into "*what*-related expectations" and "expectations about *when* events may occur."[48] This corresponds to the distinction of event qualities and their temporal relations and occurrences. HURON's differentiated taxonomy of expectation distinguishes several degrees of internalization, from anticipation of reoccurrences of qualities, just stored in short-term memory, to long-term abstract knowledge of and familiarity with templates and schemas of component patterns.[49] Such cognitive top-down strategies tend to balance the perceptual input in favor of an integrated viable shape.

[45]Deutsch, 1982, p. 102. The original work is from Deutsch, D. (1975). "Two-channel listening to musical scales". In: *Journal of the Acoustical Society of America* 57, pp. 1156–1160.
[46]Patel, 2008, p. 112
[47]Patel, 2008, pp. 159 ff.
[48]Huron, 2006, p. 175 (emphases in source)
[49]cf. the following section 2.1.3, as well as sections 2.2.1, 2.3.3, and 2.4.

2.1.3 Component hierarchies

Musical rhythm can emerge from compound melodies: a homogeneous musical line or voice is amenable to be perceptually segregated into different layers of immanent voices. In other words, a pattern of melodic accents (see sections 2.3 and 2.4) can constitute a component rhythm which is hierarchically dependent on an underlying melodic stream. More generally, the temporal patterns of component rhythms consisting of any kind of accents may be perceived as superordinate rhythmic layers. Unaccented events are perceptually grouped in between adjacent accents (section 2.3). Compound melodies similarly arise from melodic accents, melodic contour, pivotal pitches or dynamic, harmonic or metric stresses.

> we are able to hear compound melodies, and hence perceive a series of different durations within a musical surface consisting of even articulations[50]

In other words, rhythms of compound melodies or immanent voices emerge from surface rhythms, if pitch structure or other component patterns give rise to them. The homogeneous stream of eighth notes in figure 2.4 (a) contains melodic jumps over an octave and more. The auditory effect (b) is that of a *hocket*: two voices exhibiting complementary rhythms. Pitch proximity plays a central role here, as more generally for processes of stream segregation in homogeneous rhythmic surfaces.[51]

FIGURE 2.4: Compound melody after London (Petersen, 2010, p. 291)

When component patterns are repeated or closely variated, another instance of rhythmic hierarchy may become sensible. In their *metrical preference rule 1*, LERDAHL and JACKENDOFF propose that the perceived grouping and meter[52] tends to assimilate to parallel structures.[53]

[50] London, Justin (2000). "Rhythm". In: *The New Grove Dictionary of Music and Musicians*, 2nd Ed., Vol. 21, p. 279 (cited in Petersen, 2010, pp. 290 f.)

[51] cf. Deutsch, 1982, and the previous section 2.1.2.

[52] The perception of grouping and meter and their relation will be discussed in depth in the course of this study.

[53] See also Lee, 1991, pp. 71 ff., and Temperley and Bartlette, 2002

> MPR 1 (Parallelism) Where two or more groups of parts of groups can be construed as parallel, they preferably receive parallel metrical structure.[54]

Accordingly, repetitions and variations of short rhythmic or melodic motives can impose their boundaries on grouping, and induce as well the feeling of rhythmic depth. Like compound melodies and accent rhythms, patterns of perceptually confined motives constitute patterns of temporal intervals, possibly perceived as additional rhythmic layers. Notions of the most rhythmically relevant types of repetition and parallelism in music (which is further discussed in section 2.4) include motivic ostinati, figuration and sequences. Beside the described perceptual properties, ostinati and figuration influence hearing on more abstract cognitive levels, as a function of internalized knowledge.

> Figuration is often shared between works. [...] When a figuration pattern is commonplace, a listener may begin to experience it as an independent schema.[55]

HURON exemplifies this with a typical accompaniment ostinato from MOZART (Figure 2.5). Note that a figure can vary to a certain extent without loosing its figural identity. The left-hand ostinato in Figure 2.5 changes in pitch but not in melodic contour and rhythm. The combination of the latter two aspects is perceived as the figural identity in this case. This kind of flexibility allows for a patterning of figures or ostinati into sequences. A musical sequence may outrun the rhythmic domain: such a technique is used to create larger formal units. Other components may play a role as well, such as harmonic progressions within sequences of harmonic patterns or chord arpeggios. Repetition can also enhance grouping and stream segregation by frequency proximity (see the previous section 2.1.2), as an accumulation of evidence over time.[56]

FIGURE 2.5: W. A. MOZART, *Sonata in C*, K. 545, first movement (cf. Huron, 2006, p.256)

In section 2.4, some effects of the described higher-level rhythmic components are surveyed, which involve metric ambiguity and metric conflict.

[54]Lerdahl and Jackendoff, 1983, p. 75
[55]Huron, 2006, p.256
[56]cf. Deutsch, 1982, pp. 120 ff.

2.2 Pulse sensation

The sensation of a pulse involves the feeling of recurrence, of a steady and approximately even movement, that is, a cyclic or periodic movement. We can measure time by counting pulses which are perceived as equal temporal units. Hence, pulse can as well be regarded as a sensation of meter in a basic sense. Meter perception in a more elaborated sense can be hierarchically organized as the sensation of simultaneous pulses in a harmonic relation. This happens for instance, when every second pulse in a homogeneous pulse sensation – symbolized by [x x x x x x x x...] – gets more attention than the preceding or succeeding pulse, resulting in a simple metric hierarchy: [X x X x X x X x...] (see section 2.3.1).

In the present section, pulse sensation is examined in the context of the cognitive flexibility, which accounts for metric malleability. In the later course of this study, specific musical challenges for this ability are then explored in terms of metric conflict and related phenomena. In section 2.1.1 we discussed acoustic cues, which can provoke perceptual segmentation of a musical signal into temporal intervals. In contrast, time is subjectively marked by pulse perception, and subdivided into isochronous periods on a cognitive level. Successive time spans are perceived as equal durations. Pulse thus implies categorical perception: slightly unequal durations may be judged as isochronous, regular pulses, up to a certain, contextually influenced limit of flexibility.[57] Deviations from perfect regularity within this limit may be interpreted as expressive forms of rhythmic *entrainment*.[58] Pulse sensation and entrainment are aspects of a dynamic process. A cognitive pattern is initially excited, which guides the interpretation and contextualization of the succeeding rhythmic flow, further enhancing or contradicting the pattern. Regularly recurring cues are not necessary to induce this process. We are also able to detect regularity in complex rhythms by pattern recognition, a more general cognitive ability, which is verified for different senses.[59] Hence, the relation of a pulse sensation to an acoustic stimulus can be a complex matter. This is further examined at the symbolic level in section 2.2.3.

> Musical pulse sensations are generally illusory in that they do not necessarily indicate the presence of single periodic sound sources.[60]

In the following sections, the "illusory" character of pulse sensations is differentiated as the sensitive correlate of multileveled cognitive processes, which interact in both bottom-up and top-down directions. Top-down processes have a particular impact, as grouping processes (section 2.3) are also influenced by schemes which are acquired through long-term exposure and musical experience. While the most recent cognitive

[57] Madison and Merker, 2002, see also section 2.3.4.
[58] London, 2012
[59] Parncutt, 1994, p. 433, Toussaint, 2013, pp. 198 ff., surveys musical pattern recognition in the context of gestalt psychology, including concepts like the *primal sketch* in the theory of visual perception.
[60] Parncutt, 1994, p. 433

theories on this topic – for instance, *predictive coding theory* – involve neuronal processing,[61] former approaches were often inspired by principles from gestalt theory.[62] Cognitive activity, involved in pulse sensation, is temporally oriented in two directions. Entrainment involves the projection[63] of expectations for future events. When expectations are satisfied, for instance by a match of a rhythmic event with a projected pulse, it may be reinforced as a backward effect in memory.[64] Finally, these properties of cognitive behavior allow for immediate rhythmic synchronization to musical processes which evoke pulse sensation.

2.2.1 Entrainment and dynamic attending

The exploration of general cognitive abilities and activities, which underlie pulse sensation and meter perception, is reflected by the development of diverse theories and models. These may be summarized into two main types.[65] *Interval theories* are based on an information-processing framework, where pulses of an "internal clock" are counted to estimate intervals between incoming events. The system can reduce processing effort, as interval durations can be stored in memory and used as expectations for current timings. *Entrainment theories* assume that rhythm is attended by a dynamic cognitive process, and particularly relate to the *Dynamic Attending Theory* (DAT) by JONES et al.[66] DAT "is a generalization of entrainment theory, whereby the internal driven rhythm is conceptualized as an attentional rhythm".[67] Entrainment can as well be considered as a source of pulse sensation, and, more generally, DAT is convincingly applied to musical rhythm and meter perception.[68] Perceived temporal regularities are translated into an internal cognitive pattern in time. In other words, pulse perception is a response to a musical pattern by means of a synchronized pattern of attention.

Human entrainment to rhythm can be regarded as a special case of the natural phenomena of entrainment and resonance, for instance of coupled pendula.[69] It is also embedded in our general, temporally oriented attentional behavior, as we are ecologically challenged to anticipate future sensory events. Specialized musical anticipation and

[61] cf. Vuust and Witek, 2014

[62] Deutsch, 1982, p. 101, pp. 126 ff., shows how events can be "perceptually replaced" while they are missing in the acoustic stimulus and calls this an "auditory continuity effect". See also Rosenthal, 1992, p. 75, and Parncutt, 1987

[63] cf. Hasty, 1997

[64] Desain, 1992, pp. 444 f., claims that *expectancy* is coupled with *retrospective reinforcement*, as further discussed in the following section 2.2.1.

[65] McAuley, 2010, pp. 168 f.

[66] cf. for instance Jones and Boltz, 1989, Barnes and Jones, 2000

[67] McAuley, 2010, p. 171

[68] For instance London, 2012, see Vuust and Witek, 2014, p. 2

[69] cf. McAuley, 2010, p. 170, and Sethares, 2007, p. 148: "In 1665, the Dutch scientist Christian Huygens [...] noticed that when two clocks were mounted near each other on a wooden beam, the pendulums began to swing in unison."

synchronization abilities stem from those broader skills.[70] Thus, as rhythmic entrainment facilitates synchronization, it implicates and enables attentional synchronization to rhythm. This process involves the detection of temporal regularities in a rhythm, and guides adaptation to those regularities.

Pulse sensation implies a periodic attentional pattern and is often accompanied by spontaneous use of rhythmic body movements.[71] The notion of entrainment covers these aspects as parts of an embodied cognitive activity. As in empirical studies about pulse perception, bodily activities are often "more pronounced among participants with more musical training".[72] The "use of the body in time keeping should not be regarded as a primitive expedient, but may be an intrinsic part of human entrainment to isochronous stimulus trains."[73] Furthermore, entrainment enables complex musical coordination.[74]

The close relation of attentional entrainment and motor responses to isochronous sequences (which are apt to induce pulse sensation) can be verified on the neural level. CHEN et al., using fMRI, show that the functional connectivity between auditory and dorsal premotor cortices correlates with rhythmic interaction of the auditory and motor system. They suggest, that the "activity in the dPMC may represent the integration of [...] auditory information with temporally organized motor actions."[75]

To focus on the specific processual implications of pulse sensation, it is noteworthy that anticipation and synchronization emerge very quickly and inextricably from the feeling of pulse.[76] As motor anticipation is required to bodily synchronize with a rhythm, pulse sensation implicates perceptual anticipation. The synchronization of a perceived regularity with an internal attentional pattern causes expectations about what may follow: the synchronized regularity is expected to continue. Expectation makes possible and simplifies the anticipation of future events and continuation of the temporal structure of synchronization. Thus, the whole process tends to be circular and self-enhancing, as anticipation facilitates, enhances and accelerates synchronization on both cognitive and motor levels.

Varying explanations can be found in the literature about the details of the induction

[70] London, 2012, pp. 10 f., Huron, 2006 (see section 1.2.1, footnote 48)

[71] cf. section 1.2.1

[72] Madison and Merker, 2002, p. 206

[73] ibid.

[74] cf. London, 2008: "'Musical entrainment' crucially involves the coordination of entire metric cycles, that is to say, of matching the downbeats (as well as the overall beat pattern) [...] so that a complex motor action (e.g., a complete rhythmic figure) may be produced in time with others".

[75] Chen, Zatorre, and Penhune, 2006, p. 1779. Moreover, "metric organization (via intensity accentuation), modulates motor behavior and neural responses in auditory and dorsal premotor cortex. Auditory–motor interactions may take place at these regions with the dorsal premotor cortex interfacing sensory cues with temporally organized movement." (p. 1771)

[76] See for instance Fraisse, 1982, Hasty, 1997, Noorden and Moelants, 1999, Parncutt, 1994.

process and its celerity. PARNCUTT already regards a pair of events as a sufficient perceptual input for a pulse entrainment.[77] FRAISSE reports experimental evidence for this assumption: an isochronous sequence of events could be accompanied from the third sound on. This also works with more complex rhythms: repetitive rhythmic figures could be accompanied from the second repetition ("third pattern") on.[78] Such observations support the assumption that it is not events (respectively onsets), but the IOIs between them, which constitute the relevant information to be processed in preparation for the reaction.[79] If two events are perceived, a third event will be expected to follow most likely by the same interval after the second event as the interval between the first two events.

A theory by DESAIN[80] is based on the same assumption. It generally describes rhythm perception as an activity of extrapolating a perceived temporal structure into the future, by a complex pattern of expectations. This pattern is displayed as a function of *expectancy*. It gets modified as soon as there are new temporally relevant events perceived. According to PARNCUTT and FRAISSE, DESAIN suggests that processing a single IOI already causes an expectation.

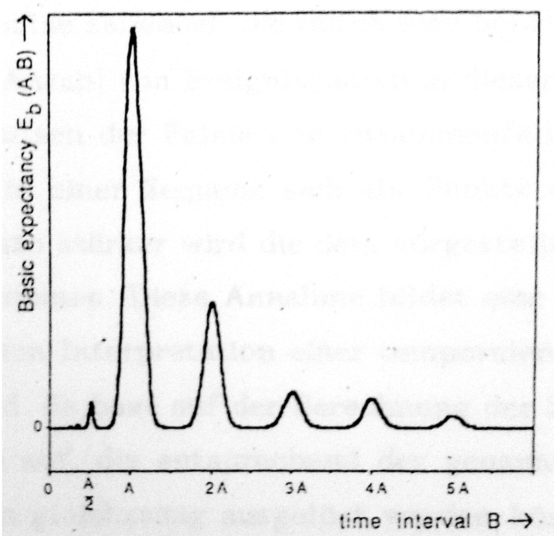


FIGURE 2.6: Basic Expectancy (Desain, 1992, p. 447)

Figure 2.6 displays the temporal structure of a *basic expectancy* as a function of the ratio between two IOIs.[81] If interval A is perceived between two events, interval B is expected between the second, currently last perceived event, and a third expected event.

[77] Parncutt, 1994, p. 434
[78] Fraisse, 1982, p. 155
[79] In contrast, Huron, 2006 argues in the opposite direction, as mentioned in section 2.1.1 (cf. "event-related binding").
[80] Desain, 1992
[81] The basic expectancy function is defined by the sum of several Gaussian curves, building up around every relatively simple IOI ratio.

The expectancy level for the third event reaches its maximum when the two intervals between the three events would have a 1:1 ratio. Higher integer ratios are assumed to be expected as well but on successively lower levels. This implies an anticipation of further pulses, the expectation of a continuing pulse, as induced by two initial events.

DESAIN regards expectancy as a consequence of an attentional process, unfolding bidirectionally in time. Peaks of attentional energy leave memory traces, and thus, events coinciding with those peaks may also be better remembered. Moreover, this can lead to a "reinforcement of past events by new data."[82] Hence, the concept of expectancy "has no time direction in itself."[83] A consequence of this bidirectionality "is that the sum of the corroborations of each event in a pattern by a virtual new onset is the same as the expectancy of that onset generated by the pattern."[84] In DAT, the intrinsic role of short term memory in entrainment processes is similarly specified, as working memory "is captured by the current period of the oscillator; this provides an internal estimate of sequence rate, a running memory of the sequence's time intervals."[85]

Pulse sensation and metric entrainment indicate a dynamic cognitive activity which creates a particular mental structure of time. Under that condition, rhythmic figures are perceived as molded into a temporal scheme. Conversely, rhythmic figures can also trigger, enhance or challenge the process of entrainment. The notion of metric malleability reflects that a rhythm can be perceived within different metric frameworks, that is, it can induce different entrainment behaviors. The concept of entrainment thus adds a dynamic perspective to the rather static notion of a framework. Metric ambiguity and conflict can emerge from this dynamics. If a dynamic process of attending reaches a stable pattern of entrainment, it imposes this pattern on what follows. On the other hand, if the following input challenges the pattern, it can get instabe, or the pattern can even dissolve. However, particular behaviors accompanying entrainment, allow smooth or resistant cognitive processes, even when complex and conflicting metric situations arise (see section 2.3.3).

While pulse sensation can be regarded as a basic form of dynamic attending, entrainment can be driven from complex situations as well. The character of entrainment may be differentiated in situations of different complexity, but there are some important common features and constraints.

[82]Desain, 1992, p. 439. This concept is influenced by M.R. Jones, regarding expectancy and memory as closely related. According to Jones, remembering can be considered "as a dynamic attentional process unfolding in negative time." (p. 444, quoting p. 574 of Jones, M.R. (1981). "Only time can tell: On the topology of mental space and time". In: *Critical Inquiry* 7, pp. 557–576) Desain, 1992 consequently proposes to estimate "the influence that a new incoming event might have on the prior context. The new event can support one or the other of the previous interpretations and in retrospect contribute to a limited extend to the salience of already perceived stimuli." (pp. 444 f.)

[83]Desain, 1992, p. 445

[84]Desain, 1992 p. 449

[85]Barnes and Jones, 2000, p. 263

- *response delay*: as discussed, the ability of immediate rhythmic synchronization indicates that patterns of entrainment emerge quickly. Nevertheless, ambiguous situations and conflicting cues can defer entrainment to different degrees.
- *flexibility*: entrainment is adaptive to specific properties of rhythmic input. It may emerge on different levels of temporal regularity, and displays a certain elasticity in resonating with complex rhythmic stimuli.
- *periodicity range*: rhythmic entrainment only responds to periodicities within a limited temporal window – according to common estimates between 100 ms and about 5 seconds.[86] Around the logarithmic mean of this range (ca. 600 ms), a peak of entrainment sensitivity correlates with the phenomena of preferred tempo and spontaneous tempo. Section 2.2.2 provides further information on these issues.

A certain flexibility of entrainment can be observed both in ecologically valid settings of listening to music, and laboratory experiments with exactly designed stimuli. For instance, BUCHLER introspectively examined modes of rhythmic entrainment, cued by compositions of John ADAMS. He concludes that, even while listening to complex temporal irregularities, entrainment still can lead subjects to feel a sense of meter. When "something that sounds metrical fails to meet our strictest definitions of meter, we ought to begin questioning our definitions and/or examining the elasticity of our metrical fabric."[87] MADISON and MERKER investigated temporal constraints of pulse sensation and recognition of irregularity in sequences of identical sounds.[88] In judgements about 10-event sequences with 570- to 630-ms nominal IOIs they found an area of tolerance between 3.5% and 8.6% of temporal deviation: within such sequences, although irregularity is already noticed, pulse sensation is still felt.[89] They

> assume that the experience of pulse is closely related to the ability to synchronise, and that it should, therefore, not be rigidly dependent on physical isochrony. On the other hand, too large a tolerance will be inefficient by leading to predictive imprecision and wastage of resources on seemingly useful information which turns out not to allow any prediction at all.[90]

Anticipation and prediction are thus functionally connected to pulse sensation and entrainment. They allow flexible synchronization within the mentioned limits. BARNES and JONES propose "two aspects of stimulus-driven attending, namely attentional pacing and temporal capture".[91] The first refers to the interplay of an entrainment activity

[86] cf. London, 2008: "Within this range there are several sub-ranges in which our entrainment characteristics [are] different (these are 100-400ms, 400ms to 1 second, and 1-2 seconds), and as a result our rhythmic perception and performance is quantitatively and qualitatively different in these ranges." See also section 2.2.2 and London, 2012, pp. 27 ff.

[87] Buchler, 2006, p. 709

[88] Madison and Merker, 2002

[89] Other measurements involving contextual hearing, suggest considerably higher tolerances under specific conditions, see section 2.3.4

[90] Madison and Merker, 2002, p. 201

[91] Barnes and Jones, 2000, p. 301

with the expectation of immediate continuation of that pace, as described above. The second aspect addresses flexible adaptations "to violations of these expectancies."[92]

Experimental evidence for adaptation is provided in several studies of synchronization performances, where participants are asked to tap regularly along structurally isochronous or more complex, metrically structured sequences which comprise temporal perturbations or timing deviations.[93] Two aspects of adaptation can be distinguished, namely: period and phase adaptation. A phase shift corresponds to a timing perturbation which changes dynamic attending at a relatively small scale, and does not affect the period of an attentional pulse. In contrast, changing periodicities in a sequence must be tracked more extensively, as the pace or tempo affects a sequence on a larger scale. Phase adaptation is found to be performed much more rapidly than period adaptation.[94]

Some specific biases are commonly found in this context, concerning timing mechanisms in rhythm production. REPP et al. report biases in perception, and production of cyclically repeating two-interval rhythms, depending on the ratio between the two IOIs: "rhythms with small ratios were produced with increased ratios, and timing perturbations in these rhythms tended to be harder to detect when they locally increased the ratio than when they reduced it. The opposite held for rhythms with large ratios."[95] This seems to contradict the principles of *assimilation* and *distinction* in rhythm perception and production (section 2.3.4), examined by FRAISSE.[96]

REPP et al. proved that synchronization to complex IOI ratios can naturally be accomplished, as such temporal complexities systematically occur in musical performance timing.[97] For instance, BARALDI et al. measured rhythmic timing and synchronization, between musicians performing a local Transylvanian variant of an *aksak* rhythm. Aksak rhythms are based on two duration units in a ratio commonly conceptualized as 2:3 ("orthodox") or 3:4 ("heterodox"). They found that the ratio is variably performed in between the two concepts, although the musicians had no trouble to synchronize to each other in a flexible way.[98]

[92]ibid.

[93]See e.g. Large, Fink, and Kelso, 2002, Repp, London, and Keller, 2011

[94]Barnes and Jones, 2000, Repp, London, and Keller, 2011

[95]Repp, London, and Keller, 2011, p. 227

[96]Fraisse, 1982, p. 167

[97]Repp, London, and Keller, 2011, p. 239

[98]Baraldi, Bigand, and Pozzo, 2015. This fact "may suggest that the musicians conceive two 'blocks' of durations (S and L), which are largely independent one from the other, that is, they do not rely on a common underlying pulse. In other words, the aksak rhythm – or at least, this Transylvanian version of aksak – should be regarded in terms of two independent duration units, as proposed by Brăiloiu (1973), rather than in terms of a unique smaller subpulse, as suggested by other scholars such as Arom (2004). Nevertheless, comparison with other studies should be advanced carefully, since we are in presence here of a slow tempo aksak, less studied than the fast tempo aksak rhythms analyzed by these authors." (p. 275, referring to Arom, S. (2004). "L'aksak: Principes et typologie". In: *Cahiers de Musiques Traditionnelles* 17, pp. 11–48; and Brăiloiu, C. (1973). *Problèmes d'ethnomusicologie*. (A. L. Lloyd Trans.). *Problems of ethnomusicology*. Cambridge: Cambridge University Press, 1984.)

A common type of entrainment models describe the flexibility of an attentional pulse by the behavior of an adaptive oscillator.[99] "In principle, an entraining oscillator can modulate both its period (*period adaptation*) and its phase (*phase adaptation*) in response to an unexpected event (a perturbation)."[100] To a certain extend, oscillator models coincide with DESAIN's model of expectancy.[101] An oscillator involves expectation through its periodic activity which can be projected to anticipate the following. Hence, adaptation involves the modification of an "expectancy profile"[102] by adjusting the mentioned parameters of period and phase. In this way, an oscillator instantiates an attending rhythm by "entraining, i.e., 'locking into' the ongoing time structure."[103]

> The perception of meter from an entrainment perspective requires a multiple-oscillator model in which each oscillator corresponds to each metrical level.[104]

Musical meter can be described as a complex, hierarchically organized, form of entrainment.[105] LONDON identifies both entrainment and meter as cognitive frameworks for rhythmic perception and performance.[106] Hence, the entrainment perspective contributes to the understanding of meter as a subjective, and therefore potentially ambiguous phenomenon.

Hierarchical entrainment can be dynamically established in different depths. A rhythm can accordingly embody more or less metric levels. LONDON gives such an example which is shown in figure 2.7. Two melodies represent different depths of a metric hierarchy: the dots beneath the systems indicate metric levels in the style of LERDAHL and JACKENDOFF.[107] Melody (a) articulates beat and measure levels, whereas melody (b) additionally renders a layer of beat subdivision. LONDON argues that time signatures do not fix the type of meter. Rather, since examples (a) and (b) "give rise to dissimilar forms of entrainment, they are in fact *different* meters."[108]

---

[99]Oscillator models of metric entrainment relate to JONES' theory of dynamic attending (DAT). For an instructive overview, see for instance Sethares, 2007, pp. 147 ff.

[100]Barnes and Jones, 2000, p. 293 (emphases in source)

[101]The model proposed in Desain, 1992, is not explicitly based on the notion of entrainment, but, in my opinion, it fits to the dynamic attending account of expectancy. Another coincidence between both approaches is the representation of time as a continuous function. Reconsider also the relation between the concepts of memory in expectancy theory and DAT, as discussed above.

[102]ibid.

[103]Barnes and Jones, 2000, p. 261

[104]McAuley, 2010, p. 191

[105]London, 2012, London, 2008, Buchler, 2006, McAuley, 2010

[106]London, 2008, p. 2 f: "In many [...] musical behaviors, entrainment involves a coordinated *set* of entrainments, some of which may be in 1:1 relationships with the driving oscillator(s), but with other ratios necessarily obtaining between other component periodicities. [...] Like entrainment, meter can be regarded as the temporal framework which guides our listening and performance. It is subject to the same perceptual and cognitive limits as entrainment more generally. It is hierarchically structured, with different performance and perceptual attributes accruing to different structural levels (i.e., subdivisions, beats, and measures)." (emphasis in source)

[107]Lerdahl and Jackendoff, 1983

[108]London, 2012, p. 17 (emphasis in source)

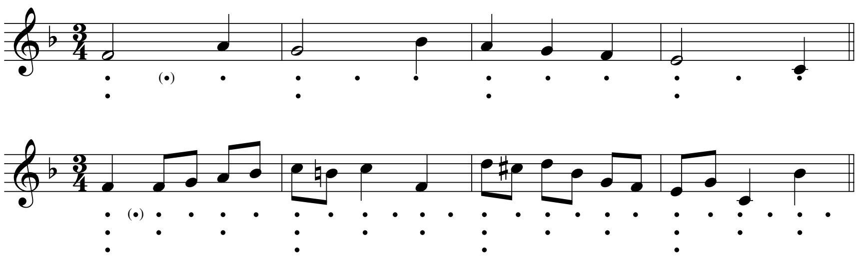


FIGURE 2.7: Different metric depths (London, 2012, p. 18)

Accordingly, DAT assumes dynamic and flexible forms of entrainment to metric hierarchy. Metric levels are "thought to vary in salience in relation to each other".[109] LARGE et al. provide experimental "evidence for a dynamic and flexible internal representation of [...] metrical structure."[110] Synchronization to random, metrically structured rhythmic sequences, can be accomplished by tracking the metrical hierarchy, even in spite of tempo (period) and phase perturbations. Subjects were asked to tap according to three different metrical levels, which were first triggered by corresponding isochronous patterns. They could further track those levels when unpredictable, temporally distorted patterns followed, which metrically coincided to the trigger only on the lowest (or fastest) level (400 ms). Synchronization at different levels exhibited varying qualities. The perturbations caused typical delays in synchronization ("next-cycle adaptation") at particular metric levels, but significant "synchronization disturbances were rarely seen at higher metrical levels."[111] Tracking the periodicity of a particular level seems to imply simultaneous monitoring of other levels. LARGE et al. conclude "that the behavioral flexibility reflects the simultaneous representation of three periodicities within a single coordinative structure [...], similar to a music-theoretic metrical structure".[112] Thus, the relation of psychological reality to metric structure is an important motivation for research into formal properties of meter, as presented in section 4.2.

The dynamics of interactions between several attentional pulses, which monitor different levels of a metric hierarchy for temporal tracking, implies both bottom-up and top-down processes. The former can be thought of "as a source of low-level attentional control",[113] focusing on local rhythmic properties. The latter corresponds to the adaptive process of anticipation, "based on internal oscillations".[114] In case of complex metric entrainment, top-down processes induce complex expectancies, leading to metric accentuation (see sections 2.3 and 4.3). Moreover, the results of LARGE et al.

[109] Vuust and Witek, 2014, p. 2: "In this way, meter emerges as a consequence of the reciprocal relationship between external periodicities and internal attending processes."
[110] Large, Fink, and Kelso, 2002, p. 3
[111] Large, Fink, and Kelso, 2002, p. 13
[112] ibid.
[113] Barnes and Jones, 2000, pp. 261 f.
[114] ibid.

showed that subjects were "able to recruit and synchronize oscillations corresponding to different metrical frequencies at will."[115] It is thus evident that attentional processes can be modulated by conscious control, affecting the quality of the perceived meter. The attentional focus can voluntarily shift away from phenomenally dominant, that is, rhythmically embodied metric strata.[116]

To summarize, the perceptual emergence of a metric hierarchy corresponds to simultaneous pulse sensations, resonating to different metric levels. Hence, entrainment is dynamically shaped, both by the properties of a stimulating rhythm, and by the subjective aspects of cognition, such as spontaneous adaptive mechanisms[117] and voluntary attentional control.

As mentioned above, entrainment turns out to be highly flexible and adaptive to complex musical settings. LONDON discusses metric entrainment by illustrating possible entrainment behaviors in concrete musical situations. His inquiry, "how much meter a listener will generate in response to a particular rhythmic surface"[118] leads to a categorical distinction about musical rhythms. A rhythmic figure is *metrically overdetermined* when there is more rhythmic information than needed to entrain on higher metric levels, such as the beat or measure level. On the contrary, if the rhythmic surface does not articulate a salient metric level, "listeners will have to generate a periodicity [...] that is not phenomenally present in the music."[119] The rhythmic structure of such music is thus *metrically underdetermined*. Consequently, the depth of metric entrainment stands in a complex relation to rhythmic features, as filtering or extrapolation of metric levels from the phenomenal input are intrinsic to dynamic attending and entrainment processes. Metrical overdetermination does not mean that the associated rhythmic differentiation is metrically irrelevant:

> the metric foreground – that is to say, the subdivision levels of the beat – has [...] significant import on our metrical attending and hence the meaning and motional qualities of a musical gesture.[120]

Entrainment theories employ oscillators to model the periodic development of attentional energy. On their part, oscillators introduce a polarity between "in-phase" and "out-of-phase". Articulated beat-subdivision levels perceptually fill up the out-of-phase moments between peaks of attentional energy.

While this section focuses on theories and models of dynamic attending and entrainment, several other approaches provide similar cognitive perspectives on pulse and

[115] Large, Fink, and Kelso, 2002, p. 16
[116] Large, Fink, and Kelso, 2002, p. 15: subjects "are able to produce taps corresponding to beats that may or may not be marked by a stimulus event on any given cycle." See also section 2.2.3 and Parncutt, 1994.
[117] More precisely, the discussed aspects of attentional rhythm: expectancy, anticipation and temporal capture (adaptation to violations of expectations).
[118] London, 2012, p. 72
[119] London, 2012, p. 75
[120] London, 2012, p. 98

meter sensation.[121] Computational models of beat induction, surveyed in section 5.1, partly implement those theories, or are based on a more pragmatic set of rules. Their requirements vary between the simulation of the cognitive behaviors, discussed here, and practical applicability for issues of *music information retrieval*, like beat tracking on a symbolic or audio level. For the purpose of the quantitative model of metric malleability, proposed in section 5.2, both aspects – cognitive modeling and computational application – are relevant.

2.2.2 Tempo and temporal thresholds

> Rhythm does not remain invariant across changes in tempo; rather the rhythm emerges at a specific tempo.[122]

Musical tempo is not simply communicated by the number of rhythmic events within a fixed time span. It emerges as a subjective aspect of rhythmic entrainment. Tempo is thus not inherent in a rhythmic texture, but is carried by a subjective pulse, occurring at a specific rate. In fact, if rhythm induces pulse sensation, it retroacts to shape the rhythmic gestalt which implies a subjective sense of tempo. As a factor of entrainment, tempo sensitivity may improve anticipation because it "helps listeners [to] track musical events as they unfold in time and enables predictions about when future events are likely to occur."[123] On the other hand, if tempo is constituted by an emergent cognitive process, it is basically ambiguous and may be affected by voluntary control of rhythmic engagement.[124] In this context, cognitive tempo has to be distinguished (1) from the function of "tempo" as a means of musical coordination and synchronization, and (2) from temporal density of musical events, as it will be discussed in the following.

(1) Subjective tempo can emerge independently from communicated tempo, that is for instance, a musical interpretation of a notated tempo in a score. Usually, both may be parallel throughout the major corpus of written music, but the potential to create a "malleable relationship between the two tempo types"[125] is of special interest to this study. The same cognitive tempo can be evoked by different notational variants,[126] or a music, notated in a certain tempo, can evoke different emergent tempos. BENADON uses the term *conceptual tempo* to indicate "an abstract durational grid over which different emergent tempos may be manifested".[127]

[121] As for instance, the theory of projection, proposed by Hasty, 1997, and the rule-based approach of Lerdahl and Jackendoff, 1983.
[122] Handel and Oshinsky, 1981, p. 9
[123] McAuley, 2010, p. 172
[124] cf. for instance Clayton, 1997, pp. 7 f.
[125] Benadon, 2004, p. 564
[126] An apropos example is provided by Gotham, 2015a, pp. 24 f., which is cited in the context of section 4.1.2 (see figure 4.1).
[127] Benadon, 2004, p. 564

(2) As suggested, tempo essentially stands in a complex relation to the temporal density of a rhythmic texture, that is, the amount of rhythmic component events (section 2.1.1) per time unit. Sometimes both aspects are misleadingly regarded as commensurate to each other. FRAISSE, for instance, notes that tempo "corresponds to the number of perceived elements per unit time, [...] a definition based on frequency".[128] If cognitive tempo is linked to a pulse sensation, it is inherently metric. Consequently, the relation between the tempo function of pulse rate and the density of surface rhythms can vary to a great extent. Two examples shall illustrate this. First, in baroque harpsichord music, pieces implying slow paces are commonly interpreted by means of opulent embellishment (see figure 2.8). Indeed, as harpsichord tones have a very fast decay, arpeggios, trills, grace notes, and so forth, are appropriate to suggest sustained sound. Anyhow, this leads to an amount of rhythmic event density which may stand in opposition to slow tempo.

FIGURE 2.8: Beginning of *Les Tendres Sentiments: Rondeau* from ROYER, *Pièces de clavecin* (Premier Livre, 1746)

The example in figure 2.8 belongs to the *Pièces de clavecin* (Premier Livre, 1746) by the French composer Pancrace ROYER. In an interpretation by Christophe ROUSSET (Decca, 1993), part *A* (first ten bars) lasts ca. 35 seconds, suggesting a mean pace of ca. 51 quarters per minute. As chords are often arpeggiated and embellishments are articulated *largo*, the approximate number of clearly distinguishable attacks in that part is about 120, a mean of 12 per bar. Though, most of the attacks do not contribute to a rhythmic structure which suitably communicates tempo. In fact, they make it more difficult to follow the pace because the embellishments take effect as "rhythmic noise"[129] rather than supporting frequency. In this example, the slow pace of the quarters is nevertheless recognizable, contrasting the lively rhythmic surface.

Second, the difference between musical tempo and rhythmic event density is reflected in Hindustani music theory by the term *lay* or *laya*, which covers both aspects and even

[128] Fraisse, 1982, p. 151

[129] London, 2012, p. 196: "maintaining a simple isochronous meter, where one needs to [...] discount various kinds of rhythmic noise such as trills or grace notes, may prove [...] challenging."

their relationship.[130] CLAYTON demonstrates the flexibility of *lay* by a discussion of the extraordinary tempo range of contemporary Hindustani music, specified from around 10 to 700 *matras* (beats) per minute. It will become obvious that this range is much wider than both the possible range of cognitive tempi and the range of effective metric durations. Hence, the rate of *matras* merely expresses a conceptual tempo in the sense of BENADON.

> In practice, what happens at very slow tempi is that the matra is too long to be regarded as the 'beat' - musicians count units of 1/2 or 1/4 matra, and tabla players articulate these divisions when playing the *theka*.[131] In effect, the 1/2 or 1/4 matra pulse takes over the original role of the matra, and the matra takes over that of the *vibhag*[132] [...].[133]

Figure 2.9 illustrates the same *tal* in two different tempi, the faster *madhya lay* and the very slow *vilambit lay*. The rhythmic articulation of a metric layer below the *matra* in the latter shifts the relation of conceptual tempo and rhythmic density. In fact, both are reduced, but by a different amount. Another consequence of the slow *vilambit lay* is that the highest metric layer (*avart*) becomes ineffective as its duration (ca. 40 s) is not recognizable anymore as a metric period, as discussed later in this section.

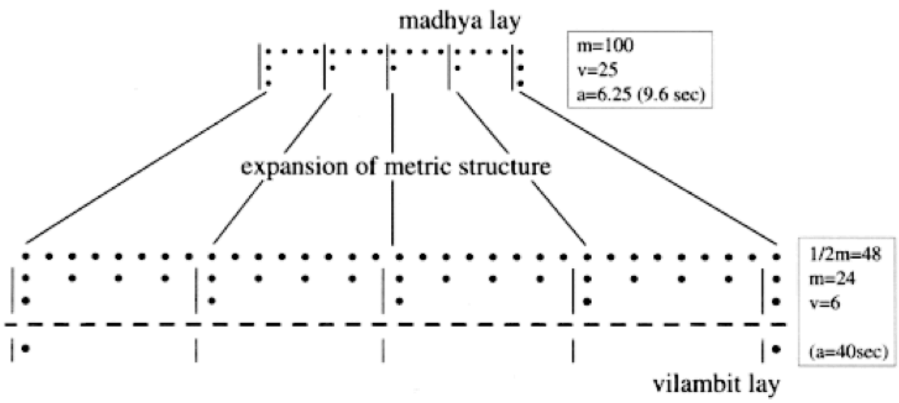


FIGURE 2.9: Clayton, 1997, p. 13: "The expansion of the 16-matra tintal to a tempo of 24 matras/minute. m=matra, v=vibhag, a=avart; pulse rates are given in MM. The lowest level, the avart, becomes divorced from the metric structure at slow tempo (vilambit lay)."

These two examples show that tempo, like pulse, is distinct from rhythmic articulation. They give as well birth to a sense of equivocality. Many aspects of musical interpretation may have impact on tempo sensation. A pace can be suppressed, distorted or

[130] cf. Clayton, 1997, p. 7 of the translation

[131] The theka is a basic sequence of drum strokes, associated with a particular *tal*. *Tal* can denote a certain rhythmic-metric pattern (cf. section 3.2.2) or the rhythmic system of Hindustani art music as a whole.

[132] Vibhags are metric sections, which are not necessarily of equal length.

[133] Clayton, 1997, p. 12. The footnotes are added here to clarify the Indian terms according to references elsewhere in the source.

shifted in relation to the overall rhythmic movement. In the context of what follows, it is consequently assumed that tempo exclusively emerges by means of pulse sensation and metric entrainment. Otherwise we merely perceive temporal event density. Tempo and event density should thus be regarded as independent from each other.

As mentioned, tempo is specified by the rate of a perceived pulse. Musical intuition suggests that tempo conversely has a fundamental impact on the subjective character of that pulse. A wide range of studies provide corresponding empirical evidence, combining observations about perceptual judgements and motor behaviors related to tempo.[134] The temporal limits and preferences of pulse sensation and motor entrainment also constitute basic conditions for meter as a cognitive framework. A metric fabric integrates the emergence of pulse sensation and tempo under certain constraints, which will be further identified.

Pulse sensation occurs within a particular temporal range, which is limited to periodicities affording rhythmic entrainment. Studies on motor synchronization have proven that "there is an optimum tempo and a upper and lower threshold for repeating bodily movements elicited by music."[135] To get a clear idea of those limits of rhythmic sensitivity, and the perceptual qualities of different periodicities, it is essential to distinguish complex metric entrainment from synchronization behavior in the context of simple isochronous stimuli.

> The lower limit for meter – that is, the shortest interval that we can hear or perform as an element of rhythmic figure – is about 100 ms. Conversely, the upper limit is around 5 to 6 seconds, a limit set by our capacity to hierarchically integrate successive events into a stable pattern [...]. These upper and lower bounds can be regarded as a kind of *temporal envelope* for meter.[136]

LONDON also refers to the lower limit for metrically relevant IOIs as "the 100 ms metric floor."[137] It corresponds to the minimal IOI between consecutive, perceptually distinguishable rhythmic events. Nevertheless, shorter intervals are produced in rhythmic performance, but then onsets get blurred as in embellishments like drum rolls or trills.[138] Thus, IOIs below 100 ms cannot be perceived as metric time units. Their sizes cannot be compared as their fast progression limits the mode of perceptual processing.

Estimates of the "upper limits" for entrainment and meter perception diverge, as they refer to different situations. Synchronization to isochronous sequences with IOIs beyond two seconds is reactive, and hence imprecise.[139] Single events are perceived as

[134] See e.g. Parncutt, 1994, Noorden and Moelants, 1999, and Handel and Lawson, 1983

[135] Noorden and Moelants, 1999, p. 43

[136] London, 2012, p. 27 (emphasis in source). See pp. 25 ff. for surveys of studies on tempo sensitivity.

[137] London, 2012, p. 134

[138] See section 2.1.1. Noorden and Moelants, 1999, p. 53, observe that, when events "follow each other too quickly, it will be difficult to determine their temporal order or even to distinguish them as individual elements (loss of temporal acuity)."

[139] Unless, an interpolation strategy is applied.

isolated as they do not afford entrainment or anticipation.[140] Entrainment to larger periodicities can thus only be induced by internally structured temporal patterns. The "hierarchical arrangement of our temporal attention allows us both to expand the upper bound of the temporal window from 1.5 to 2 seconds for individual inter-event intervals to a longer, composite span of time, and to integrate qualitatively defined, very short intervals into a rhythmic attention-action framework."[141] Another important factor is memory integration.

> In order to link two successive tones we need to be able to keep them simultaneously in our perceptual working memory, sometimes called the perceptual (or subjective) present. However the exact range of this perceptual present is difficult to establish.[142]

Van NOORDEN and MOELANTS report different estimates about the range of the perceptual present between 3 and 5 seconds.[143] LONDON, referring to MICHON and others, notes that "the psychological present [...] as the time interval in which sensory information and concurrent behavior are to be integrated within the same span of attention [...] may vary; typically 2 to 3 seconds, its upper limit may reach 5 to 7 seconds".[144] He assumes that the variance of approximations in the literature "may be related to perceptual context – different rhythmic patterns may afford longer or shorter metric frameworks."[145] This is in line with HASTY's theory, according to which meter emerges from *projection* of *mensurally determined* durations. These are durations perceived as metric units or time spans. "Beyond two seconds, mensural determinacy rapidly deteriorates in [...] simple environments."[146] Though, a specific musical context can alter metric entrainment, as "limitations or conditions for the grasp of duration arise from our adaptation to a particular environment".[147] To summarize: the question "how slow a pulse can become to be felt as rhythmic motion" cannot be answered independent from a particular context. It is indeed possible to entrain to periodicities up

[140] McAuley, 2010, p. 173, London, 2012, p. 29, Fraisse, 1982, p. 156

[141] London, 2012, p. 190

[142] Noorden and Moelants, 1999, p. 54

[143] ibid: "Fraisse (1957) found a maximum of about 5 seconds, but also stated that our perceptual present only seldom exceeds 2 or 3 sec. More recently, converging evidence from different fields points to a window of temporal integration working up to 3 seconds (Pöppel, 1996)." (referring to Fraisse, P. (1957). *Psychologie du temps*. Paris: Presses Universitaires de France; and Pöppel, E. (1996). "Reconstruction of subjective time on the basis of hierarchically organized processing system". In: *Time, internal clocks and movement*. Ed. by M.A. Pastor and J. Aritieda. Amsterdam: Elsevier Science, pp. 165–185)

[144] London, 2012, p. 30, referencing Pöppel, Ernst (1972). "Oscillations as possible Basis for Time Perception". In: *The study of time*. Ed. by J.T. Fraser, F.C. Haber, and G.H. Müller. Berlin: Springer Verlag, pp. 219–241; Michon, John A. (1978). "The Making of the Present: A Tutorial Review". In: *Attention and Performance* VII. Ed. by J. Raquin. Hillsdale, N.J.: Erlbaum; and Fraisse, Paul (1984). "Perception and Estimation of Time". In: *Annual Review of Psychology* 35, pp. 1–36.

[145] ibid: "if 2 seconds is the limit for hearing successive events as temporally connected outside of a metric hierarchy, then it makes sense that the absolute value for a measure might be from about 4 to 6 seconds (that is, twice or three times the length of the slowest possible beat)."

[146] Hasty, 1997, p. 86

[147] Hasty, 1997, p. 147

to 6 seconds if an appropriate context is provided – that is: a rhythmic structure which also evokes regular motion or pulse sensation at faster rates.

The temporal window for pulse sensation and metric entrainment includes regions of periodicities with distinct characteristics affecting rhythmic engagement. The thresholds, or rather, the transition zones between those regions are not fixed. They show some flexibility, depending on the rhythmic context, and vary considerably between different subjects.[148] Nevertheless, converging evidence is registered for fundamental differences between these regions.

Periodicities faster than about 200 – 300 ms are processed differently than slower ones. An IOI of "approximately 200 ms is needed for two tones being perceived as fully independent, without influencing each others loudness."[149] Temporal discrimination also works differently when periods exceed 250 ms. In this range, the *just noticeable difference* (JND) between two successive durations, grows with period duration as a constant proportion (according to WEBER's law) of about 2.5%. A constant absolute JND of ca. 6 ms is found between durations below this period range.[150] Motor synchronization behaviors also show differences related to this temporal threshold. Tapping "one to one" along isochronous sequences is hard to be precisely performed at high speeds with IOIs below 200 ms.[151] The range "between 200 and 300 ms can be regarded as a transition zone in which controlled tapping is possible, but still relying on more or less automated processes, without the possibility of anticipation and immediate adjustment."[152] Presumably, the speed limit of controlled motor behavior is only partly due to kinetic constraints, and reflects as well the perceptual limits of temporal discrimination and integration.

A peak of sensitivity and preference for tempi around 100 bpm (periods around 600 msec) is reported in a wide range of studies.[153] Periodicities in this range typically attract more attention and more obviously provoke entrainment than considerably slower or faster ones. Hence, a *moderate* tempo seems to demand less cognitive effort to be tracked than much faster or slower tempi. These effects correspond to the average of individual tempo preferences. Experimental studies of "preferred tempo have generally emphasized either spontaneous motor measures or perceptual measures."[154] Both exhibit inter-individual variability but distribute in the vicinity of the

[148] London, 2012, p. 27
[149] Noorden and Moelants, 1999, p. 54: "Experiments in the determination of the audibility threshold [...] show that the ear integrates energy over time within a time frame of roughly 200 ms. Loudness increases if a tone gets longer until the threshold of 200 ms is reached (Gelfand, 1981). This implies that we can only speak about a fully developed tone if it lasts at least 200 ms." (reference: Gelfand, S.A. (1981). *Hearing – an introduction to psychological and physiological acoustics*. New York: Marcel Dekker, Inc.) See also Povel and Okkerman, 1981 and section 2.3.2 about "echoic" memory.
[150] London, 2012, pp. 33 f., Noorden and Moelants, 1999, p. 54
[151] McAuley, 2010, p. 172
[152] Noorden and Moelants, 1999, p. 54
[153] See e.g. Fraisse, 1982, Noorden and Moelants, 1999, Parncutt, 1994
[154] McAuley, 2010, p. 173

mentioned frequency.[155] McAuley et al. "found a large, positive, correlation between SMT [spontaneous motor tempo] and PPT [preferred perceptual tempo], i.e., near 0.75. Such correlations support the view that motor and perceptual tempo preferences have a common psychological basis."[156] Other indicators for an increased sensitivity at preferred rates, are tempo discrimination and synchronization tasks, which are performed best at those rates.[157]

Nevertheless, particular cues in a musical signal, such as perceptually salient component rhythms (section 2.1.1), can distract from periodicities in the moderate-tempo region, and call attention to other temporal ranges.[158] Due to the recognition of rhythmic pattern structure or of motivic analogies, our temporal focus is often biased to higher metric levels which group together several periods of lower levels (see sections 2.3 and 2.4).[159]

2.2.3 Pulse salience

The salience of a pulse sensation may be defined as the relative strength or intensity of that sensation in the context of rhythmic entrainment.[160] As remarked in section 2.2.1, the sensation of pulse correlates with regular peaks of attentional energy in an entrainment pattern. Rhythmic cues in music potentially evoke pulse sensations at different periods and phases. One listener may "latch" on a different pulse than the other.[161] Furthermore, simultaneous monitoring of metric levels in a rhythm may lead to a concomitant perception of embedded pulses, associated with those levels. It may be possible to estimate the relative strengths of the latter, because each of the simultaneous pulse trains may contribute, to a different extent, to an attentional pattern. Attention is guided by the most salient, that is, the most attention-demanding pulse sensation. To

[155] Fraisse, 1982, pp. 153 f. In tapping experiments, the IOI of two consecutive strokes in spontaneous performances of a periodic sequence varied between 330 and 880 ms. Fraisse asserts 600 ms as the most representative IOI duration. The spontaneous tempo highly varies between individuals but is measured as personally rather stable. Judgements about the most natural tempo, neither fast nor slow, of an isochronous sequence show as well a mean interval of 600 ms.

[156] McAuley, 2010, p. 174 (terms in brackets added). Measures of preferred tempo also reveal influences of age and musical experience: both increasing age and musical expertise (especially in childhood) shift the preference to slower tempi (pp. 174 f.).

[157] Parncutt, 1994, pp. 449 f., Fraisse, 1982, p. 155: "synchronization is most regular for intervals of 400 to 800 msec.", McAuley, 2010, p. 192: "Studies of tempo discrimination and tempo production reveal violations in Weber's law consistent with the concept of a preferred tempo. With respect to perception, JNDs for tempo tend to be a minimum in a range of tempi (optimal tempo region) that is centered on 600 ms, with thresholds for single interval sequences around 6% and those for multiple-interval isochronous sequences around 2% in ideal listening conditions. Weber fractions for variability in tempo productions tend to mirror threshold results observed for perception."

[158] Moelants and McKinney, 2004, p. 558: "Quite often a majority of the listeners judges a metric level far from the resonance frequency most salient."

[159] In the words of London, 2008: "For coordination of human action/interaction, entrainment most crucially involves coordination of the higher-level component periodicities (in the 1-3 second range) of a hierarchically-organized set of rhythms."

[160] Parncutt, 1994, Parncutt, 1987, Gotham, 2015a

[161] If the periods/rates of these pulses differ, listeners may have also different tempo percepts.

measure the intensity of a pulse sensation, PARNCUTT gives a probabilistic definition of pulse salience, and connects perception to motor reaction.

> The salience of a pulse sensation is a measure of the probability that a listener will tap out that pulse when asked to tap the tactus (main or underlying beat) of a rhythmic sequence.[162]

Several computational models of pulse salience, employed in models of metric interpretation, have been suggested (see section 5.1). Some emphasize aspects which are important for the computational approach to metric ambiguity and malleability, presented in section 5.2. PARNCUTT's model estimates pulse salience as mainly affected by two factors, which he calls *pulse-match salience* and *pulse-period salience*.[163] The former is derived from pattern recognition behaviors, and the latter reflects the influence of pulse rate, basically in accordance with what is stated in the previous section 2.2.2. Thus, pulse-period salience is meant to be a quantitative estimate of the subjective differences of pulse character in regard to pulse rate. Van NOORDEN and MOELANTS proposed a similar model which concentrates on the issue of human body resonance to tempo.[164] In the context of tempo attractors for meters, GOTHAM picked up the concepts of pulse-period salience and tempo resonance to propose individual pulse salience as a factor of *meter salience*, which depends on tempo and metric type.[165]

As entrainment implicates anticipatory behavior, pulse salience is closely related to expectancy and memory. Pulse perception induces an expectation: essentially that the pulse will continue. Pulse salience and expectancy may grow proportionally in the course of a continuing pulse which confirms its self-induced expectation. It may also be affected retrospectively, if the assumption by DESAIN, made in the context of his general expectancy model (see section 2.2.1), holds for pulse salience. In terms of "the influence that a new incoming event might have on the prior context", he proposes that the "new event can support one or the other of the previous interpretations and in retrospect contribute to a limited extend to the salience of already perceived stimuli."[166] Expectancies are thus projected into the future, and as well into the past. The assumption that "a later event can facilitate or inhibit the memory of past ones"[167] may as well be applied to pulse salience. A rhythmic development which enhances or inhibits a pulse sensation may influence the memory of its salience. This may be an important aspect of the dynamics of metric conflict (section 2.3.3).

We recognize patterns regardless of their complete or partial perception. Moreover, we can discern implicit patterns caused by regularities in complex superpositions. This

[162] Parncutt, 1994, p. 433. The terms *tactus* and *beat* are used as synonyms. It is assumed to be equal to the pulse perceived as most salient. It also indicates the perceived tempo.
[163] Parncutt, 1994. These aspects will be further developed in section 5.2.1.
[164] Noorden and Moelants, 1999
[165] Gotham, 2015a
[166] Desain, 1992, pp. 444 f. See also footnote 82 in section 2.2.1.
[167] ibid.

well researched phenomenon, proposed in the context of visual gestalt theory as the *principle of good continuation*, also applies to acoustic perception.[168] Two functionalities of this cognitive strategy can be distinguished in regard to pulse sensation. First, we are able to integrate fragmentary cues in a permanent succession of perceived pulses, and second, a pulse or a metric entrainment pattern can emerge from a complex rhythmic succession of acoustic cues. These modes of entrainment may be associated to rhythmic situations which are metrically underdetermined or overdetermined (section 2.2.1).

> Like all perceptual patterns, a pulse may be perceived even when some elements of the pattern – the equally spaced events – are missing.[169]

To formalize the impact of pattern recognition on pulse salience, PARNCUTT assumes that the perceptual salience of a pulse grows with the number of "events" matching that pulse. Thus, he uses the terms *event* and *event salience* to describe the influence of component properties (section 2.1.1) on pulse perception: "*pulse* and *metre* percepts in music are determined by (and therefore predictable from) configurations of event percepts".[170]

However, certain constraints might be taken into account. As perceived event saliences are dynamically processed in immediate memory, they progressively fade away:[171] "each pair of event percepts in short-term memory contributes to the salience of one and one only pulse percept, whose period and phase are determined by the temporal positions of the two events."[172] Hence, the "psychological present" or "short-term memory" (see section 2.2.2) acts as a limit and saturating space for pulse salience.[173] This is especially interesting regarding cyclic rhythms, as the length of a cycle may or may not exceed this limit. Consequently, the cycle itself can or cannot influence pulse perception: if the cycle period is too long, it "cannot evoke a feeling of pulse, as no more than one of the events of such a pulse can be 'stored' in a single 'chunk' of short-term

[168] Deutsch, 1982, pp. 126 f., describes several cases where an "auditory continuity effect" occurs, for instance, when softer, or even physically missing tones are perceived to form continuing streams, although they are obscured or masked by louder ones. See also Rosenthal, 1992, p. 75, and Toussaint, 2013, pp. 195 ff.

[169] Parncutt, 1987. p. 132

[170] Parncutt, 1987, p. 127 (emphases in source)

[171] See section 2.2.1 (cf. Desain, 1992, Hasty, 1997, London, 2012).

[172] Parncutt, 1987. p. 132 (see also Noorden and Moelants, 1999, p. 52). Similar approaches are developed by Rosenthal, 1992 and Volk, 2004 (see also Fleischer, 2002), proposing a comparable algorithm to explore metric structure in musical scores independent from notated time signatures. "Pulse layers arising from equally spaced notes' onsets (called *local meters*) serve as a starting point for the definition of two different types of *weights* which model the metric structure expressed by the notes." (Volk, 2004, p. 435, emphases in source) See also the notion of *metric coherence* (section 4.4.1, footnote 227).

[173] cf. Parncutt, 1987: "a limited time period, metaphorically called the 'psychological present', 'short-term memory' or 'echoic memory'. The duration of short-term memory is normally of the order of a few to several seconds." (p. 133, referring to Neisser, U. (1967). *Cognitive Psychology*. New York: Meredith; Glucksberg, S. and Cowen, G.N.Jr. (1970). "Memory for nonattended auditory material". In: *Cognitive Psychology* 1, pp. 149–156; Crowder, R.G. (1970). "The role of one's own voice in immediate memory". In: *Cognitive Psychology* 1, pp. 157–178.)

memory”.[174]

As mentioned in section 2.2.1, analytical approaches to possible cognitive processes underlying pulse sensation and metric interpretation have to take into account both bottom-up and top-down strategies. The former can be conceptualized as an extraction or extrapolation of periodicity from a perceived rhythmic structure. The latter corresponds to a template-matching or pattern-matching activity, that is, the rhythmic structure is compared with an internalized repertoire of familiar templates.[175]

> In unfamiliar metrical territory we cannot resort to template matching strategies, and in these cases we must proceed by a more general process of first extracting the relevant periods from the musical surface and then finding the metrical framework that optimizes our attention to them.[176]

Hence, the application of one or the other type of process depends not least on musical experience, that is, the depth of the listener’s internalized repertoire of rhythmic and metric patterns. Possibly, both strategies can be activated by the same instance, when similar patterns are recalled, but, immediately, their discrepancy with the currently perceived pattern suggests to reprocess the whole structure. Bottom-up processing may thus interfere, if counter-evidence occurs during template matching. Moreover, top-down processing is only conceivable for pattern durations which do not exceed the psychological present. Although it operates within the same time window, pattern matching may not lead to the same conclusions as maintaining an established metric framework. Termed *metric priming*, this is further examined in section 2.3.3. An established metric framework “primes” the cognitive processing and its corresponding pulse sensations: yet another top-down inference on the rhythmic percept.

To generate empirical data for the already mentioned model of pulse salience, PARNCUTT designed a palette of cyclic stimuli – short rhythmic figures with a temporal structure of a maximum of three different IOIs in ratio 1:2:3. These are represented in figure 2.10. Note that the cyclic pattern (e) equals our example pattern [1-2-3] – introduced in section 1.1 and further investigated in section 5.2 – if both are regarded as *metric rotations* of each other (see section 4.1.3). The sequences were presented with randomly chosen percussion samples. Within one sequence each sound event was identical. Each of the six sequences was presented six times with six different event densities,[177] logarithmically distributed between 50 and 400 events per minute (0.83 to 6.7 per sec.). PARNCUTT states that such a distribution of different numbers of events/notes per time unit provides a better balance within the spectrum of musical tempi, than a similar distribution of different beats per time unit. However, it was the listeners task to perform

[174] Parncutt, 1987, pp. 133 f. (quotation marks in source), see also London, 2012, p. 30: “The constraint on the scope of larger temporal patterns is related to our sense of the psychological present.”
[175] cf. London, 2012, pp. 67 ff., see also Desain and Honing, 2003, and Huron, 2006
[176] London, 2012, p. 68
[177] See section 2.2.2 for a discussion of *event density* versus *tempo*.

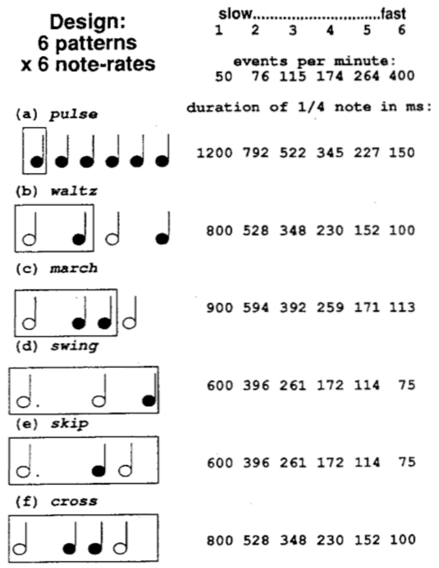


Fig. 1. Reference table for stimuli used in the experiments. Each pattern is given a name for ease of identification. The boxes define rhythmic cycles or measures. The name *cross* refers to a cross-rhythm of 1/2-notes and 3/4-notes.

FIGURE 2.10: Experimental stimuli from Parncutt, 1994, p. 415

a certain beat rate by tapping along the sequences, and thus, they collectively produced the actual distribution of speeds.[178] Among participants, different pulses were beaten along most versions of the sequences. This suggests that there is no definite perception of a most salient pulse: the tactus or beat may thus be ambiguous in most cases of such sequences, if we assume that the tapped pulse corresponds to a subject's most salient pulse perception. The ambiguity seemingly increases with rhythmically more complex sequences, and/or a higher density of events. The empirical results of the described experiment are reproduced in appendix A. They are discussed at greater length in the course of the present study, as they provide evidence for some important features regarding the metric interpretation of temporal patterns and its inherent ambiguity (see, amongst others, section 3.2).

PARNCUTT reveals a close link between the salience of pulse sensation and pulse rate, by a statistic overview of his data in regard to the pulse durations of all 792 tapped beats. Figure 2.11 shows a logarithmic distribution of the intervals of all tapped pulses. The crosses represent the numbers of responses with pulse IOIs within units of 0.1 on the logarithmic scale. This set of all tapped periods ranges around a mean duration of 0.71 sec. The mean deviation is 0.224 on the logarithmic scale which corresponds

[178]For details of the experimental design, see Parncutt, 1994, pp. 414 ff.: 22 listers tapped along the "main or underlying beat" of the 6 cyclic sequences (see figure 2.10) in 6 different densities of events. The entry-points of the cycles were randomly chosen for each of the 22 * 6 * 6 = 792 trials to level or grade their influence. As mentioned, the first two events of a sequence can already trigger a pulse sensation. This illustrates the possible influence of the entry-points in cyclic sequences on their perceived *metric rotation* or *metric phase* (see sections 1.1 and 4.1.3).

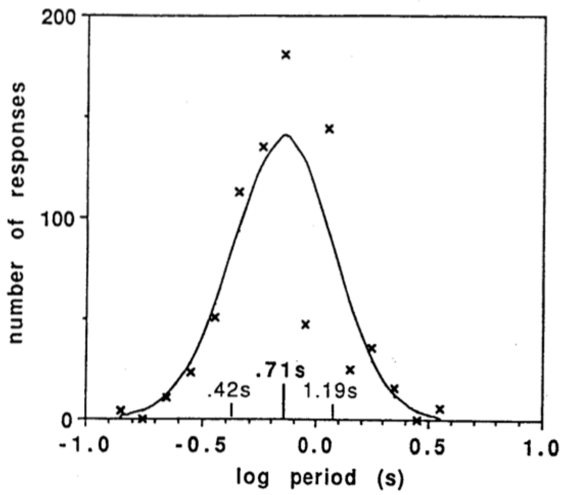


Fig. 3. Further results of Experiment 1 (pulse salience). Crosses: distribution of the periods of all 22 × 36 = 792 pulse responses. Vertical axis: number of pulse responses falling in categories of width 0.1 (in $\log_{10}$ of pulse period in seconds). Curve: The corresponding normal distribution.

FIGURE 2.11: Existence region of pulse sensations (Parncutt, 1994, p.420)

to the time interval 0.42 to 1.19 sec. PARNCUTT interprets the mean duration – equal to a tempo of 81.5 bpm – as a point of gravitation: the listeners tendency to beat in a moderate tempo implies a higher salience of pulse sensation in this temporal range. He consequently defines an "existence region of pulse sensation", and assumes a general relation between the pulse rate and the salience of a pulse sensation:

> DEFINITION: *The existence region of pulse sensation is a range of periods within which isochronous sequences are perceived to be musically rhythmic (or imply movement).*
> ASSUMPTION: *The median, or (logarithmic) center, of the existence region of pulse sensation corresponds to a moderate musical tempo. In general, the closer the tempo of a pulse to moderate tempo, the greater the salience of the corresponding pulse sensation.*[179]

The ranges of the spontaneous and preferred tempo (see section 2.2.2) correspond as well with the logarithmic mean of the existence region of pulse sensation. This *moderate pulse period* may be perceived as natural and agreeable, because it may induce a most salient and unambiguous pulse sensation. This may be a requirement for a spontaneous tapping or beating. Additionally, timing can be most precise in this range (see section 2.2.2). According to the described empirical distribution, PARNCUTT defines the relation between pulse rate and pulse salience, called *pulse-period salience* in his model, as a Gaussian function resembling the curve in Figure 2.11 (see equation 2.1).[180]

[179]Parncutt, 1994, p.436 (italics in source)
[180]Parncutt, 1994, pp. 437 f.

$$S_p = exp(-\frac{1}{2}(\frac{1}{\sigma}log_{10}(\frac{p}{\mu}))^2) \quad (2.1)$$

In equation 2.1, p denotes the pulse period, whereas the variables μ and σ account for the preferred *moderate pulse period* and the standard deviation of the logarithm of p, "a measure of the width of the existence region".[181] Modeling his experimental data (which are partly reproduced in appendix A), PARNCUTT estimated the "typical" values for μ around $700ms$ and for σ around 0.2.[182] He describes the function of equation 2.1 as a bandpass-filter modeling the saliences of pulse sensations within their "existence region". The saliences rise to the logarithmic mean (moderate tempo) of this temporal region from both sides (slowest and fastest pulse sensations), and obtain only negligible values outside this range.

As mentioned above, the concept of pulse-period salience has been further explored in subsequent studies.[183] In the context of the present study, it is adopted as an essential building block for the model of metric malleably, proposed in section 5.2.

2.3 Grouping and accentuation

Grouping is an extensive topic in psychology and cognitive sciences. Basically, grouping processes are automatic cognitive mechanisms which may be consciously controlled to a very limited extend. Hearing as well involves spontaneous grouping of elements into larger perceptual entities.[184] The way in which perceptual grouping interacts with properties of musical sequences, was frequently set in relation to principles stated in gestalt theory.[185] Groups as gestalts are in a way abstracted from the rhythmic flow (see chapter 4. They may be varied and still identified. DEUTSCH calls this the "production of invariance under transposition".

> Given that we are presented with a set of first-order acoustic elements, how are these combined so as to form separate groupings? If all first-order elements were indiscriminately linked together, auditory shape-recognition operations could not be performed. There must, therefore, exist a set of mechanisms that permits the formation of [...] linkages between certain elements, and inhibits the formation of such linkages between others.[186]

[181] Parncutt, 1994, p. 438
[182] Parncutt, 1994, pp. 440 ff.
[183] cf. Noorden and Moelants, 1999, and Flanagan, 2008. Gotham, 2015a proposes a series of adaptations of Parncutt's formula (see equations (1) to (4) in Gotham, 2015a), specifying $\mu = 600ms$ and $\sigma = 0.3$ to suit his model of "attractor tempos for metrical structures" (see section 5.2.3).
[184] The decision, wether these elements are component events or IOI may be left open for now (cf. sections 2.1.1, 2.1.2, and 2.2.1).
[185] See e.g. Fraisse, 1982, p. 161, Jones, 2010, pp. 8 f., McLachlan, 2000, p. 61. Tenney and Polansky, 1980 developed an algorithmic theory on grouping, influenced by gestalt theory.
[186] Deutsch, 1982, p. 100

Systematic triggering of those cognitive mechanisms by acoustic cues of confined complexity, reveals robust correlations between grouping and acoustic patterning. For instance, there are strong indications that elements with increased intensity initiate a group and pitch series are grouped according to repetitive units in otherwise plain sequences.[187] However, a more complex interplay of durational proportions, intensities and pitch structure, inevitably leads to equivocality. That is: perceptual grouping may differ from person to person and arise in unpredictable ways. In other words, perceptual groups are malleable. This fact leads to assume that grouping is also an important factor in the context of metric malleability. It is therefore essential for this study to examine potential musical properties and perceptual preferences which resolve malleability in grouping processes.

The temporal window for grouping is commonly estimated in the same range within which metric structure can be processed. A group, as a perceptual unity, is noticeable within the duration of the psychological present (ca. 4-6 seconds, see section 2.2.2).[188] It is not only for this reason that the relation of grouping and meter is widely discussed in the context of rhythm perception. LERDAHL and JACKENDOFF assume that meter and grouping are independent mechanisms in the cognitive processing of rhythm, and suggest two different sets of rules for each.[189] As part of *A Generative Theory of Tonal Music* (GTTM in the following), this classical proposal has exerted a great influence on subsequent computational models of rhythm perception. Both sets of rules comprise subsets of *well-formedness rules* and of *preference rules*. The former enumerate formal constraints under which grouping and metric structures can be cognitively construed. The latter specify criteria for judgements about the likelihood of particular construals, when more than one would be formally appropriate. The concept behind these rules yields two implications. First, in accordance with results described above, grouping processes resolve rhythmic equivocalities according to perceptual preferences. Second, these equivocalities are differentiated with respect to grouping and meter.

Metric entrainment may result from grouping processes or vice versa. Particular dynamics between grouping and meter thus arise from ambiguous sources. The analytical separation of both perceptual strategies by LERDAHL and JACKENDOFF is also related to the widely discussed opposition between serial and periodic grouping of rhythm.[190] Figural or serial grouping "depends primarily on the serial proximity in time, pitch and timbre of temporally adjacent events [...]. Gaps between serial groups tend to be the largest gaps available in a given sequence".[191] Periodic grouping corresponds to

[187] McAuley, 2010, pp. 184 f.
[188] Fraisse, 1982, pp. 157 f.
[189] Lerdahl and Jackendoff, 1983
[190] Parncutt, 1994, pp. 411 ff. Corresponding terms in use are "figural" and "metric coding" (Povel and Essens, 1985, see also McAuley, 2010, p. 183), as well as "metrical" versus "non-metrical representation" (Essens and Povel, 1985)
[191] Parncutt, 1994, p. 412. See also Fraisse, 1982, pp. 162 ff. and Yu, Getz, and Kubovy, 2015 (cf. section 3.2.5).

metric entrainment and pulse sensation, and "usually depends on the relative timing and perceptual properties of *nonadjacent* events".[192] This distinction reveals another dimension of malleability, as "the relative importance of periodic and serial grouping depends on the listener."[193]

Imagine the rhythm [x.x.xx.x.x.x] – representing one of a couple of well-known patterns, extensively discussed later in this study[194] – in a cyclic, repetitive form, rendered monotonically at a pulse duration about 250 ms. It may be serially grouped as two groups of two events, separated by a single event and by two single events respectively. It could also be periodically grouped in three or four pulses. If the starting point of the cycle is assumed on the first pulse displayed, the former would yield the four groups [x.x] [.xx] [.x.] [x.x] and the latter would result in the three groups [x.x.] [xx.x] [.x.x]. All three perceptual structures highly contrast with each other. They are nevertheless equally plausible. A crucial difference between serial and periodic grouping is the fact that periodic groups can definitely begin "empty" – that is, not marked by an event – which is hardly the case in serial grouping. As periodic grouping implies metric entrainment, these empty beginnings correspond to *virtual pulses* or *syncopations* (see sections 2.3.3 and 4.3.1). The tendency for a specific periodic grouping may be weakened by an empty beginning. When the elementary-pulse duration is about 250 ms, the duration of whole example cycle above is about 3 seconds. Hence, it can easily be grasped as an entity within the perceptual present, though the cycle period does not have much impact on pulse sensation. It is thus unpredictable if a pulse will emerge, and if yes, which periodic groups will occur. Grouping preferences consequently change at different rates. At faster rates, the example cycle may preferably be grouped in two half-cycles. In this way, empty beginnings can even be avoided when the groups start at the sixth and the twelfth pulse: [x.x.x.][xx.x.x].

At this point, we merely note the fact that serial (or figural) grouping is a more general cognitive process, whereas pulse perception and metric entrainment specifically occur as part of a periodic grouping activity.

Grouping implies a perceptual differentiation of the grouped elements, and some elements may be perceived as *accented*. This can happen as a subjective consequence or implication of grouping, independent from component properties. Nevertheless, contextual musical properties like temporal lengthening, peaks or leaps of loudness or of melodic contour, bring about accent perception just as well. Therefore, the complexity of the phenomenon has led to diverging views in music theory and psychology, and has caused considerable confusion. Several types of accents were differentiated and

[192]ibid. (emphasis in source)
[193]ibid.
[194]This pattern, exhibiting interval structure [2-2-1-2-2-2-1], and known for instance as one of the "standard" patterns in in sub-Saharan African rhythm (section 3.2.2, footnote 122), will be frequently revisited throughout this study. Its structural properties specifically support metric ambiguity (see sections 3.2.2 and 4.1.1, where also its structural equivalence with the diatonic scale is discussed). In section 5.3.1, it will be identified as a prevalent instance of a malleable rhythmic *necklace* (sections 1.1 and 4.1.3).

associated with the mentioned musical parameters – and others – but it proved difficult to consistently describe the role of those accent types in the interplay of grouping and meter.[195] From the perspective that grouping and meter are emergent phenomena, it is less surprising that accentuation similarly resists causal determination. LERDAHL and JACKENDOFF subsume "any event at the musical surface"[196] which causes a sense of emphasis as *phenomenal accent*. They suggest two other categories: *structural accents* indicate gravity points in a wider musical context, such as cadences or closures, and *metrical accents* denote relatively strong beats in a metrical context. The relations between these accent categories also reflect the subjectivity and ambiguity of an emergent metric pattern.

> Phenomenal accent functions as a *perceptual input* to metrical accent – that is, the moments of musical stress in the raw signal serve as "cues" from which the listener attempts to extrapolate a regular pattern of metrical accents. If there is little regularity to these cues, or if they conflict, the sense of metrical accent becomes attenuated or ambiguous. If [...] the cues are regular and mutually supporting, the sense of metrical accent becomes definite and multileveled.[197]

The notions of phenomenal accent and the extrapolation of subjective metrical accents match the ideas of attention, expectancy, anticipation and metric entrainment. In fact, different terminologies describe the same issue under consideration. LERDAHL and JACKENDOFF assume an extrapolation of perceptual input, and thus distinguish "active" cognitive processes from "passive" sensory reception. This seems particularly appropriate regarding voluntary engagement with musical stimuli. The subjective resolution of metric malleability can indeed be attended by conscious control, and is at least guided by extrapolation, respectively anticipation processes. Hence, we will continue to regard the cognitive mechanisms involved in rhythmic structuring as temporal patterns of relative attention. In this framework, accent is basically a temporal peak of attention. In the words of JONES, an "accent is anything that is relatively attention-getting in a time pattern."[198] JONES suggests an integrated approach to accents as relative attentional peaks, called *joint accent structure* (JAS).[199] JAS is connected to DAT (see section 2.2.1) as it describes a cognitive activity. Basically, melodic and temporal accents are supposed to be identified in pitch and time relationships, specifying a higher order dynamic structure from which the temporal peaks of attention emerge. This interplay level enables "accent coupling"[200] which introduces a differentiation of attentional strength, that is, combined melodic and temporal accents are perceptually stronger than those relying on just one aspect. The theory predicts "that coincident

[195] See e.g. Berry, 1976 and Benjamin, 1984
[196] Lerdahl and Jackendoff, 1981, p. 485
[197] ibid. (emphasis and quotation marks in source)
[198] Jones, 1987a, p. 623
[199] Jones, 1987a, p. 625
[200] ibid.

(concordant) melodic and temporal accents (corresponding to a simple joint accent structure) should lead more efficient entrainment and a stronger perception of metrical structure than conflicting (discordant) accent timing (corresponding to a complex joint accent structure)."[201] This has proven empirically evident, as melodies with a concordant JAS are easier to perform, and their metrical structure is more clearly perceived, and less ambiguous.[202]

LONDON adapts the ideas of JAS and DAT to his notion of metrical accentuation. The divergence among musicological perspectives could be resolved, as

> regarding meter as a form of anticipatory behavior finesses the problem of metrical accent. Rather than seeking a phenomenal basis for metrical accent – whether by brute force in the form of dynamic emphasis, or due to the tonal interpretation of events – metrical accents are generated by the listener via his or her attending process. The degree of metrical accent is correlated with the relative strength and temporal focus of our entrained temporal expectancies.[203]

Being aware of the ambiguities inherent in grouping and accentuation processes, their exploration in real musical engagement and listening situations is difficult. Empirical research on the interaction of multiple dynamic musical features, and its influence on grouping and accentuation, has thus rarely been done so far.[204] "This endeavor remains challenging in part because it involves striking the right balance between experimental control and real musical complexity."[205] To improve the musical validity of experimental methods, "two factor designs that allow for systematic manipulations of each of two types of musical components"[206] are a reasonable starting point for further development. Nevertheless, the fact that most "research in this area has focused on temporal accents and has used either very simple or even isochronous sequences"[207] provide plenty of useful empirical evidence, according to the abstraction level of the theoretic studies, presented in chapters 4 and 5. Temporal variations of sequences containing identical events already induce complex accentuation phenomena. The following section 2.3.1 deals with the perception of isochronous sequences of which only IOI-duration, respectively event rate, is varied. Insights into the influence of temporal differences between IOIs in such sequences are discussed thereafter in section 2.3.2.

[201] McAuley, 2010, p. 187
[202] ibid., referencing, amongst others, Ellis R.J., and Mari Riess Jones (2009). "The role of accent salience and joint accent structure in meter perception". In: *Journal of Experimental Psychology: Human Perception and Performance* 35, pp. 264–280.
[203] London, 2012, p. 24
[204] Honing, Bouwer, and Háden, 2014, p. 318
[205] Jones, 2010, p. 8
[206] ibid.
[207] Honing, Bouwer, and Háden, 2014, p. 318

2.3.1 Subjective grouping of isochronous sequences

> [...] given an even/undifferentiated series of pulses we tend to hear them in groups (usually twos) and with a subjective sense of accent. This sense is wholly endogenous, and not guided by any phenomenal cues.[208]

Already in 1894, BOLTON[209] examined this phenomenon – the *subjective grouping* of monotonous sequences with isochronous IOIs – which he called *subjective rhythmization*.[210] Within a certain range of IOI duration,[211] listeners spontaneously group elements of such sequences into larger rhythmic units. The first element of a group is often accented, that is, its intensity and length is perceptually augmented.[212]

Subjective grouping "is an aural illusion that exists in the mind of the perceiver, but cannot be measured with scientific equipment in the outside world, since it lacks a concomitant acoustic signal."[213] "It presents a clear cautionary message to researchers attempting to simulate human responses to music",[214] as cognitive processes are not part of the physical world. Hence, towards a heuristics of metric malleability, subjective grouping is another variable to be taken into account when isochronous passages occur in rhythms.

At his time, BOLTON precisely marked the limits of IOI durations permitting subjective grouping ($0.115s$ - $1.58s$). These values can be compared and are similar to the temporal boundaries of rhythmic motion perception (section 2.2.2). POVEL refers to a range from $0.125s$ to $1.5s$, outside which the shorter intervals between events lead to the impression of internally ungrouped or unstructured rhythmic cascades. Longer intervals result in perceptually isolated events.[215] The limits of the existence region of pulse sensation (section 2.2.3), defined by PARNCUTT, indicate nearly the same range. Hence, pulse perception and subjective grouping occurs in the same temporal region.[216]

Subjective groups of isochronous elements tend to contain more elements with shorter IOIs or vice versa, "suggesting that there may be an intrinsic preferred total duration for each group".[217] Van NOORDEN and MOELANTS assume "a periodic fluctuation in the perceptual system"[218] with a characteristic frequency corresponding to the length of subjective groups. "One could call this the resonance period for hearing subjective

[208] London, 2008, p. 7
[209] Bolton, 1894
[210] Fraisse, 1987, p. 8
[211] respectively, event rate (to avoid the term *tempo*, which was characterized as subjective as well in section 2.2.2)
[212] Fraisse, 1982, p. 156
[213] Toussaint, 2013, p. 196
[214] Sethares, 2007, p. 6
[215] Povel, 1984, p. 330
[216] Parncutt, 1994, pp. 418 ff.
[217] McAuley, 2010, p. 185
[218] Noorden and Moelants, 1999, p. 46

rhythmization."[219] This is supported, amongst others, by the data provided by PARNCUTT, introduced in section 2.2.3 and reproduced in figure A.1: the shorter the IOIs of an isochronous sequence, the less elements were tapped along, with a preference to tap on every second or every fourth event. Thus, the intervals of $0.345s$ and $0.227s$ are mainly grouped by two, and an interval of $0.15s$ is mainly grouped by four (see figure A.1 (a)). With these interval durations also, a grouping by three, or another group size like five or eight, is possible. These groupings or pulses are purely subjective, as the sequence does not contain any "phenomenal" cue for a certain grouping. Hence, pulse sensation and subjective grouping are closely linked to each other. An effect of preferred tempo (sections 2.2.2 and 2.2.3) can also be derived from these observations, as subjects tend to avoid tapping very slow or very fast.

Independent from PARNCUTT's investigation, where listeners were explicitly asked to tap pulses, spontaneous grouping of isochronous sequences usually continuously combines a fixed number of elements. Therefore, successive subjective groups have the same duration: a case of periodic grouping, leading to a second-order isochronous sequence which can trigger a pulse sensation. This may be an explanation for the mentioned accent which can be perceived at the beginning of a group, and a motivation to tap according to that regularly recurring accent. Thus, a superordinate pulse sensation, which corresponds to a *metrical* accentuation, can emerge from subjective grouping. Other cases of metrical accent will be discussed in section 2.3.4. In any case, according to their correspondence, metrical accentuation depends on the same temporal constraints as pulse sensation.

In order to explore, in a different way, the influence of the IOI duration of an isochronous sequence on the number of subjectively grouped elements, FRAISSE asked his subjects to tap isochronous sequences, and to group the strokes by three and by four.[220] On average, the produced IOIs were clearly shorter than the range of spontaneous and preferred tempo – the logarithmic mean of the existence region of pulse sensation and moderate tempo ($0.6s$ - $0.7s$). Grouping by three resulted in a mean IOI of $0.42s$, grouping by four in an average IOI of $0.37s$. Hence, element rate and number of elements in a group are proportional: there is a tendency to keep the duration of a group in the range of strong pulse sensation. PARNCUTT assumes that performers precisely adjust the tempo to optimize or maximize the total strength of all pulse sensations induced by the sequence – their "*aggregate* salience".[221] GOTHAM attests the convergence of different experimental data sets in accordance to that idea, and similarly assumes that this "appears to indicate a desire to balance the various pulse levels involved, maximising the combination rather than any individual pulse."[222] Van NOORDEN and MOELANTS, employing the same data to calculate parameter values for their model, suggest

[219] ibid.
[220] Fraisse, 1982, p. 157
[221] Parncutt, 1994, p. 438 (emphasis in source)
[222] Gotham, 2015a, p. 32

a resonance frequency for subjective grouping, which corresponds to an optimal group period around $1.1s$.[223]

Hence, the regular subjective grouping of an isochronous sequence within a certain temporal range obviously corresponds to two simultaneous pulse sensations. A basic pulse matches the elements, and a second-order pulse parallels the grouping, or rather the accentuation that comes along with it. This is a special case of a more general tendency that interacting grouping mechanisms lead to emergent tempo perception in the context of hierarchical meter. BENADON, referring to BERRY, distinguishes "two interacting components: pulse- and activity- tempo".[224]

Although primarily spontaneous and involuntary, subjective grouping is amenable for conscious control. HONING et al. describe a voluntary grouping switch from automatically accenting every other tone, to a conscious projection of subjective accent on every third tone, "thus adjusting the period of the beat to our will. This ability has been very useful in examining beat and meter perception, because it can allow us to hear a physically identical stimulus as on the beat or not, depending on the instructions".[225] From this point of view, voluntary switching of subjective grouping is also a special case of the more extensive musical capability to consciously reframe a (more complex) rhythm according to a different meter. HONING et al. however argue in a neurophysiological research context, and claim that neural activity can be more directly correlated with subjective accent and beat perception when physically neutral stimuli are employed. Otherwise, the perceptual processing of physical differences in a stimulus and the cognitive activity of subjective grouping may mutually interfere in corresponding neural patterns.[226]

BROCHARD et al., indeed, were the first to demonstrate neurophysiological evidence for subjective grouping.[227] They showed that neural responses – measured by event-related potentials (ERPs) – are significantly different, when intensity differences are introduced on single odd-numbered versus even-numbered events, in an isochronous sequence of identical tones with IOIs of $0.6s$.[228] Based on the assumption of binary subjective grouping and the perceptual salience of the first tone, BROCHARD et al. "expected that odd-numbered events would correspond to accented positions ('strong' beats) and even-numbered events to unaccented ones ('weak' beats)."[229] ERPs are supposed to have larger amplitudes when expectancies are violated at moments of higher attention. According to the hypothesis that subjective accents correspond to attentional

[223]Noorden and Moelants, 1999, pp. 46 f. The experimental data are taken from Vos, P. (1973). *Waarneming van metrische toonreeksen*. Nijmegen, The Netherlands: Stichting Studentenpers.
[224]Benadon, 2004, p. 563. Cf. Berry, 1976, p. 305
[225]Honing, Bouwer, and Háden, 2014, p. 309
[226]ibid.
[227]Brochard et al., 2003
[228]Deviant tones were presented 4 dB softer than the other tones
[229]Brochard et al., 2003, p. 363

peaks, the study also supports DAT (section 2.2.1) "in that attention seems to be deployed periodically."[230] Related theories of meter perception similarly suggest an interaction of attention and expectancy with periodic neural activity. The *neural resonance theory* assumes "neural oscillations that resonate to external events".[231] In fact, neural oscillations are detected to emerge with subjective accent, though their functional role remains to be investigated.[232] The *predictive coding theory* regards subjective grouping as a special manifestation of meter as "a key predictive model for the musical brain, shaped by statistical learning, and repeatedly challenged by the sensory input from rhythmic patterns"[233]. From this perspective, subjective accentuation is dependent on a predictive system which projects corresponding expectations on a neutral stimulus.

As indicated, the experiment of BROCHARD et al. affirms the predominance of binary patterns in subjective grouping. This exemplifies an obvious bias to expect "binary" metric structure, which is observed and discussed from different perspectives. First, metric expectations are assumed to be formed by long-term statistical learning. Duple and quadruple meters are numerically predominant in Western classical tradition, and thus "one might suppose that listeners experienced with Western music would be biased toward binary expectations."[234] However, the influence of experience and exposure to musical environments on metric expectancy and subjective grouping remains to be explored.[235] Second, the mentioned behavioral studies in the context of subjective grouping generally indicate a bias for binary groups.[236] However, these observations may be qualified because they may depend on the temporal constraints of subjective grouping. BROCHARD et al. used isochronous sequences with $0.6s$ IOIs, which strongly suggest duple grouping, as predicted for instance by the resonance model of van NOORDEN and MOELANTS mentioned above. A spontaneous triple grouping would have resulted in a group duration of $1.8s$, a rather unlikely case merely from the perspective of pulse sensation. Nevertheless, PARNCUTT states that subjective "groups of four occur more often than groups of three", and that "this effect is quite general and independent of tempo."[237] He offers an explanation based on considerations related to pulse salience:

[230] Brochard et al., 2003, p. 365. See also Huron, 2006, pp. 195 f.

[231] Honing, Bouwer, and Háden, 2014, p. 310. Neural resonance theory "can be seen as an extension of DAT and makes largely the same predictions. Like the attending rhythms in DAT, neural oscillations are suggested to be self-sustaining and are suggested to adapt their phase and period to an external rhythm." (ibid.)

[232] ibid.

[233] Vuust and Witek, 2014, p. 5

[234] Huron, 2006, p. 195

[235] Cf. Huron, 2006, p. 196: "It is conceivable that there is some innate disposition toward binary temporal grouping, but I suspect that the acoustical environment plays the preeminent role. Cross-cultural experiments could help resolve whether binary grouping is innate or learned through exposure."

[236] In Parncutt, 1994, p. 416, almost all tapping responses to an isochronous sequence correspond to groupings of two or a power of two (see figure A.1 (a)). Noorden and Moelants, 1999, p. 46, report that the data of Vos mentioned above (see footnote 223), show similar tendencies.

[237] Parncutt, 1994, p. 418. Cf. figure A.1 (a)

> HYPOTHESIS: *The salience of a pulse sensation may be enhanced by the presence of a parent or child pulse sensation, that is, a consonant pulse sensation two or three times slower or faster. In other words, there is mutual salience enhancement among consonant pulse sensations.*[238]

From this perspective, a quaternary grouping may involve three consonant pulses sensations, if they all occur within the temporal window of pulse perception. The two simultaneous pulses suggested above – a faster matching the events, and a slower corresponding to the subjective accentuation of the first element in the group – are consonant with a supposed intermediate pulse which coincides with every other event.[239] If mutual enhancement of consonant pulse sensations takes place, it would generally promote subjective groupings by a power of two. Combined with the assumption of the existence region of pulse sensations, it becomes obvious that there is a higher potential for binary consonant pulses in isochronous sequences, than of ternary consonant pulses. This aspect is further illustrated in section 5.2.4.

Subjective grouping exemplifies metric malleability in a nutshell. Two simultaneous pulses – the faster one induced by the isochronous sequence, and the slower one resulting from the subjective groups – establish the perception of a simple hierarchic meter.[240] When the groups are spontaneously or voluntarily switched, the sequence is reframed by a new “meter”. Hence, an isochronous sequence is highly malleable, and offers a rich source for metric interpretation in spite of its simplicity.

2.3.2 Temporal accents in non-isochronous sequences

> As soon as a difference is introduced into an isochronous sequence of elements, this difference produces a grouping of the elements included between two repetitions of the difference. [...] This difference can be a lengthening of a sound, an increase in its intensity, a change in pitch or in timbre, or simply a lengthening of an interval between two elements.[241]

The differentiation of element properties in musical sequences affects the formation of grouping which may interfere with endogenous processes of subjective accentuation. Bearing in mind all the components which are capable to intervene these activities, we will subsequently concentrate on grouping as a consequence of temporal differentiation in non-isochronous sequences of identical events.[242] Systematic studies on the

[238] Parncutt, 1994, p. 443 (italics in source)

[239] Parncutt, 1994 asserts that subjective grouping “appears to involve both serial and periodic grouping, where the two types of grouping are in phase – the start of each serial group corresponds to a periodic group (downbeat). In the case of grouping by fours, subjective accents occur twice per group, on the first and third elements.” (pp. 443 f.)

[240] cf. Yeston, 1976, quoted in section 1.3 (footnote 85).

[241] Fraisse, 1982, p. 157

[242] That is, monotonous sound patterns comprising IOIs with different durations.

behavioral effects of varied IOIs in such sequences reveal complex responses in terms of perception and production.[243] In spite of the artificial character of those stimuli (see section 2.3) it is thus worth to study temporal variation independently from all other acoustic and musical aspects.

POVEL and OKKERMAN explored the perception of *temporal accents* in non-isochronous two-tone cycles.[244] Such accents occur due to IOI differences in a temporal sequence. As opposed to dynamic or melodic accents, and so forth, they are not caused by differences between event properties. The empirical data represented in figure 2.12, suggest two types of temporal accent which come along with perceptual grouping.

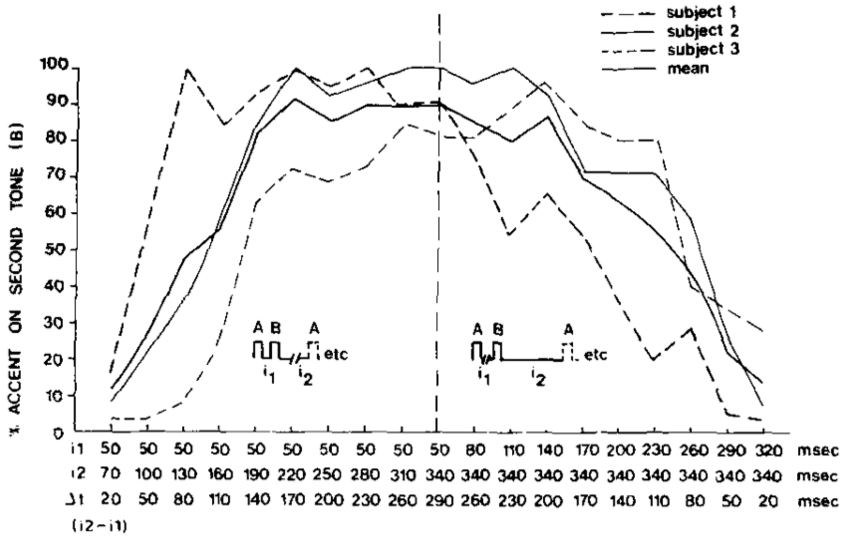


FIGURE 2.12: "Percentage of perceived accent on last tone of group in sequences consisting of repeated groups of two tones, as a function of the duration of the intervals i1 and i2." (Povel and Okkerman, 1981, p. 568) The two intervals between very short ($50ms$), identical sine tones with fast attacks ($7.5ms$), were variated one after the other (left and right from the dashed vertical line).

A longer interval divides cyclic two-tone groups as soon as it exceeds the shorter by 5-10%: the first tone of a group is then perceived as accented. Similar to accentuation through subjective grouping of isochronous sequences (section 2.3.1), this effect involves serial grouping corresponding to the gestalt principle of proximity.[245] If the interval ratio is further enlarged, the second tone of a group is heard as accented. This accent seems to be stronger than the formerly described first-tone accent. Another experiment proved that, if the second tone is accented, the loudness level of the first tone has to be raised by 4 dB to balance the sequence, that is: to perceive the two tones as equally accented. Though, a tone in an isochronous sequence is definitely perceived

[243] See e.g. Repp, London, and Keller, 2011
[244] Povel and Okkerman, 1981
[245] See e.g. Deutsch, 1982, p. 101, Fraisse, 1982, pp. 159 ff., Toussaint, 2013, pp. 198 f.

as accented with a 2 dB increase in intensity relative to other tones, and the mentioned first-tone accent can be evened out with less than 2 dB increase of the second tone in a group.[246] These differences justify the distinction of two different accentuation types. The second-tone accent is clearly induced by a considerable IOI difference, whereas the weaker first-tone accent corresponds to a perceived grouping. Accordingly, POVEL and OKKERMAN call the former an interval-produced accent, and the latter a grouping accent, which "is apparently the base response for temporally rather undifferentiated sequences."[247] The minimal interval difference to entail an interval-produced accent can not be generalized, as the data in figure 2.12 show considerable differences between individuals. Furthermore, if the shorter interval connecting the two-tone groups exceeds $0.25s$, the effect inverses, and the first tone of the group tends to be accented again. A grouping accent thus seems to be induced by all sequences with small IOI differences, independent from their absolute IOI durations within the tested range (see the sequences listed in the leftmost and rightmost columns of figure 2.12).

The question arises if this accent is resulting from the perceptual grouping, or from an orienting response influenced by the starting point of the sequence.[248] If orientation leads to an accentuation of the very first tone,[249] the repetition of the group could initiate a pulse sensation leading to a projection of accent on the initial tones of following groups, respectively cycles. Orienting responses may generally interact with structural cues for the determination of the perceived gravity point or *downbeat* of a rhythmic cycle. In an anacrustic beginning, for instance, the first accent or rhythmic emphasis is deferred to a later attack. Structurally, this is often induced by a longer IOI after the accented moment, exactly parallel to the described cases of interval-produced accent (see section 4.3.2). To explore the perceptual implications of cyclic structures independently from a possible orienting response related to the starting point, experimental setups have to be carefully designed to avoid its influence on subjects behaviors.[250] POVEL and OKKERMAN added a longer preceding tone ($0.5s$ duration) to the mentioned sequences comprising little IOI differences, and double-checked responses. Grouping accents occurred slightly less frequent, but the same tendencies remained. Thus, this type of accent is evidently conditioned by properties of temporal structure.

The interval-produced accent could be related to the temporal window of energy integration mentioned in section 2.2.2 (footnote 149), as it only occurs when the shorter

[246] Povel and Okkerman, 1981, p. 570. If only every third interval in the sequence is lengthened and three tones are grouped together, the accentuation tendencies are the same. A small IOI-difference causes an accent on the first tone of a group. "With a further increase of the third interval, in addition to the accent on the first tone a stronger accent is reported on the last tone of the group." (p. 565)

[247] Povel and Okkerman, 1981, p. 567

[248] Povel and Okkerman, 1981 call this phenomenon *induction*. See also sections 3.2.5 and 4.1.3 for related discussions, concerning metrical phase responses and metric rotation.

[249] See the experiment of Brochard et al., 2003 in section 2.3.1.

[250] For instance, Parncutt, 1994 presented his cyclic rhythms with randomly chosen entry points (cf. section 2.2.3, footnote 178), and Yu, Getz, and Kubovy, 2015 presented their cycles with decelerating speed (cf. section 3.2.5, footnote 166).

interval is less than about 250 ms. Hence, the temporal placement of the second tone could have an influence on the loudness estimate for the first tone, because it may not be fully processed in a time span shorter than $0.25s$. Though, POVEL and OKKERMAN found only little influence of backward masking, by testing similar sequences with longer tone sustains. They thus involve the widely-used concept of *echoic memory*,[251] and assume that the relative strength of a tone may be underestimated when the perception process is interrupted by a new attack. They furthermore suppose that loudness determination only starts after a monitoring period of about $0.15s$, which is needed for energy integration.[252] Interval-produced accents may thus still be noticed after preceding IOIs which are considerably longer than the monitoring period.

Echoic storage differs from memory as an aspect of cognition. Rather, "it may be regarded as a lingering of sensory experience in the absence of conscious cognitive processing."[253] Experimentally supported estimates about the duration of echoic storage vary widely, but most are located between $0.5s$ and $2s$.[254] Given this amount of variability, PARNCUTT assumes that the salience of sensory excitation in echoic memory exponentially decays in time,[255] which better corresponds to the idea of lingering than a fixed storage duration. This would also explain the inter-individual differences in accent responses, by assuming different "decays of the internal persistence of the stimulus in echoic store."[256]

PARNCUTT models the effects of echoic storage and IOI duration on accent perception, as parameters contributing to the perceptual salience of events. He concludes that, leaving aside other phenomenal properties, the more salient is an event, the greater the IOI after it. With growing salience, an event would also gain more impact on coinciding pulse sensations, respectively metrical accentuation. The data shown in appendix A provide corresponding empirical evidence, of whether events coinciding with taps are interpreted as relatively more salient (see section 2.2.3). Events followed by longer IOIs are generally more salient (in the mentioned sense) in temporal structures with mixed IOIs which are considerably longer and shorter than 250 ms.[257] When all IOIs exceed this threshold, there is no clear relation of salience and IOI.

Hence, event salience seems to be influenced by temporal structures in the vicinity of the echoic storage. This is in accordance with the observations of POVEL and OKKERMAN, as their interval-produced accent contributes to the salience of the event that

[251] Povel and Okkerman, 1981, p. 566. See also Parncutt, 1994, pp. 427 f. for an overview about the development of this concept in experimental psychology.

[252] Povel and Okkerman, 1981, p. 571

[253] Parncutt, 1994, p. 428

[254] ibid.

[255] ibid., referring to Kubovy, M. and F. P. Howard (1976). "Persistence of pitch-segregating echoic memory". In: *Journal of Experimental Psychology: Human Perception and Performance* 2, pp. 531–537.

[256] Povel and Okkerman, 1981, p. 568. These differences may additionally caused by inter-individual variety, concerning the duration of the monitoring period mentioned above.

[257] At least this is valid within the experimental condition and context of cyclic temporal sequences with identical events.

precedes the interval.[258] If various IOIs in a temporal structure have durations in the range of the echoic memory, stronger differences in the accentual effects of differently sized IOIs can be assumed. Temporally determined event salience may thus be formalized as *durational accent* which "*increases with IOI for small values of IOI and saturates as IOI approaches and exceeds the duration of the echoic store (auditory sensory memory)*."[259] Durational accent can only be effective when IOIs are considerably longer than the minimum discriminable IOI.[260] At this threshold (50-100 ms or 10-20 events per second) the effectively perceived IOI may be zero, as successive events can just be discriminated.[261] Consequently, differences of IOI duration are most effective for durational accent in the range of 150-300 ms.

Among the manifold types of phenomenal accent, discussed in sections 2.1.2 and 2.3, "durational accent appears to have the greatest impact on metrical organization".[262] Due to this assumption, PARNCUTT has formalized durational accent as an essential feature of pulse-match salience (section 2.2.3). This is further explicated in section 5.2.1 (see equation 5.4), where pulse-match salience is revised, before it is employed in the model of metric malleability as a component of *meter salience* (section 5.2.3). To that effect, durational accent becomes a fundamental aspect in the context of the present study.

2.3.3 Metric priming

> [The] presence of a specific musical metre primes the perception of specific rhythmic patterns[263]

A simple, yet significant demonstration of the influence of a metric context on grouping and accentuation, sometimes called *metric priming*,[264] is provided by POVEL.[265] Figure 2.13 shows cyclic tone sequences. Vertical lines represent equal tones with a duration of $0.05s$ and the shortest IOI is $0.2s$. The enumeration clarifies the cycles and dots show the durations of the longer intervals as integer multiples of the shortest unit.

---

[258] Interestingly, this phenomenon can be related to several aspects mentioned before: first, the notions of "short times", "long times", and "event-related binding" mentioned in section 2.1.1, second, the distinction of "duration" from other sound properties (see the beginning of section 2.1.2), and third, the notion of expectancy as a "bidirectional" process in time (see sections 2.2.1 and 2.2.3).

[259] Parncutt, 1994, p. 427 (italics in source)

[260] cf. sections 2.2.2 and 2.3.1. See also Parncutt, 1994, pp. 427 ff., about the "perceptual onset delay".

[261] Parncutt, 1994, pp. 429 f.

[262] Parncutt, 1994, p. 426. The list of arguments for that claim includes the strength of the interval-produced accent in Povel and Okkerman, 1981 and several studies "suggesting that phenomenal accent is more sensitive to changes in IOI than to changes in loudness" (ibid.)

[263] Desain and Honing, 2003, p. 341

[264] ibid.

[265] Povel, 1984

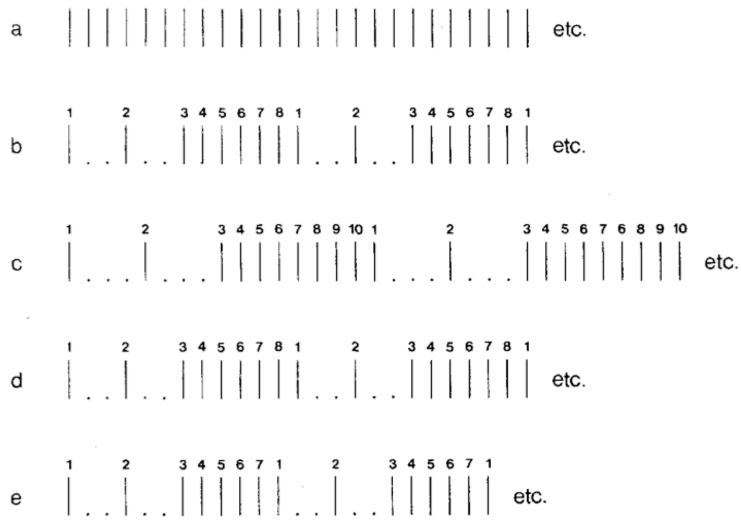


FIGURE 2.13: Tone sequences after Povel, 1984, p. 318.

Cycles b to e are alternations of isochronous sections (shortest IOI) and two longer intervals with triple or quadruple length. POVEL reports clear effects of these structures on grouping and accentuation. Although the isochronous sections are not internally differentiated, listeners of sequence b steadily perceived tones 3-8 as two three-tone groups, whereas tones 3-10 in sequence c were inevitably perceived as two four-tone groups. These groupings are very stable, compared to subjective grouping of isochronous sequences (sequence a), and lead to different accentuations. The longer intervals evidently trigger the groupings of the succeeding shorter ones.

Generally, a perceived temporal structure of any rhythmic component tends to be used as a template for what follows. Rhythm perception is a projective activity. This characteristic is already surveyed from the viewpoint of entrainment in section 2.2.1. Projection is about temporal expectancy.[266] Thus, the priming effects in sequences b and c are induced by an expectation or projection of a group duration. A simple metric context is established by a regular succession which is "filled up" by a faster succession fitting into the framework. The perceptual grouping of the faster, undifferentiated passage is reliably provoked by a preceding context, and not by its internal structure. The regular, metrical accentuation which comes along with such groupings can also be described in terms of pulse, as the two longer intervals structure the following by inducing a pulse sensation. For instance, in sequence b tone 6 coincides with a pulse entrained by tones 1 and 3 and is therefore metrically accented, which supports the perception of three-tone-groups.

[266]The notion of projection in Hasty, 1997 implies that a comparison of two successive (adjacent or non-adjacent) durations or temporal intervals can only be made after the completion of the latter.

The cognitive dimensions of priming can be further differentiated into bottom-up and top-down aspects. The immediate bottom-up effect by the extrapolation of initial interval duration is especially relevant in regard to a potentially different priming process in listening to cyclic rhythms, depending at which point or phase the cycle started (see sections 2.2.3 and 2.3.2). The perception of cyclic repetition can however establish a top-down influence on priming. Evidence for the integration of such parallel processes and their systematic relation to neuronal activity is proposed by predictive coding theory.[267] Another dimension of top-down influence is added by what we have mentioned as long-term templates and probabilistic knowledge, learned through exposure. When "listeners anticipate sounds, they often rely on heuristic knowledge – imperfect rules of thumb about what is likely to happen next."[268]

A template for temporal division may also be applied retrospectively to an initially unstructured temporal interval. Imagine a plain, continuous rendition of the rhythmic cycle [2-1] started with the longer interval of, say $0.8s$ duration. At some point, this first interval would be conceptualized as divided by the pulse induced by the second interval (with duration of $0.4s$).[269] Consecutive temporal structures can thus interact in complex ways: either a subdivision can be primed, or it can be retrospectively realized as a subdivision.

From a broader perspective of musical engagement, priming corresponds to the *conceptual maintenance*[270] of an established meter. Once a meter is induced, listeners tend to maintain it. Not only neutral successions, like tones 3-8 in sequence b or tones 3-10 in sequence c, naturally receive a contextual structure in the mind of the listener. Even in the face of structural challenges arising in succession of an induced meter, its resistance may not take serious effort if conflicting elements can be conceptualized as syncopations. LONDON gives a striking example which is reproduced in figure 2.14.

FIGURE 2.14: London, 2012, p. 16: syncopation in and out of context.

[267] Vuust and Witek, 2014
[268] Huron, 2006, p. 91
[269] cf. Hasty, 1997, p. 153, see also sections 2.2.1 and 2.2.3 about retrospective enhancement of event salience, derived from the notion of temporal "bi-directionality" of expectancy.
[270] cf. Krebs, 1987

Priming is necessarily preceding syncopation.[271] This becomes obvious when we compare pattern a in figure 2.14 with itself, embedded in a metric context (pattern b, from third bar). Pattern a would suggest the quarter attacks (two-lined c and d) to be accented by durational and tonal structure which would result in a duple meter with the first note perceived as anacrusis. In pattern b, however, the first two bars establish a metric context which is incommensurable to that interpretation. The quarter attacks

> are now heard as syncopations against the prevailing meter. That is, even though relative duration, contour accent, and tonal stability (C is the tonic pitch) fight against it, the metric framework can be maintained rather than shifted.[272]

Hence, syncopation depends on a previously established metric background and the subordinate rhythmic construction according to that framework.[273] Other types of rhythmic challenges, like hemiola (see section 3.2.4) and metrical phase conflicts (section 3.2.5) are as well based on priming effects. An example of the latter is also found in figure 2.13. POVEL notes that sequence d (identical to sequence b) was easily replicated, whereas, in reproductions of sequence e, subjects added elements, or the temporal structure was distorted. Sequence e obviously causes a conflict between the primed accentuation by the first two IOIs and the length of the whole cycle (11 units), which is not an integer multiple of the induced pulse length (3 units). A new pulse is introduced with every repetition, having the same period but another phase. Thus, a conflict arises in relation to the primed pulse, which may be resolved by the recognition of – and entrainment to – the odd cycle length.

KREBS has typecast such conflicts which may arise when an established metric framework is "conceptually maintained" in the face of a contradicting rhythmic structure.[274] According to his taxonomy, the examples just mentioned can be categorized into classes of *metrical dissonance*. KREBS assumes that the emergence of hierarchic meter basically involves *metrical consonance*,[275] a notion derived from YESTON's approach to meter as a stratification of levels of motion, shortly introduced in section 1.3. As mentioned in section 2.3.1, an isochronous sequence may induce a superordinate sequence with a slower rate, resulting from subjective grouping. These two (isochronous) sequences are metrically consonant to each other, or, the faster is completely embedded in the slower one. The feature of metrical consonance formally applies to any type of hierarchical meter, as demonstrated in section 4.2. KREBS accordingly distinguishes – like YESTON – an elementary pulse level from one or more "interpretive levels" which provide (hierarchic) groupings of pulses.[276] In contrast, metrical dissonance arises when

[271] cf. Hasty, 1997
[272] London, 2012, p. 15
[273] ibid.
[274] Krebs, 1987
[275] Krebs, 1987, pp. 100 ff.
[276] cf. Krebs, 1987, p. 103: "Metrical consonance arises from the combination of at least two levels such that each attack of every interpretive level in the collection coincides with an attack of every faster level.

metric strata or pulses simultaneously or successively occur, which are not metrically consonant to each other.

> Metrical dissonance, unlike consonance, requires the presence of at least three levels – a pulse level and at least two interpretive levels that provide conflicting groupings of the pulses.[277]

KREBS mainly discriminates two types of metrical dissonance or conflict. First, *type A dissonance* involves interpretive levels that exhibit periods in non-integer ratios, such as 2:3 or 3:4 (see left part of figure 2.15). Second, *type B dissonance* arises "from the non-aligned superimposition of [...] interpretive levels of the same cardinality"[278] (see right part of figure 2.15). In this case, none "of the attacks of these interpretive levels will coincide", [279] whereas in type A dissonance, coincidences must occur, unless the cardinalities of the interpretive levels have a common factor > 1. Then, different phase relations between the interpretive levels are possible, corresponding to type B dissonance.[280] Consequently, the two dissonance types may occur in a combined fashion.[281]

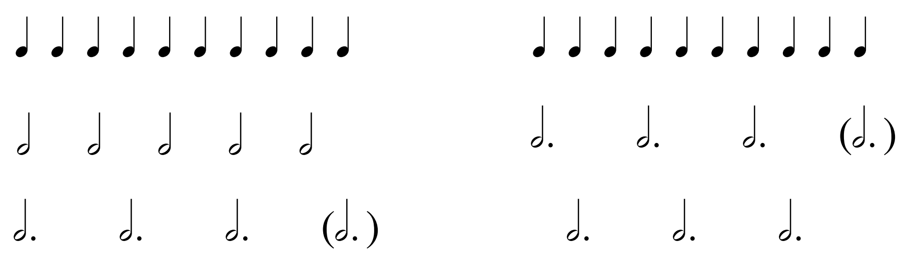


FIGURE 2.15: Types of metric dissonance, after Krebs, 1987, p. 102 (examples 2 and 3).

The concept of metrical dissonance is integrated in the following in several ways. In section 3.2, a taxonomy is suggested to classify those types of metric ambiguity, which can lead to metric conflict and metrical dissonance. Specifically, type A dissonance is associated with metric *period ambiguity* (section 3.2.4) and type B dissonance corresponds to the concept of metric *phase ambiguity* (section 3.2.5). Section 4.4 explores

Given that, by definition, levels involve regular motion, a combination of just one pulse level and one interpretive level must always be consonant".

[277] ibid.

[278] ibid. Here,"cardinality" denotes the number of elementary pulses, grouped by one period of the interpretive level.

[279] ibid.

[280] See also sections 1.1 and 4.1.3 about metrical phase shift.

[281] In regard to type A dissonance, Krebs, 1987 (ibid.) specifies that "for any two dissonant interpretive levels of different cardinalities x and y, there are (x + y) - 2 points of non-coincidence for every one point of coincidence." However, due to the fact that the greatest common divisor of x and y ($gcd(x, y)$) could be > 1, this term should be modified to $(x + y)/gcd(x, y) - 2$, provided that (if $gcd(x, y) > 1$) the two interpretive levels are "in phase".

formal relations between different hierarchic meters, and a definition of *metric contrast* (section 4.4.2) includes measures regarding both period and phase aspects of metric relations, which correspond to KREBS' dissonance types.

Regarding the metric malleability of a rhythm, it can be resolved by metric priming, as described from the perspectives of grouping and metric construal. When a rhythm gets "fixed" in a metric framework, the resulting grouping and accent structure gives additional shape to the rhythmic appearance. In the next section (2.3.4) it will also been shown that priming has an influence on *categorical rhythm perception*. Metric categories are primed, and the same rhythmic figure can be heard to match different categories, depending on the primes.

2.3.4 Rhythmic categories and metrical accent

> Rhythms are perceived categorically. That is, a number of slightly different performed rhythms will be perceived by listeners as "the same" rhythm.[282]

In musical performance, a considerable range of deviation from nominal durations is acceptable.[283] Deviations from nominally metric IOIs in musical performances are interpreted to consist of motor inaccuracies (good musicians deviate considerably less than 100 ms) and conscious, deliberate musical expression, called *rubato* or *expressive timing*.[284] Measurements of the latter reveal up to 50% deviation from notated durations.[285] A number of authors demand to study those systematic deviations in performance timing to study meter.[286] It is widely assumed that meter exerts contextual influence both on categorical rhythm perception and expressive performance. In fact, systematic performance timing is an essential means to *communicate* meter. And indeed, this involves the metric disambiguation and the dissolution of metric malleability of performed rhythm.[287]

A close relation between performance deviations and biases in temporal perception is discussed by many authors.[288] The latter are part of a cognitive process which extracts discrete rhythmic categories from a continuous signal.

[282] Huron, 2006, p. 191 (quotation marks in source), cf. Clarke, 1987
[283] ibid.
[284] "Timing" refers to the ratio/relation between notated "exact" or nominal metric durations and performed duration. Metrically equivalent durations can widely deviate from each other in performance (cf. figure 2.16).
[285] See for instance Desain and Honing, 1992, p.30.
[286] For instance Desain and Honing, 1992, London, 2012, Moelants, 1997, and Polak and London, 2014.
[287] As mentioned in sections 1.3 and 3.1.2, it depends on the context of a musical culture if disambiguation is an aesthetic goal. Though, for instance, the preference of leaving meter in abeyance in many African styles (cf. Locke, 2011) nonetheless gets along well with systematic performance deviations (cf. Polak, 2010, Polak and London, 2014).
[288] See for instance Bengtsson, 1987, Gabrielsson, 1982, Repp, London, and Keller, 2011, Sadakata, Desain, and Honing, 2006.

> The continuous time intervals in music performance are not just categorised into symbolic categories: the categories themselves have structure, as they relate to each other as rational numbers. They can be represented and coded by small integers, each signifying a multiple of a small symbolic duration.[289]

Hence, the separation of categorical and rational time relations from expressive components and variations is not only what happens when rhythm is perceived or performed. It is also exactly what is required to notate natural rhythms and to analyze and construct rhythmic structures (see section 4.1).

Categorical rhythm perception also limits and reduces the rational complexity of temporal differentiation according to our cognitive capacities. In empiric studies of spontaneous rhythmic performance and reproduction, the majority of subjects tend to simplicity. During spontaneous tapping of short rhythmic sequences, only two categorically different IOIs are usually produced in a ratio around 2:1. In reproductions of complex sequences, more than two different IOIs will often be adjusted, resulting again in two categorical IOIs.[290] FRAISSE generalizes this reduction to two temporal categories or magnitudes, and derives two rules of *good gestalt* as economic principles of categorization: *assimilation* (small differences will be further diminished or eliminated) and *distinction* (large differences will be tightened and adjusted to simple ratios).[291] Experimental support for this claim is reported by ESSENS and POVEL. Subjects, asked to reproduce sequences with alternating IOIs in ratio 3:2, augmented the longer IOI and diminished the shorter to approximate a 2:1 ratio.[292]

These observations are also reflected by statistical learning of temporal categories, which is – at least in a western musical context – limited to simple ratios.[293] This may lead to problems with rhythms containing adjacent IOIs with non-integer ratios, like the 3:2 example above. Thus, if a metric unit is divided into two parts with a temporal ratio of 3:2, it depends on additional factors if this subdivision is interpreted as an expressive variation of 2:1 or 1:1, or if two distinct metric categories are perceived, comprising three, respectively two elementary pulses. These factors can be intrinsic to the musical input, like absolute IOI durations (section 2.2.2) or additional rhythmic context. But the listener's experience and cultural preoccupation may also either enable or preclude possible interpretations.[294]

More generally, it will be shown in the following how meter emerges from rhythmic categorization, and that it is fundamental for symbolic approaches to rhythm. DESAIN

[289] Desain and Honing, 2003, pp. 342 f.
[290] Fraisse, 1982, Essens and Povel, 1985
[291] Fraisse, 1982
[292] Essens and Povel, 1985
[293] Huron, 2006, pp. 190 ff.
[294] cf. Honing, Bouwer, and Háden, 2014: "the culture with which we are familiar influences how we perceive [...] metrical structure" (p. 309, referring to Hannon and Trehub, 2005).

and HONING demonstrate the inextricable linkage of rhythmic categorization to meter as an essential aspect of rhythm perception. In this sense,

> it appears that the topic of categorization has a much wider relevance, as it reflects the transition from sub-symbolic to symbolic mental representations and thus forms a bridge from perceptual processes to cognitive ones, with rhythm perception being an intriguing domain where these levels of representation meet.[295]

Figure 2.16 shows a typical relation of performance timing and its categorization in a symbolic score. The rhythmic categories in (b) can be interpreted both as a cognitive level of abstraction of perceived performance "data" in (a) (see arrow "perception"), and as an instruction for performance. The timing of a performance usually deviates from symbolic norms, and the cognitive abstraction "separates structural from expressive components". Categorization and performance timing are thus opposite processes.

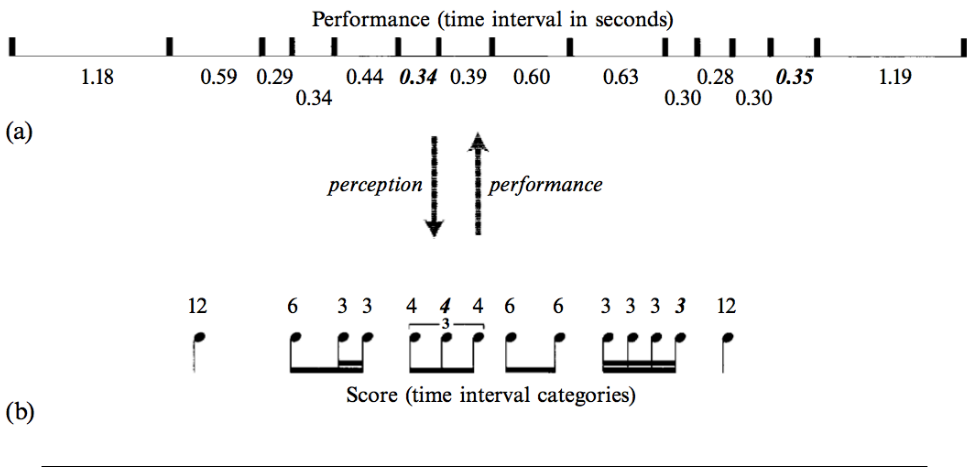


FIGURE 2.16: Performance timing, musical notation, and perceptual categorization (Desain and Honing, 2003, p. 342)

Hence, categorization can be regarded as the inverse process of expressive interpretation – the abstraction and separation of structural meaning from expressive, non-structural information in a temporal sequence. "Categorical perception functions so as to separate essential structural units of events from non-structural information".[296] To use a linguistic metaphor, rhythmic categories have grammatical meaning, whereas characteristic expressive timing deviations from the categorical norm can embody certain "phonetic" idioms. DESAIN and HONING emphasize the importance of the local temporal context for this process. The two marked (bold italic) IOIs in figure 2.16 illustrate a characteristic situation: the nominal duration of the first IOI (4 elementary time-units) is interpreted as longer, although its performed duration is even slightly

[295] Desain and Honing, 2003, p. 363
[296] Clarke, 1987, p.22

shorter ($0.34s$) than that of the second IOI (3 elementary units, $0.35s$). The local context of proximate intervals here plays an important role. Yet, it is hardly possible to perceive categorized interval durations without context.[297] Moreover, it turns out that predictive schemes for temporal categories stem from learned metric contexts which are recalled for the interpretation of current contexts.[298] Corresponding patterns of expressive timing can also be internalized as a base for temporal expectation, judgement and performance behavior.[299]

Experimental evidence for a particular influence of metric context on categorical rhythm perception can be derived from a range of studies.[300] CLARKE showed that metric priming shifts categorization processes when the IOI ratio of a beat subdivision is in between 1:1 and 2:1. He assumed a concise categorical boundary between both ratios and, depending on the metric context, an assimilation or distinction to either a simple (1:1) or compound (2:1) subdivision. Categorical perception can be defined and demonstrated empirically by the observation of non-monotonic discrimination functions.[301] Assimilation, distinction, and a categorical boundary became evident as CLARKE's data clearly show non-linear behavior in discrimination and identification tasks. In terms of priming, ratios around the categorical boundary tend to be assimilated to 1:1 in a duple-meter context, whereas, with a triple-meter prime (compound subdivision of the tactus), the same ratios tend to be interpreted as a 2:1 ratio.[302]

To gain a more comprehensive view on rhythmic categorization and on the influence of metric priming on categorization, DESAIN and HONING constructed a timing-variation space for three-IOI patterns within a fixed time span, set to $1s$. That is, all possible patterns in which the first and the last onsets are $1s$ apart and, in between, two other onsets vary in their temporal position. They conducted two experiments where the task was to map a variety of patterns from that "performance space" to a "score space" by notating the performed rhythms in common music notation.[303] 66 different patterns were used which are evenly "spread" within the performance space to cover and represent a more ecological variety compared to CLARKE. In the second experiment, similar to CLARKE, two metric contexts (a duple and a triple meter) were provided to account for possible shifts in categorizing the patterns into the score space. Figure 2.17 shows a

[297] Desain and Honing, 2003, p. 343

[298] cf. for instance Huron, 2006. Note also the parallel to pitch perception: categorical pitch classes can be distinguished from their intonation, which is processed from the "remaining" information after the pitch-class categorization.

[299] cf. London, 2012, p. 171. For instance, when musicians try to induce the impression of a deadpan timing, they nevertheless unconsciously produce systematic deviations from ideal temporal norms which are not just imprecise but show certain biases, usually judged as unbiased by listeners and the performers themselves (cf. Repp, London, and Keller, 2011). On the other hand, a precise temporal structure rendered by an electronic realization may give rise to subjective grouping (section 2.3.1) and hence to the impression of temporal inequality.

[300] cf. Clarke, 1987, Desain and Honing, 2003, Schulze, 1989

[301] Schulze, 1989

[302] For details see Clarke, 1987, p. 26.

[303] Thus, the participants were professional musicians. The rhythms were performed by computer, using General-MIDI percussion sounds. For details see Desain and Honing, 2003, pp. 348 f.

graphical representation of the two spaces. From this perspective, categorization "implies the partitioning of performance space into a set of equivalence classes: all points in performance space mapping to the same score belong to the same class. Thus, although rhythmic categories are named or labelled by a sequence of integers, they are characterised by a region, an area in performance space".[304]

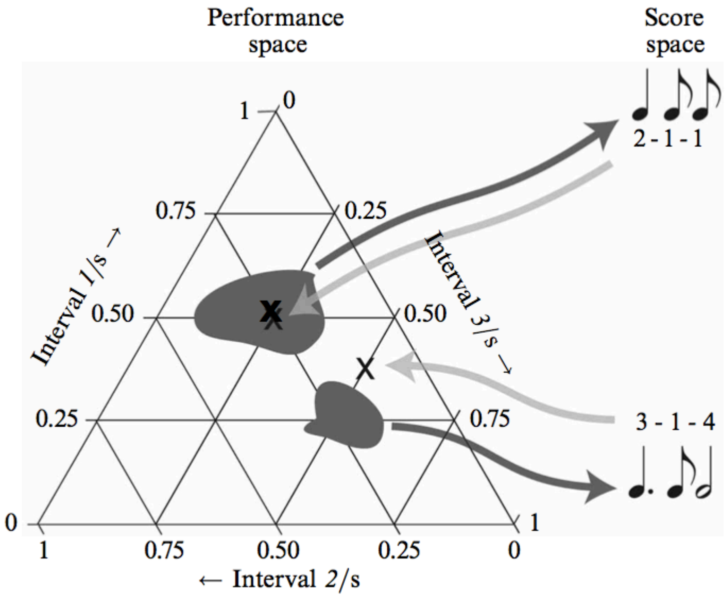


FIGURE 2.17: Performance space and score space (Desain and Honing, 2003, p. 346)

It was possible to conclude from the experiments "that the space of performances is partitioned into a small set of connected rhythmic regions"[305] in the score space. Thus, a rhythmic categorization reduces complexity and connects a perceived rhythm to a subjective system comprising musical knowledge and experience.

> The categories differ in size, according to the complexity of the rhythm, and their location is not centred around the position of a mechanical rendition of the category. The shape of a rhythmic category in performance space is quasi-convex.[306]

HURON notes that regions of simple categories like [1-1-1], [1-1-2], and [2-1-1], for which he found a high frequency of occurrence in databases of Western music, are much larger than more complex categories like [2-1-3] and [1-1-4]. Possible, even more complex or rare categories are simply never responded. "For these areas of the rhythmic map, other more common rhythms could simply encroach into these 'unused' regions."[307] Thus, the way in which the rhythmic space is "parceled" into categories

[304] ibid.
[305] Desain and Honing, 2003, p. 362
[306] ibid.
[307] Huron, 2006, p. 194

reveals a strong correspondence between the familiarity of a category and the size of its region in that space. As HURON puts it, due to statistical learning, the categorization of rhythmic patterns is influenced by the actual distribution of these patterns in a cultural context. The more frequently a rhythm occurs, the better it is processed and identified.[308]

The "large and robust effect of metric priming"[309] on rhythmic categorization found in the experimental results of DESAIN and HONING may be summarized, in regard to the purposes of this study, as follows.

- A duple-meter prime yields similar categorization responses as if there were no prime: a clear indication that listeners assume simple (binary) metric hierarchies (see section 2.3.1). Binary metric subdivision seems to be an attractor for categorizing rhythms. It may be regarded as the default "mental prime", a framework which is supposed in the absence of counter-evidence (at least in western contexts).
- A triple-meter prime considerably shifts the categorization of the performance space. Whereas simple rhythmic categorizations like [1-1-2] and [1-2-1] are supported by a duple meter, they almost completely disappear in the triple-meter condition. On the contrary, the category [1-3-2] is recognized almost only with a triple-meter prime.[310]
- Two particular categorizations are resistant in face of different metric primes. The triplet [1-1-1] is as easily recognizable with a duple-meter prime as in other conditions, probably due to its simplicity and familiarity.[311] [2-1-3] turns out as an interesting case of a metrically malleable categorization, as it is equally well identified in both metric contexts. The difference here lies in the timing preferences for each metrical interpretation.[312] "This is an important point for any theory of rhythm perception, as the interaction between the processes of rhythmic categorisation and beat induction need to be made clear."[313]

[308]ibid. and p. 201

[309]Desain and Honing, 2003, p. 362

[310]The highest differences in response proportions are between duple-meter and triple-meter primes, and between no prime and triple-meter prime. The "stimulus that is perceived most differently in the two metres is [0.210, 0.474, 0.316]. In duple metre it is heard mostly as 1-2-1 (64% of the responses; with another 10% identifying it as 1-3-2), in triple metre it is interpreted mostly as 1-3-2 (36% of the responses, not one participant identifying it as 1-2-1)."(Desain and Honing, 2003, p. 362)

[311]The same would probably hold for the duplet [1-1] with a triple-meter prime.

[312]Duple-meter interpretation prefers a slightly later timing of the second onset than a triple-meter interpretation. This can be concluded from "the relative position of the 2-1-3 area in the performance space" (Desain and Honing, 2003, p. 361. The area is centered around [0.341, 0.179, 0.480] for duple-meter primes and around [0.329, 0.196, 0.476] for triple-meter primes).

[313]Desain and Honing, 2003, p. 361

- Over all presented patterns, entropy distributions[314] of categorization correlate rather poorly between the different metric contexts. The differences generally suggest that "the process of rhythmic identification is helped by presenting an appropriate metre and hindered by another one."[315]

To this point, three factors can be specified which may interact to determine rhythmic categorization. The first lies in the subjective tendency to perceptually simplify the temporal ratios in a rhythmic pattern according to two opposite principles. Assimilation and distinction lead to durational categories which relate to each other in rather simple integer ratios like 1:1 (assimilation) or 2:1 (distinction). Second, if a metric context primes the perception of successive patterns, rhythmic categorization is facilitated by appropriate relations between the metric prime and the rhythmic pattern to be categorized. If the accentual pattern of the meter maintained by a listener contrasts with the accents suggested by a following rhythm, categorization may get difficult and is accompanied by the feeling of rhythmic tension or conflict (see for instance POVEL's experiments, described in section 2.3.3). Third, the processes of assimilation, distinction and metric priming are influenced by the familiarity of the rhythmic patterns and contexts in question. If a naturally performed rhythm is close to a familiar category of rhythmic pattern or scheme, it may be identified in a top-down fashion which may override the other mentioned factors.

CLARKE compares this process of complex interaction with "picking up the best fit for an array of perceptual information in a multidimensional space".[316] From this point of view, categorical rhythm perception can be regarded as finding "invariant properties in a multidimensional perceptual space for the durational information of a musical performance".[317] A meter may be identified as such an invariant, and, having it established, perceptual judgements on other dimensions follow.

In section 2.2.1, entrainment and pulse sensation are described as aspects of the emergence of a flexible attentional pattern which is imposed on a rhythmic succession. In a way, this activity also involves rhythmic categorization, as it is a kind of subjective measuring which constitutes meter. Indeed, pulse sensation can be regarded as a form of categorical perception, because the temporal units which correspond to single pulse percepts form rhythmic categories. Consecutive durational quantities, perceived as isochronous, are categorized likewise: isochrony is in itself a perceptual category. The

[314]Entropy here means the variety of identified rhythms/categorizations of a particular pattern (reciprocal to the amount of agreement), which is also proposed as a measure for the difficulty of rhythm identification in a certain metric context. In section 5.2.2, a measure of information entropy is also applied to distributions metric interpretations.

[315]Desain and Honing, 2003, p. 360

[316]Clarke, 1987, p. 31

[317]Clarke, 1987, p. 33. See also section 2.3 for the perspective on grouping as the "production of invariance under transposition" (Deutsch, 1982, p. 100), resembling the notion of categorical perception "as equivalent to Gibsons's idea of the perception of invariants under transformation." (Clarke, 1987, p. 33, referring to Gibson J.J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin Co.)

flexibility in pulse attribution and judgements of isochrony, mentioned in section 2.2.1, can thus be described as the categorical identification of isochrony within a rhythmic space. "Pulse is the subjective experience of isochrony",[318] as MADISON puts it.

Categorical thresholds between different rhythmic interpretations emerge from rationalizing the temporal intervals in a rhythm. They can be displaced by different metric contexts. A metric context can also be regarded as the superposition of another interpretational scheme, based on metrical accentuation. Metric contexts induce pulse sensations and regular attentional patterns which are conceptually maintained. Forthcoming events will be matched and judged according to these patterns. Hence, the metric reinterpretation of a rhythm due to its metric malleability is facilitated by a differentiation or a "decoupling" of the durational and the accentual structure of a rhythm.[319]

Pulse sensation leads to metric grouping and accentuation when it imposes a higher level of periodicity on a rhythmic surface. The most salient periodicity is the beat or tactus. From this perspective, CLARKE's experimental setup consists of metric primes for either even or uneven categorical subdivision of the beat.[320] Thus, the metric context determines the categorical boundary between even and uneven subdivision. Metrically compliant categories will preferably be perceived in ambiguous situations when events fall close to categorical borders. Without additional metrical cues there may be no further distinctions within the category of uneven subdivision. For instance, 2:1 and 3:1 could be categorically equal,[321] and a double dotted rhythm (7:1) may be considered as a sharp version of a single dotted one (3:1). A previous subdivision could nevertheless function as a more specific metric prime to distinguish different unequal subdivisions, as it indicates the relation of two adjacent metric levels. Accordingly, CLARKE reports systematic deviations in reproductions of complex unequal beat subdivisions of the category "long-short" in duple and triple-meter contexts. Reproductions of the same interval pairs consistently tended to a 3:1 ratio in a duple-meter context, and to a 2:1 ratio in a triple-meter context.[322]

LONDON further distinguishes types of beat subdivisions and connects them to the temporal constraints of metrically significant time spans (see section 2.2.2). He hypothesizes that "hearing a beat requires at least the *potential* of hearing a subdivision."[323]

[318] Madison and Merker, 2002, p. 201

[319] These two rhythmic aspects are for instance asserted and differentiated by Fraisse, 1982, p.160, and Clarke, 1987, p.23.

[320] Clarke, 1987, p. 30: "When listening to music that belongs to a metrical tradition, we perceive duration relations between notes as belonging to two basic classes: *even* and *uneven* divisions of a metric timespan." (emphases in source)

[321] In the words of Clarke, 1987 (ibid.), we perceive different uneven rhythms as different points on a continuum.

[322] Clarke, 1987, p. 29

[323] London, 2012, p. 35 (emphasis in source). This assumption is substantiated by supposing that the motor action of physical beating, comprising the two phases of "upbeat" and "downbeat", is implied in beat perception. See ibid., footnote 7 (p. 201), London, 2012, p. 38, and Toussaint, 2013, pp. 195 f.

The 100-ms limit of metrically significant durations thus functions as a bottom-up constraint for the shortest possible interval of the beat level. As a consequence, an unevenly subdivided beat needs to be longer than about 300 ms when the subdivision level shall represent a metric level. Different variants of uneven subdivision then require different temporal preconditions.[324]

It also seems to depend on its categorical identification if a subdivision of a metric unit is capable to embody a metric level, that is, to induce a pulse sensation. PARNCUTT claims that a "sequence of alternating IOIs in the ratio 2:1 appears to evoke a pulse sensation of period 1, but alternating IOIs in the ratio 3:1 (e.g., a dotted rhythm) do not."[325] This can be related to the above-mentioned evidence for a particular significance of the 2:1 ratio in categorical rhythm perception and performance timing: whereas a repetitively presented 3:1 ratio does not seem to be identified and accurately reproduced without additional context,[326] ratios of 2:1 are better processed than others in any contextual condition.[327] From the perspective of pulse salience, this supports PARNCUTT's assumption, that parent or child pulses (consonant pulses two or three times slower or faster) enhance each other (section 2.3.1). In this case, a pulse sensation on the compound (ternary) beat-subdivision level would be the reason for the better categorization of 2:1 ratios.[328] A repetitively presented 3:1 ratio does not feature a cue for a simple (binary) subdivision level of period 2, which could support the level of period 1 as a parent level. In other words, it is *metrically underdetermined* (see section 2.2.1), whereas a 2:1-ratio sequence is not. However, when pulse sensation is involved, the categorization process also depends on absolute durations of the relevant temporal intervals, as mentioned before.

Both pulse sensation and categorical perception can be regarded as foundations of metrical accentuation. The example above can be reinterpreted from this perspective, as the second event receives metrical accent in the [2-1] cycle, but not in the [3-1] cycle. Moreover, metrical accents may be differentiated to indicate a metric hierarchy which arises from their relative strengths. As demonstrated in section 2.3.2, the first event in the [2-1] cycle will be more accented if the second interval is less than about 300 ms (durational accent). In other words, the rhythmic pattern is cognitively associated with a hierarchic metric structure. In its simplest form, metric grouping and accentuation emerges as an alternation of strong and weak beats, a colloquial expression for hierarchic entrainment.[329]

[324]For details see London, 2012, pp. 34 ff.
[325]Parncutt, 1994, p. 444
[326]cf. Povel, 1984, p. 319
[327]Essens and Povel, 1985
[328]Parncutt, 1994, pp. 442 ff.
[329]Theories of hierarchic meter and computational models, involving the relative weighting of metrical accents in a hierarchy, are further discussed and developed in sections 4.2 and 4.3. See also section 2.2.1 about depths of metric entrainment (cf. London, 2012, pp. 15 ff.).

PARNCUTT explicitly examines the metric implications of categorical rhythm perception and proposes a one-to-one relation between metrical accents and temporal categories.[330] This approach integrates the discussed questions of metric levels and beat subdivisions, since metrical accents here correspond to elements of pulse sensations. Thus, for instance, the shorter IOI of repetitive patterns like [3-1] or [5-1] is not perceived as a temporal category, and the second event does not carry metrical accent, or is associated with a pulse sensation, compared to the [2-1] pattern.[331] From this perspective, metric hierarchies are resulting from the superposition of consonant pulse levels. The perception of metrical accent is thus categorical.[332]

LERDAHL and JACKENDOFF basically associate beats or metrical accents with abstract time points.[333] In their metric notation system, they are marked by dots (see figure 2.7). Though, the nature of metrical accents, discussed so far, involves temporal latitude. According to LONDON, they are embodied by attentional peaks in a regular, but flexible and continuous attentional pattern.[334] As discussed in section 2.2.1, attentional "patterns of time" (see section 2.1.1) are the cognitive foundation of pulse sensation and meter perception.[335] Thus, the notion of categorized temporal locations as metric units or "positions", which can carry metrical accent seems to be more adequate. Event onsets matching such a temporal category are associated to a pulse belonging to a certain metrical level. If an onset occurs close to, or within categorical transition zones, it can be interpreted in regard to the categorical metric context either as expressive or faulty timing (premature or slack).

To conclude this section, the crucial aspects of categorical rhythm perception in regard to meter and metric malleability are summarized in the following.

- *Entropy, complexity and conflict*: rhythms which cause a higher entropy (variance of categorical interpretation) are ambiguous and have more potential to mold into different categorical and metrical frameworks. They hence may be the most malleable rhythms. On the other hand, they may cause perceptual difficulties, as rationalization, respectively categorization, can fail or they cause conflicts with metric primes. It seems that the distinction of categorical and expressive temporal proportions is easier in the context of simpler relations. Counterexample: in *aksak* music there are many uncertainties about the conception of the rhythm

[330]Parncutt, 1994, pp. 446 f.

[331]ibid: the differences between such patterns are thus "perceived on a continuous scale."

[332]It is important here, to stress that categorical perception and pulse sensation depend on presentation rate: the mentioned patterns may all be perceived as uneven beat subdivisions, if the cycle duration is apt to induce a sense of beat or pulse. The [2-1] cycle may however be too short for a metrical categorization of the shorter interval, or too long to induce two levels of pulse sensation.

[333]Lerdahl and Jackendoff, 1983

[334]London, 2012, p. 21 and pp. 79 ff. "Although attentional peaks have some temporal spread, I will refer to them as time points in the sense that they serve to mark determinate temporal locations." (p. 83) In the current context, "determinate temporal locations" could also be expressed as categorized metric positions.

[335]London, 2012, p. 79

and the categories formed within the local cultural contexts of perception and performance (see section 2.2.1).

- *Metric priming*: "without context, categorisation is hardly possible. [...] even experienced musicians have difficulty in recognising and reproducing relatively simple ratios like 1:5 or 3:4 if they are presented in isolation. While in music practice these ratios (and even much more complex ones) are quite common, one always needs some context to be able to recognise and perform them well."[336] Appropriate metric primes effectively reduce ambiguity and "increase the conformance of the responses. Thus, presence of a metrical context eases the formation process of rhythmic categories".[337] This was shown as experimentally proved, especially by the study of DESAIN and HONING discussed in this section. Hence, metric context and priming is important for rhythmic categorization.

- *Expectancy and anticipation*: The model of expectancy developed by DESAIN (see sections 2.2.1 and 5.1.3), involves a theoretic approach to the way, a context could facilitate expectation and recognition of more complex ratios in rhythmic patterns. It could thus be employed as a common basis to model categorical rhythm perception and metrical accent.[338] It may generally be assumed that, in differentiated metric contexts, the processes of projection and anticipation become more differentiated, and result in complex and hierarchic metrical accentuation. This perspective is supported by the theories of projection[339] and DAT (sections 2.2.1 and 2.3), as well as LONDON's concept of "meter as a form of anticipatory behavior [that] finesses the problem of metrical accent",[340] already quoted in section 2.3. Against this background, categorization – as specifying temporal locations for events "marked by peaks of attentional energy"[341] – is governed by expectancy and anticipation. It is thus a part of a self-organizing, embodied activity – "a constructive process in which the mind projects a categorical pattern on an input stimulus."[342]

- *Categorization and meter induction*: nevertheless, the question of how categorization and expressive performance communication together induce meter, remains an open field for research.

> The question of the relation between the rhythmic-categorisation process and the metre-induction processes is still open. Are these processes

[336] Desain and Honing, 2003, p. 343, referring to Sternberg S., R.L. Knoll, P. Zukofsky (1982). "Timing by skilled musicians". In: *The psychology of music*. Ed. by Diana Deutsch. New York, London: Academic Press, pp. 181–239.
[337] Desain and Honing, 2003, p. 363
[338] Desain, 1992, p. 452
[339] cf. Hasty, 1997
[340] London, 2012, p. 24
[341] ibid.
[342] Toussaint, 2013, p. 305

> modular and conducted in sequential order, or are they better understood as one integrated process? [...] Our finding that the rhythmic categories depend on metre can only demonstrate that the categorisation process is open to induce metre. However, since the areas of timing that define a category for the ambiguous rhythms, which fit both metres equally well, were shown to be systematically different, the timing information itself might provide a cue for metre induction. This communication of metre by expressive timing [...][343] suggests that metre induction and rhythmic categorisation are closely interrelated processes.[344]

- *Formalization*: categorical rhythm perception can be characterized by the coherent separation and abstraction of rational rhythmic information from expressive performance variance. Thus, it can serve as a basis and legitimation for a more abstract notion of rhythm to which we resort in this study. Formal analysis and modeling should nevertheless be informed by the perceptual limits of categorization and pulse perception.[345]

2.4 Complex interaction of rhythmic components

So far, this chapter examined aspects of rhythm and meter as cognitive activities, which are relevant for the metric ambiguity and malleability of rhythm. The involved cognitive behaviors and their relation to properties of the perceptual matter[346] – the input for these processes – were analyzed and considered apart from each other. We emphasized, first, that rhythmic engagement is characterized by the cognitive integration of manifold perceptual cues. Second, that rhythmic engagement is embodied by some sort of entrainment involving levels of grouping and accentuation. Phenomenally, rhythmic gestalt is constituted by the temporal interaction of perceptual components, guided by the cognitive processes of categorical perception (section 2.3.4) and the stratification of component patterns (section 2.1).

In this section, we consider some more complex phenomena through which the interplay of stimulus components and cognitive processing can involve ambiguity. Individual reactions to ambiguity are of special interest for this study. Metric malleability can be regarded as latent ambiguity. It can become obvious in different ways, for instance as a spontaneous rhythmic gestalt flip by switching from one metric framework to another, or as a feeling of tension or conflict between a maintained framework and

[343] Reference to Sloboda J. A. (1983). "The communication of musical meter in piano performance". In: *Quarterly Journal of Experimental Psychology* 35, pp. 377–396.
[344] Desain and Honing, 2003, p. 363
[345] cf. Parncutt, 1994, p. 446
[346] The focus on temporal properties corresponds to the formal level of the studies presented in chapters 3 to 5.

current rhythmic activity. Recall that rhythmic activity can be embodied by any component, that is, by temporal patterns of changes or by other dynamic processes on a component level (see section 2.1).

As a first instance, it is "well-known that repeated patterns of pitch and rhythm can affect the perception of metrical structure."[347] TEMPERLEY and BARTLETTE developed a model of the relation of "parallelism" and meter, concluding that meter and parallelism have mutual influence onto each other. Parallelism involves not only exact repetition but also sequences (section 2.1.3), repetitions of intervallic and temporal patterns, or repetitions of melodic contour, and so forth. Parallel structures can induce meter perception, and in turn, meter can have a backward effect on the perception of following parallel structures. If a parallel goes against a perceptually maintained metric framework, it is either almost not noticed, as in the example in figure 2.18,[348] or, with "sufficient strength", it "can stand out against a previously established meter and can even cause us to modify our metrical perception accordingly."[349] TEMPERLEY and BARTLETTE illustrate this last case with the cross-figuration of four sixteenth notes against a prevailing 3/8 meter in the measures 33-35 in figure 2.19. Thus, from our perspective, the most interesting cases of parallel structures give rise to ambiguous situations, when it is not clear if and how a listener will modify his or her sense of meter.

FIGURE 2.18: Bach, Suite for Violoncello in E-flat major, Allemande, measures 5-6 (Temperley and Bartlette, 2002, p.147)

Parallel musical structures induce a sense of periodicity within the temporal range of pulse sensation and metric entrainment. Our flexibility regarding entrainment and pulse attribution also allows for the recognition of rough or distorted periodicities in quasi parallel structures. As already mentioned in section 2.2.1, BUCHLER examined erratic motivic *ostinati* in the music of John ADAMS, exploiting and challenging our flexibility to attribute a sort of "metricity" to these structures.[350] In other words, those

[347] Temperley and Bartlette, 2002, p.117
[348] Temperley and Bartlette, 2002, p. 147, note that this figure "features an exactly repeated three-sixteenth-note motive, G-F-Eb (marked with brackets), but because the two occurrences are not parallel with respect to the prevailing meter the parallelism is hardly heard. (Rather, we hear the first G-F-Eb as parallel to the F-Eb-D beginning on the second quarter note beat of the measure; and these two segments *are* parallel.)"
[349] ibid.
[350] Buchler, 2006

FIGURE 2.19: Bach, Sonata for Violin in G minor, Presto, measures 24-35 (Temperley and Bartlette, 2002, p.143)

ostinati induce a sense of polarity between strong and weak moments, a characteristic feeling of rhythmic energy, that accompanies metric entrainment.[351]

In music, temporal, melodic, and harmonic structure can induce multiple accentual patterns. The requirements and properties of temporal accents are extensively discussed in section 2.3.2. Given the notion of accent as a relative attention peak, melodic accents can arise from contrasting relations in melodic contour and interval structure. Melodic peaks, intervallic jumps, or other deviations relative to the local context call relatively more attention. Hence, in figure 2.19 the melodic accents in measures 33-35, caused by ascending melodic jumps against downward scales, elicit a periodic attentional pattern which is strong enough to produce a clearly perceived conflict with the established meter.

Tonal accents in melodies evoking harmonic implications constitute an interesting case of interplay with a listeners perceptual or musical predispositions. HURON evaluated the occurrence frequencies of melodic intervals, related to three classes of metric contexts in a large sample of western common practice tonal music.[352] Certain ascending and descending intervals occur more likely in strong-weak contexts whereas others tend to be employed for weak-strong metric positions. Except the fact that it is not clear if the metric frameworks specified by notation always correspond to perceived metrical accent structures, statistical evidence is given as follows. Descending intervals generally tend to be metrically positioned in strong-weak constellations (higher/initial pitch is metrically strong). As a major exception, the descending major seventh is predominantly used the other way around (lower/target pitch is metrically strong). Amongst ascending intervals within the octave range, a striking difference is found between the major third and the perfect forth. The latter is typically positioned for the higher target pitch to be metrically strong, suggesting a dominant-tonic closure, whereas the former most often is employed vice versa, making the lower/initial pitch metrically strong).[353]

[351] cf. Dadelsen, 1994

[352] Huron, 2006, pp. 297 ff. "The sample included nearly 6,000 melodies from the Essen Folksong Collection (Schaffrath 1995)." (p. 403, footnote 28)

[353] To a lesser extend, the perfect fifth, minor second, major sixth, and major second occur with similar tendencies like the perfect forth.

These values correspond to subjective accents which listeners perceive when stimulated by repeating isochronous patterns of major thirds and perfect fourths. "In the case of the thirds, listeners tend to hear the lower note in each third as accented. In the case of the fourths, listeners tend to hear the higher note in each fourth as accented."[354] HURON concludes that this perceptual predisposition emerges from statistical learning of the described probabilities and resultant expectancies about metric implications of melodic intervals. Specific predispositions may be acquired in particular cultural contexts. The described tonal accentuation behaviors seem to be an outgrowth of the western tradition. However, the assumption that particular, culturally embedded, musical experience shapes top-down processes of metric interpretation and expectancies may be generally valid.

HURON investigated the effect of tonal accents when they are placed in contradiction to metric expectancies. Such "tonal syncopations" – as all other types of syncopation – can create rhythmic tension against an established metric framework.[355] An extensive survey on the topic of syncopation would go beyond the scope of this study (see section 4.3.1 for a short discussion of syncopation models). This vast topic has been approached from many perspectives in music theory and psychology, and has received multiple descriptions.[356] Nevertheless, it is generally an important interaction feature of grouping and accentuation processes. In the context of this study, basically any type of subjective accentuation, which locally or punctually challenges an established metric framework, could be subsumed under the term syncopation. Thus, in contrast to the central topic of metric ambiguity and malleability, syncopation only emerges when a metric context is already active and in a robust state (see figure 2.14 in section 2.3.3).

Some general heuristic ideas may be formulated about the interplay of multiple types or levels of accentuation, and possible consequences for metric interpretation. Rhythmic tension, but also metric ambiguity and conflict, can arise from contradictory accent patterns in different rhythmic components. It was reviewed in section 2.3, that the properties of *joint accent structure* (JAS) affect processes of entrainment and metric interpretation, as suggested in the context of DAT (section 2.2.1). Specifically, "metrical clarity ratings are greater for melodies with a concordant JAS than for melodies with a discordant JAS."[357] In the terminology of POVEL, the interplay between temporal structure and pitch structure "either give rise to one and the same grid [...], or they

[354] Huron, 2006, p. 298. Starting point influences were avoided by presenting the tone sequences slowly fading in and out, "so that there was no sense of which pitch 'began' the sequence." (p. 300)

[355] cf. Huron, 2006, pp. 294 ff.

[356] See for instance Pressing, 1997, Sioros and Guedes, 2014, Temperley, 1999, Volk, 2008, Witek et al., 2014.

[357] McAuley, 2010, pp. 187 f., referring to Ellis R.J., and Mari Riess Jones (2009). "The role of accent salience and joint accent structure in meter perception". In: *Journal of Experimental Psychology: Human Perception and Performance* 35, pp. 264–280.

may yield competing grids".[358] In the latter case, the process of metric construal is less predictable. The subjective weighting of accent saliences, influencing the choice of one of the competing accent patterns, or a combination thereof, may prove decisive for meter.[359]

We could summarize as follows the most important aspects of the general perspective we have drawn in this section. Metric grouping and accentuation emerge when the metric malleability of a musical pattern is resolved through the process of metric interpretation. Metric ambiguity arises when this process can equally lead to different interpretations and attentional patterns.[360] This ambiguity can be distinguished from metric conflict and metric dissonance (section 2.3.3), as a consciously felt tension between competing interpretations suggested by the perceived musical structure. Such cases will be further discussed in the following chapter 3. Metrical dissonance, as described by KREBS, provokes a cognitive reaction which determines the ordering of dissonant levels: one has to be chosen as the metric background for the other. This process can be guided by the interaction of multiple factors.

> When none of the levels in a dissonant collection represents the primary consonance, a variety of factors might nevertheless make one of the levels predominate. For example, if only one of the levels is "in tune" with the harmonic rhythm, or if only one of the levels receives strong dynamic emphasis [...] particularly in type A dissonances, the conflicting levels might be heard as equally prominent, given that neither conforms to the primary consonance.[361]

Metrical dissonance in a musical pattern corresponds to the amenability of a perceptual flip of the metric framework.[362] Especially in sub-Saharan African traditions, rhythmic fabrics are often designed to support these kinds of gestalt flip. A metric background can spontaneously become (part of) the rhythmic foreground and vice versa.[363]

More generally, individual metric construals are determined at once by properties of the rhythmic input, and by cognitive dispositions. The dynamic process of metric interpretation involves complex interaction of these two factors. Metric priming, for instance, relies on an already established framework. However, a metric framework – that is, an attentional pattern – may not be simply triggered by preceding metric cues.

[358]Povel, 1984, p. 332. "Numerous examples can be found in music where the different determinants of the grid, temporal structure, accents, and pitch structure give rise to different grids thus creating a special rhythmical tension." (pp. 332 f.)

[359]cf. McAuley, 2010, pp. 187 f.

[360]cf. London, 2012, pp. 99 f. It was mentioned before that systematic timing nuances in performed music can prevent metric ambiguity (see sections 1.1 and 2.3.4). The involved rhythmic structures may however be amenable for metric reinterpretation (metric malleability).

[361]Krebs, 1987, note 14 (p. 120). "Additional interesting features would no doubt come to light if the relationship of metrical consonance and dissonance to pitch structure were investigated." (p. 119)

[362]cf. Flanagan, 2008

[363]See for instance Locke, 2011, and London, 2012, pp. 136 ff.

Attentional processes also refer to internalized patterns, and can moreover be voluntarily controlled. This leads to a constant interference between involuntary bottom-up processes induced by input properties, and internally activated, and sometimes conscious, top-down acts of musical engagement. Within this bundle of activities, which aspects will reach conscious observation may depend on weighting or filtering processes, which cause the transition from pre- to post-attentive stages.[364] Basically, attention and awareness hence have different temporal characteristics: conscious and active processes are slower than reactive, automated ones.

The individuality (or particularity) of attentional processes is a potential source of metric ambiguity on two levels. On the inter-individual level, it can lead to a variety of possible attentional patterns, respectively metric construals, even of rhythms which strongly suggest a specific interpretation. On the individual level, attentional patterns can be spontaneously transformed, or changed by conscious effort, resulting in a perceptual switch of the metric framework (for instance, in subjective grouping, see section 2.3.1). The former can be specified as inter-individual metric ambiguity, and the latter is a possible consequence of metric malleability.

In this chapter, insights about rhythm perception and cognition were collected and prepared as a basis for further investigations on metric ambiguity and malleability[365]. It might be obvious by now that there is no simple relation of stimulus complexity and ambiguity. Phenomena like the subjective grouping of isochronous sequences and categorical rhythm perception, show that rhythm and meter emerge from autonomous cognitive behaviors. Simple stimuli can elicit complex percepts (subjective grouping), and complex temporal structures are rationalized for economic cognitive processing (categorical rhythm perception) involving expectancy and memory. Nevertheless, on the level of temporal structure, there exists some evidence that ambiguity rises proportionally with stimulus complexity and density ("tempo", see section 2.2.2). Other levels of complexity, involving discordant interplay of components, such as a discordant JAS or tonal syncopation (see above), were only introduced in order to further emphasize what can be illustrated already on the temporal level: a quantitative heuristics about the relation of rhythmic properties and metric interpretation is necessarily limited, as the complexities discussed necessarily lead to unpredictability. Moreover, while rather schematic mechanisms of automated levels in rhythm cognition are accessible for systematic investigation, conscious structuring and interpretation of a malleable aural input by voluntary control introduces another autonomous subjective level.

Hence, in order to be able to develop reasonable propositions, the remainder of this thesis will deal with restricted aspects of rhythm in terms of temporal structure. Anyhow, it has to be kept in mind that temporal structure is embedded in multidimensional

[364]cf. Deutsch, 1982

[365]The notion of metric ambiguity, developed in this study, differentiates between metric malleability as latent ambiguity, inter-individual ambiguity and forms of metric conflict.

musical formations. In real musical engagement, listeners adapt to multiple levels of structure and resolve equivocalities by a plausible or viable mental construction. However, the prototypical level adopted in this study, corresponds to the level of many musicological rhythm analyses. HASTY, for instance, constantly refers to his awareness of the equivocality of his examples, when dealing only with abstracted temporal aspects of rhythm: "attentiveness and the countless qualifications and relevancies that are excluded from these general, schematic figures will all work together in an actual musical event to provoke a decision."[366] In other words, the interplay and the relations of all musical components can lead to an unequivocal interpretation, which moreover depends on the subjective factor of "attentiveness". Thus, by analyzing rhythm on a reductive schematic level, and by abstracting from particular components, this study deliberately focuses on temporal structure as a genuine and powerful source of metric ambiguity.

[366] Hasty, 1997, p. 88

Chapter 3

Metric ambiguity and metric malleability

In this chapter, a taxonomy is advanced which distinguishes structural types of metric ambiguity and malleability and reflects different levels of subjectivity. A systematic analysis of these aspects is necessary to establish a methodological basis for the quantitative approach to metric malleability suggested in chapter 5. LONDON's related terminology, basically introduced in section 1.1, as well includes a differentiation of metric malleability into types which are structurally correspondent to the types of metrical dissonance, forwarded by KREBS (see section 2.3.3). However, malleability and dissonance are distinguished by the fact that malleable rhythms or melodies do not *per se* cause metrical dissonance or other obvious challenges for metric interpretation. They merely have the potential to be interpreted in different metric frameworks. Indeed, metrical dissonance does not occur until a rhythmic fabric explicitly suggests contradictory frameworks with a similar vividness. LONDON illustrates the different ways in that a melody can be metrically malleable by three examples, in each of which the same melody naturally fits in two contrasting metric frameworks.[1] In the first example, a melody fits equally well in a 2/4 and a 3/4 meter. In a second example another melody starts with a series of eighth-notes. Once the first note is interpreted as an anacrusis in 4/8 meter, but the counterexample shows that the same melody can likewise begin with a downbeat in 4/8 meter. A third instance convincingly casts an identical melody in both 4/4 and 9/8 meter. Let us suppose, these alternative interpretations would emerge at the same time and compete in the mind of a listener. Then, the first and the third example represent type A dissonance (dissonance between metric period durations) and the second example correlates to type B dissonance (dissonance between metric phases).[2] These issues will be further examined in this chapter by relating the types of metric malleability, ambiguity, dissonance, and conflict to musical instances which also invite complex interaction of these aspects. The focus will however remain on abstracted temporal patterns of rhythmic cycles, including the absolute durations

[1] London, 2012, pp. 99 ff.

[2] See section 2.3.3 and Krebs, 1987.

of the temporal intervals which have proven impact on metric interpretation (see section 2.2.2).

As outlined in section 1.1, the space of metric ambiguity of a particular rhythmic pattern generally corresponds to the set of plausible metric frameworks for that pattern. Thus, the metric malleability of a rhythm is closely related to its metric ambiguity, and consequently, the types of ambiguity elaborated this chapter are also valid when regarded as types of malleability. Though, the concept of metric ambiguity may be more specifically related to the inter-individual level, that is, ambiguous rhythms are likely to be interpreted in more than one metric framework among a group of listeners. To that effect, we may attend (to the rhythms of) a musical performance

> as a collection of individuals rather than as a cohesive group. We may share an aesthetic experience together, but we will lack the sense of really "getting it" the same way, of having the music engender the same sorts of motional and emotional responses in each other.[3]

LONDON here puts the issue into terms of "individual" and "collective". If people in an audience get entrained in diverging ways by music featuring ambiguous rhythms, aesthetic experience fans out individually. Moreover, if individuals get aware of the metric ambiguity of a rhythmic fabric, it induces perceptible tension, or even metric conflict. Individual listening in turn is shaped by collective cultural heritage in form of listening experience and exposure to music. Corresponding ramifications on metric interpretation are as well considered in the following. The rhythms of some musical traditions, for instance, rather seem to be predicated on serial grouping than on hierarchic metric grouping (see section 3.2.1).

It was discussed in section 2.1 that in music, the interplay of temporal patterns and other components like pitch and dynamics often disambiguate the equivocalities inherent in temporal patterns. The same effect is achieved by "expressive" timing in musical performance which interacts with categorical rhythm perception (section 2.3.4). Though, within particular aesthetic contexts the contrary may be desired. A theoretic exploration of the metric ambiguity and malleability of temporal patterns is restricted to some level of abstraction. However, the knowledge, generated in this way, may be transferred back to "real" music to aid both analysis and composition.

3.1 Variety of metric interpretation

In section 1.1, metric ambiguity was defined as the diversity of subjective ways to perceive a rhythm, involving the establishment of different metric frameworks. This may be a simple ambivalence between two plausible meters but can also extend to a greater

[3] London, 2008

intersubjective variety – a "polysemy" of multiple metric interpretations. Accordingly, I will follow the approach of FLANAGAN "to model metrical ambiguity as the degree of variety of listener responses to a given rhythm."[4]

Meter can be regarded as a mode of attending[5] in terms of entrainment (section 2.2.1), including the "capacity for interactive sensory-motor synchronization and coordination, in a musical context."[6] From this perspective, metric ambiguity can "be considered as a partial inhibition of our capacity for entrainment"[7] which may occur in parallel with differences of metric interpretation between subjects. Hence, metric ambiguity could also be related to subjective feelings of rhythmic complexity and irregularity, as pulse sensation and entrainment may be disturbed by rhythmic complexity. Though, it seems to be more adequate to distinguish metric irregularity or "metric displacement"[8] from metric ambiguity.[9] The former challenges subjective individual entrainment while the latter causes interpretation differences on the intersubjective level. As it was reported in section 2.2.3, inter-individual differences were examined by PARNCUTT at the more specific level of pulse sensation. His study on effects of particular rhythmic patterns on tactus induction reveals intersubjective variety which is, to a certain degree, systematically related to rhythmic structure.[10] Thus, perceiving rhythmic patterns as metrically organized raises the intrinsic and latent potential of metric ambiguity, here called metric malleability.

> Any given sequence of note values is in principle rhythmically ambiguous, although this ambiguity is seldom apparent to the listener.[11]

This quotation of LONGUET-HIGGINS and LEE basically characterizes ambiguity in the sense of metric malleability. As mentioned in the introduction (section 1.1), LONDON

[4]Flanagan, 2008, p. 635

[5]cf. London, 2012, p. 4: "if 'meter [is] a mode of attending,' then rhythm is that to which we attend." (paraphrasing Gjerdingen, Robert O. (1989). "Meter as a Mode of Attending: A Network Simulation of Attentional Rhythmicity in Music". In: *Integral* 3, pp. 67–91.)

[6]London, 2008

[7]ibid.

[8]ibid.

[9]In this context, several authors, as Huron, 2006, London, 2008, and Petersen, 2010 exemplarily discuss "contrametric" structures in the music of Stravinsky: "segments of his music exhibit a systematic organization whose purpose is to actively subvert the perception of meter. Although his music remains highly rhythmic, some of Stravinsky's most distinctive passages are methodically 'contrametric' in structure" (Huron, 2006, p. 333, also cited in London, 2008).

[10]Parncutt, 1994, cf. section 2.2.3 for an overview and appendix A for the experimental data set.

[11]Longuet-Higgins and Lee, 1984, p. 424. See also Povel, 1984, p. 325: "since every pattern can be described within several grids, every pattern is ambiguous. In practice, however, ambiguity will only arise if two (or more) structural descriptions are feasible that are equally economical." London, 2012 comments on this passage that Povel, 1984 "equates metric malleability with rhythmic ambiguity; for him, such ambiguity obtains precisely when there are two or more equally plausible metric grids for a given rhythmic pattern [...]. As we will see [...], malleability and ambiguity, though related, are not quite synonymous" (Note 2 on page 199). A paraphrase of Longuet-Higgins and Lee, 1984 in Rosenthal, 1989, p. 325, states that "rhythms are inherently ambiguous".

defines metric malleability as "the property by which many melodic or rhythmic patterns may be heard in more than one metric context".[12] He distinguishes metric malleability from metric ambiguity and metric dissonance (section 2.3.3) which can force us to reconfigure and reorientate our metrical interpretation.[13]

> Some patterns strongly evoke a single metrical construal; metrically malleable patterns may evoke more than one metric framework. However, on a given perceptual occasion a musical figure can be metrically construed in only one way.[14]

LONDON develops a precise notion of metric malleability by characterizing when malleability involves ambiguity or not. Rhythmic figuration may initially be metrically neutral when it lacks of distinct cues for a particular metric interpretation. However, it may be suggestive to subjective grouping (such as in isochronous sequences, see section 2.3.1) or metrically primed by the previous context (section 2.3.3). Both cases imply metric malleability, not metric ambiguity.[15] If a performance of a malleable rhythm is capable to communicate a particular meter,[16] or if a metric context was established before, in all likelihood the listener would maintain it.[17]

> Therefore, it is not quite right to simply claim that malleable melodies are metrically ambiguous – they may be, but only under certain contextual and performance conditions.[18]

LONDON mentions "deadpan performances" of malleable passages as opportunities "for listeners to latch on to differing meters during the same performance."[19] In such cases, metric malleability becomes metric ambiguity. To this effect, the meaning of metric ambiguity gets more strict. Indeed, the notion of malleability serves to dissolve the general usage of metric ambiguity into more precise aspects, and to relate particularly to the temporal structure of rhythm, expressed on a symbolic level. Inherent ambiguities of temporal patterns are often resolved by other rhythmic components (section 2.1).[20] Nevertheless, in some aesthetic contexts it is desired to keep the meter equivocal (see footnote 16 and section 3.1.2) or to use the described potential to provoke metric conflicts in the listener (section 2.3.3).

[12] London, 2012, p. 99

[13] cf. London, 2012, pp. 70 ff.

[14] London, 2012, p. 75

[15] cf. London, 2012, pp. 13 ff.

[16] London, 2012, pp. 22, including footnote 9 (p. 200): "Part of an effective performance involves communicating [...] metric structure to the listener, especially if the musical surface is metrically malleable. This presumes that rhythmic and metric clarity is desired [...]. But in some contexts metric ambiguity may be desired; therefore one cannot simply require all correct performances to clearly project a metrical structure."

[17] cf. London, 2012, p. 100

[18] ibid.

[19] London, 2012, p. 99

[20] Longuet-Higgins and Lee, 1984, p. 425: "criteria that might lead a listener to favor a particular rhythmic interpretation."

A crucial feature of metric malleability motivating the present study is "a corollary kind of metrical ambiguity, which is the amenability of a rhythm to gestalt flip, in which a listener reevaluates the location of the beat."[21] In other words, metric malleability gets peculiarly apparent through metrical gestalt flip, transforming the latent malleability inherent in a rhythmic pattern into a dynamic perceptual phenomenon. This is another way to describe the metrical recasting of a rhythm, initially exemplified in section 1.1. In section 2.2.1, the notion of meter as an attentional framework was explicated. Temporal patterns get cast in this framework as soon it is established by the interaction of the perceptual and cognitive systems with a "rhythmic surface".[22] A metric framework holds as long it is continued and projected on forthcoming events. The projection may withstand contrasting cues, or may come into conflict with another framework suggested by these cues. Rhythmic tension mainly arises from this insisting behavior to keep an established meter going. Metrical gestalt flip can result from such conflicts or it occurs as a rather spontaneous reaction to metrically malleable input.[23] In particular cultural contexts, the repetition of appropriate rhythmic structures is a specific means to facilitate perceptual impulses of the latter kind.[24]

> There is a difference in metrical ambiguities in Western versus non-Western musics (and some Western minimalist music). In contexts involving the continued repetition of a dense, multilayered rhythmic pattern, as is common in some forms of African drumming, the continued repetition gives listeners the opportunity to metrically reconstrue the pattern.[25]

Additionally, voluntary control can be a determining factor for the dynamics of gestalt flip or metric recast (see for instance section 2.3.1). In contrast to the initial emergence of a metric framework which is usually rapid and unconscious, gestalt flip makes metric malleability obvious and amenable for conscious manipulation.

3.1.1 Polyrhythm

A polyrhythm is commonly defined as a rhythm resulting from the superposition of periodic streams in a non-integer period ratio.[26] The simplest form of a polyrhythm is thus a combination of isochronous pulses in a 3:2 ratio. This structure can also be represented by one composite rhythm cycle [x.xxx.] which comprises both pulses [x.x.x.] and [x..x..], also known as *cross rhythm*. The first attack marks the point of coincidence between the two pulses which indicates the common cycle pulse. Accordingly, a 4:3

[21] Flanagan, 2008, p. 635

[22] cf. footnote 16

[23] Gestalt flip is also known in the visual domain as a spontaneous foreground-background flip or as a flip of perspective. Two well-known examples are the *Necker cube* and *Rubin's vase*.

[24] see e.g. Locke, 2011, and Flanagan, 2008: "This amenability [to gestalt flip] has been convincingly advanced as one reason for the pull of African and African-derived rhythms on listeners." (p. 635)

[25] London, 2012, p. 109

[26] See for instance Handel, 1984, Sethares, 2007, pp. 67 ff., and Vuust and Witek, 2014, pp. 7 f.

ratio can be represented by [x..xx.x.xx..], a 5:3 ratio by [x..x.xx..xx.x..], and so forth.[27] Repetitive polyrhythms or cross rhythms, either performed by multiple voices or as a cross rhythm by one voice, can give rise to gestalt flips or metric recasting effects. In other words, they "sometimes cause perceptual shifts in which the metric model can be reinterpreted in a different way."[28] Each of the dissonant pulses can serve as the metric background for the other(s).[29]

> Polyrhythm is a type of bistable percept in the auditory domain, which relies on competition between different predictive models to achieve its perceptually characteristic effect.[30]

There may be more than two possible metric interpretations inherent in a polyrhythm. On the one hand, more complex superpositions of pulses with non-integer ratios like 3:4:5 or 2:3:7 generally afford more different pulses to serve as the tactus or metric reference level. But already in the simple case of a 3:2 polyrhythm, indicated in the first line of the listing below as [x.xxx.], there are four plausible metric references which are more or less suggestive, depending on absolute pulse durations.[31] They are listed below as isochronous rhythms with IOI durations in the range of 450 to 900 ms, which afford the strongest pulse sensations (cf. section 2.2.3). The correspondent ranges of elementary-pulse durations are indicated in parallel.

[x . x x x .]

• · · · · ·	(elementary-pulse period < 200 ms)
• · · • · ·	(elementary-pulse period ca. 200 - 300 ms)
• · • · • ·	(elementary-pulse period ca. 300 - 450 ms)
• • • • • •	(elementary-pulse period > 450 ms)

Expressed in words, the tactus reference of a 3:2 polyrhythm (or the rhythmic cycle [x.xxx.]) may be conceptualized as either of the three-pulse or two-pulse periods inherent in the rhythmic attacks, or as a "cycle beat" (with short minimal-pulse periods below 200 ms), or as a cross pulse comprising any of the attacks and two metrically accented rests. Many other meters are possible, presupposed that arising tactus rates would be moderate. Such meters may include two or more rhythmic cycles, or represent phase shifts of the tactus, as represented by the following two interpretations of the pattern.

[27] Note that this class of rhythmic cycles exhibits mirror symmetry, which gets more obvious when notated as interval sequences: [2-1-1-2], [3-1-2-2-1-3], and [3-2-1-3-1-2-3]. These sequences are "non-retrogradable" (cf. section 4.2.2).

[28] Vuust and Witek, 2014, p. 7

[29] See also Handel and Oshinsky, 1981, Handel and Lawson, 1983, and Handel, 1984

[30] Vuust and Witek, 2014, p. 10. The term "bistable percept" (cf. Vuust and Witek, 2014, p. 7) refers to Pressing, Jeff (2002). "Black Atlantic rhythm: its computational and transcultural foundations". In: *Music Perception* 19, pp. 285–310.

[31] cf. Handel and Oshinsky, 1981 and Handel and Lawson, 1983

[x . x x x . x . x x x .]

● · · · ● · · · ● · · ·

· ● · · ● · · ● · · ● ·

The first one, which is more plausible at relatively fast pulse rates, meshes three tactus periods with two rhythmic cycles. The second interpretation, although matching the cycle length of the rhythmic pattern, introduces metric off-beats. That is, the rhythm feels syncopated or shifted in relation to the metric framework. Hence, a polyrhythm as a particular case of a rhythmic cycle bears the possibility of multiple metric interpretations. However, the perceptual challenge of competing metric frameworks may not be obvious to the listener.

> Yet when confronted with such patterns, listeners do not tend to sense ambiguity, rather, they quickly focus attention on one or the other tempo and hear that as the implied beat. Perhaps even more striking is that these polyrhythms tend not to be perceived as two separate "lines" at all. Rather, they are heard as a single chunk of rhythm. Thus three-against-two is more clearly described as "long-short-short-long" [...] than in terms of its constituent parts.[32]

This statement reminds us about another way to conceptualize rhythmic structure, already discussed in section 2.3. Instead of a metric framework, the sequential or durational pattern of the rhythm itself becomes the conceptual reference pattern. In terms of sequential grouping, the cyclic pattern [x.xxx.] may be perceived just as a group of three onsets alternating with a single onset. If the elementary-pulse duration is about 200-300 ms, the two resulting types of IOI within the pattern (200-300 and 400-600 ms) are short times and long times in the sense of FRAISSE (see section 2.1.1). Hence, three onsets form a coherent group and a single onset becomes isolated. In other words, rhythmic foreground and metric background melt to a single layer of temporal structure, a similar approach like the mentioned possibility to focus on the elementary pulse. This latter behavior, however, provides a basis for the emergence of groupings or higher metric levels.[33] Consequently, grouping mechanisms and the formation of a metric framework can dynamically interact (see section 4.3.2).[34]

[32]Sethares, 2007, pp. 67 f.

[33]Higher levels are then ambiguous.

[34]Due to the restrictive view regarding only durational patterns, many complex musical features which can emerge from combined structures of polyrhythm and polyphony, are excluded from the current discussion. Nevertheless, this should be kept in mind, as for instance, it is particularly interesting to reflect "how African instrumental music exploits the many illusions that are present in aural perception". (Toussaint, 2013, p. 198, referring to Kubik G. (1962). "The phenomenon of inherent rhythms in East and Central African instrumental music". In: *African Music* 3(1), pp. 33–42.) See Toussaint, 2013 (ibid.) for a short survey of research into the perceptual and cognitive foundations of *inherent rhythms* – arising "from the interaction of different rhythms played on different instruments, or on the same instrument but with different tones" (ibid.) – such as *auditory stream segregation* and *gestalt* processing. See also section 2.1.2 about tonal illusions.

LONDON sums up former empirical research about polyrhythm perception and discerns two metric strategies. Listeners

> will either (1) extract a composite pattern of all of the rhythmic streams present and then match it to a suitable metric framework; or (2) focus on one rhythmic stream and entrain to its meter, while treating the other rhythmic stream(s) as "noise". The choice of strategy is correlated with the relative tempos of the component streams.[35]

The "either-or" perspective may be generalized to structural conceptions of any rhythmic ensemble including multilayered structures, that is, only one framework could be established at a time. Nevertheless, it may be flexible to change successively, either as an immediate gestalt flip or as different metric interpretations in the course of several listening occasions. In other words, metric malleability could become obvious only dynamically in temporal succession, not as an immediate and holistic feeling. "Thus, there is no such thing as a poly*meter*",[36] the simultaneous perception of different metric frameworks. Though, the conflicts which are described in the context of metric priming (section 2.3.3) suggest the contrary, as well as "various music-theoretic claims regarding metric dissonance and other instances in which one putatively hears two meters going on at once."[37]

VUUST and WITEK report that conscious orientation to one or the other possible metric framework in a polyrhythm gets manifested in characteristic neurophysiological activities. The "considerable effort to sustain [an] internal metric model while the rhythmic input deviates from it"[38] involves brain areas, the activities of which have been associated with language prosody and bistable percepts. The reported "studies thus suggest that these areas may serve more general purposes than formerly believed, such as sequencing or hierarchical ordering of perceptual information (BA 47) and predictive model comparisons (BA 40)."[39] The effort to establish a predictive model in such and similar cases "leads to a hightened state of arousal to search for a meaningful resolution of the musical tension."[40] In other words, ambiguity gets resolved by constructing a contextual framework. To paraphrase HANDEL's suggestion to use polyrhythms to study rhythm,[41] polyrhythms are an interesting area to study metric ambiguity and its

[35] London, 2012, p. 67. This is substantiated by quoting a communication with Jeff Pressing, who "has pointed out that even when using a composite pattern to produce a polyrhythm, performers will still favor one stream, such that is serves as the metric ground for the other. Expert performers may become adept at using either stream as the metric ground at moderate tempos, but at tempo extremes they too are apt to favor one stream as the metric ground." (note 2 at p. 201)

[36] ibid. (emphasis in source)

[37] ibid.

[38] Vuust and Witek, 2014, p. 8

[39] ibid. "BA" stands for "Brodman's area".

[40] Toussaint, 2013, p. 291, relating to the notion of *gestalt despatialization*, brought forward in McLachlan, 2000, and referencing p. 5 of Thaut, M.H. (2008). *Rhythm, Music, and the Brain*. New York, NY: Taylor & Francis.

[41] Handel, 1984

resolution by contextual perception. The heuristic approach to metric malleability in this study is similarly based on the condition that the "effect of any factor [depends] on the values of each other factor".[42]

3.1.2 Inter-cultural variety

Rhythms which prompt a variety of metric interpretations among different subjects were basically characterized as metrically ambiguous. Comparative studies provide as well substantial evidence that listener's metrical frameworks differ not only because of individual preferences. Culturally trained ways of metric coding exhibit typical preferences and tendencies.[43] This is also reflected in musical styles which developed in particular contexts. Some quotations and examples may show that musical practices deal with metric ambiguity and malleability in specific ways. It was already mentioned in the previous section that repetition is an appropriate means to allow for the possibility of metrical gestalt flip. In the context of an analysis of dance music from northern Ghana, LOCKE assumes

> the presence of the aesthetic goal of simultaneous multidimensionality; that is, the creation of a musical surface that can be heard from multiple perspectives at the same time. It asserts that this goal is achieved through systematic means; that is, through the workings of a musical syntax [...], especially the idea of meter as a matrix of beats.[44]

The concept of "simultaneous multidimensionality" contrasts with the notion mentioned in the previous section. If a simultaneous perception of different metric frameworks by one subject – a *polymeter* perception – is not possible, LOCKE's term can only be interpreted as metric ambiguity, that is, multiple perspectives of different listeners at the same time. In fact, it is moot if polymeter perception can take place and a decision seems to depend on the analytic perspective. From the perspective of metric malleability, however, polymeter perception could be interpreted as another possibility of the subjective awareness of metric malleability. Though, it would not require to define or analyze metric malleability in a different way. We will thus refrain from a detailed discussion of this dispute.

While in many repetitive African styles metric ambiguity, that is, the possibility of multiple subjective perspectives is part of the aesthetics, in Hindustani music it is foreign.

> In comparison with many other metric systems, tal is peculiarly explicit. Tal is not simply an inherent quality of the music, to be inferred by the

[42] ibid. (p. 465)

[43] See for instance Hannon and Trehub, 2005, Huron, 2006, and Temperley, 2000.

[44] Locke, 2011, p. 70. Locke, 2011 posits "the 'metric matrix' as a heuristic concept that tracks patterns of accentuation" and lists "features of musical design that open for listeners the opportunity to creatively hear music of this style." (p. 48)

> listener unaided. If this were the case listeners might find two or more metric interpretations of a single piece to be equally valid, an idea alien to Indian musical thought. In Indian music it is not uncommon for a musician to be employed primarily or even exclusively to keep tal [...]. Singers count out the tal, members of the audience join them, and there is no choice or subjectivity involved in metric interpretation.[45]

From a global, contemporary perspective, musical performance and listening might deal with meter between these two extremes. TEMPERLEY makes an interesting claim about a specific cultural difference in metric listening behaviors. Reactions to tensions between an established meter and contradicting rhythmic cues seem to differ in Western and African (sub-Saharan) perception. The former "involves shifting the metrical structure in order to better match the phenomenal accents, while the African perception favors maintaining a regular structure even if it means a high degree of syncopation."[46] These tendencies are coupled to prevailing musical properties in regard to rhythm and meter. While many styles of African music develop from "a greater ability to maintain a steady beat despite conflicting accents", Western developments entail "a greater sensitivity to metric shift in the music."[47] This notion obviously contrasts to the idea that metrical gestalt flip is a typical phenomenon facilitated by African rhythmic culture. It rather seems to reflect a way of creative hearing, an emergent phenomenon beyond specific cultural exposure. On the one hand, individual metric interpretation is always informed by a particular musical and cultural background. On the other hand, it is flexible to adapt to stylistic specialities by creative processes to make sense of the formerly unknown. LONDON's *Many Meters Hypothesis*[48] relates those aspects to the structural constraints of meter discussed in section 4.2.

> A listener's metric competence resides in her or his knowledge of a very large number of context-specific metrical timing patterns. The number and degree of individuation among these patterns increases with age, training, and degree of musical enculturation.[49]

Thus, the variety of metric interpretation has to be differentiated into inter-individual and inter-cultural variety. In contemporary globalized music perception these aspects can furthermore intertwine. These factors as well prevent an idealistic, culturally independent theory of metric malleability, and again remind us that is is more reasonable to pursue an open, heuristic approach.

[45] Clayton, 1997

[46] Temperley, 2000, p. 79, relating these differences to different weightings of the *metric preference rules* proposed in Lerdahl and Jackendoff, 1983. From this perspective, "the African mode of perception gives relatively more weight to the regularity rule, and relatively less to the accentual rules. [...] The principles involved in the two modes of perception are the same - Africans have no kind of 'sense' that Western listeners do not have; all that differs is the relative weightings of the different rules." (ibid.)

[47] ibid.

[48] See London, 2012, pp. 182 ff.

[49] ibid.

3.2 Types of metric ambiguity and malleability

This section provides a taxonomy of metric ambiguity and malleability. A differentiation of ambiguity types is suggested which includes notions about metric structure and formation, put forward in the theoretical and perceptional models, which are referred to in the course of this study. A reconsideration of *tactus* (section 2.2.3) and its inherent ambiguity will serve as a basis for further discussion of ambiguity involving a hierarchy of two or more metric levels.

As it was shown by the data collected by PARNCUTT (see figures A.1 and A.2 in appendix A), the metric ambiguity of quite simple rhythmic patterns already leads to a stochastic distribution of tactus choices by a number of subjects. The emergence of a subjective referent-time level evoked by a rhythm is influenced by a preference for pulse sensation in the vicinity of moderate tempo.[50] Tactus ambiguity corresponds to the perceived properties inherent in a rhythmic sequence. As we consider only temporal aspects of these properties, interval relations and absolute IOI durations cause durational accents (section 2.3.2). Such and other phenomenal accents[51] influence pulse sensation and entrainment, introducing metrical accentuation. Both combine to an attentional pattern (section 2.2.1) and theoretically different accents merge into a joint accent structure (JAS, see sections 2.3 and 2.4) of component accents and metric accents, a holistic rhythmic gesture or gestalt. Tactus ambiguity may also be conceptualized as a stochastic reference to rhythmic complexity (section 1.2.3).

Rhythms causing different pulses tapped by different subjects may also induce several pulses simultaneously sensed by one subject. If a rhythmic sequence induces several metrically consonant pulse sensations at a time, we perceive an alternation of strong and weak beats.[52] In other words, a hierarchic, stratified meter is established.[53] In the terminology of PARNCUTT, "tactus" and "downbeat" coincide, and thus, the tactus implies a lower level of metric subdivision (see section 2.3.4). More technically, we may assume a correlation between the emergence of a more complex meter and a flatter inter-subject distribution histogram of tapped pulses according to a specific rhythm.[54] This is supported by the observation that

> it is possible – indeed, normal – for a listener to attend to several pulses at once, noticing all of them while tapping out only one (usually a relatively slow one).[55]

[50] cf. section 2.2.3 – Parncutt's probabilistic definition of the most salient pulse sensation.
[51] For instance, melodic accents.
[52] cf. Lerdahl and Jackendoff, 1983
[53] cf. Yeston, 1976
[54] Distribution histograms and measures for the comparison of distributions are further discussed in section 5.2.2.
[55] Parncutt, 1987, p. 134

The tapping data in figures A.1 and A.2 suggest different types of ambiguity which can be systematically related to KREBS' taxonomy of metric dissonance types (section 2.3.3). The simultaneous or successive perception of metrically dissonant periodicities individually challenges the establishment or maintenance of a metric framework, as it was discussed at a global level in the previous section. The variety among tactus choices of different subjects represents those dissonances on an inter-individual level. Basically, KREBS' type A dissonances correspond to individual choices of different tactus periods in a non-integer ratio whereas type B dissonances relate to individual choices of different downbeat phases of a tactus period. Different periods, for instance, in 3:2 ratio are tapped in figure A.1 (a) and (b), as well as in figure A.2 (d), (e), and (f).[56] Varieties of downbeat phases of the same tactus period are articulated to all presented rhythms, except figure A.1 (a).[57] For instance, among the responses in figure A.2 (e), five of six possible shifts of the cycle period, two of three possible shifts of the half-cycle period, and all two possible shifts of the third-cycle period occur. Hence, both types of ambiguity can also be combined in the relations of chosen beats, like (3,0) and (2,1) in figure A.2 (e).

A third type of metric ambiguity can be distinguished in regard to the reference level. As hierarchical meter emerges from the perception of several consonant pulses, a tapped beat indicates an attentional focus on a certain metric level within a perceived metric hierarchy. The reference level can vary between subjects although several levels may be commonly perceived. With increasing event density, reference levels tend to be chosen according to higher metric levels.[58] This reflects the preference to choose a tactus in a moderate tempo. For instance, the rhythm in figure A.1 (b) provokes tappings according to three consonant levels, depending on event rate: the elementary level (1,-), the cycle level (3,0), and a hyper-metric level comprising two cycles (6,0). The reference-level ambiguity increases when event rates similarly support the perceptual salience of adjacent metric levels. Rhythm (b) for example, presented at event rate 2, induces similar amounts of tapping responses according to the elementary level and the cycle level. The same sequence at event rate 6 provokes equal responses to the cycle level and the two-cycle level. Parallel effects can be observed for all other rhythms in figures A.1 and A.2.[59]

[56] In figure A.2 (f) the ratio 3:2 occurs between tactus periods of 6 and 4 elementary pulses.

[57] The isochronous sequence in figure A.1 (a) does not exhibit a structural periodicity above the elementary pulse, and thus, a downbeat can not be shifted.

[58] See for instance the isochronous sequence in figure A.1 (a): with increasing event density the number of different tapped pulses increases. This can be interpreted as an increasing number of higher metric levels perceived. Thus, the sequence becomes the elementary level of a metric hierarchy which is the more complex the shorter the elementary pulse duration (until it reaches the metric floor of 100 ms).

[59] The tapping data for rhythms (b) and (c) do not suggest a linear relation of event density and ambiguity. Cyclically repeating durational accents guide pulse perception: the responses to (b) are least ambiguous in moderate tempo, where they mainly correspond to the cycle (3,1). Slightly more ambiguity is found at slow and fast tempos, including phase ambiguity of period 3 (rhythm (b)) and period 4 (rhythm (c)).

It is important to note that varying preferences for reference levels do not imply that different metric hierarchies are perceived. The variance rather relates to LONDON's differentiation of *tempo-metrical types*.[60] He defines metrical types as consistent metric hierarchies.

> A *metrical type* can be specified by (1) the cardinality of the N cycle, and (2) the arrangement of its component subcycles.[61]

"N cycle" indicates the elementary or "fastest" metric level, which comprises N elementary pulses. Subcycle levels include specific pulses ("time-points") of the N cycle. For instance, subcycles of an 8 cycle may be specified as 1-3-5-7 (subcycle 1) and 1-5 (subcycle 2). This results in a metric hierarchy including four levels, that is, the 8 cycle, two subcycles, and the total metric time span. Each two adjacent levels stand in a binary relation, that is, their period durations form a 2:1 ratio. Tempo-metrical types further individuate metrical types "on the basis of the absolute value of the IOIs of its N cycle and component cycles."[62] The temporal constraints and thresholds of pulse and meter perception discussed in section 2.2.2, limit the construction of tempo-metrical cycles, as it can be seen for the 8 cycle in table 3.1. Four tempo-metrical types can be distinguished, as specific temporal thresholds roughly align between columns (250 ms is the threshold for beats implying possible subdivision; 600 ms is the gravity point of pulse salience, and 2000 ms is the limit for pulse sensation.)

TABLE 3.1: Cyclical timings (in ms) for the 8 cycle (after London, 2012, table 5.1, p. 96)

N-cycle IOI	Subcycle 1	Subcycle 2	Total Span
100	200	400	800
125	**250**	500	1000
150	300	**600**	1200
175	350	700	1400
200	400	800	1600
225	450	900	1800
250	500	1000	**2000**
300	**600**	1200	2400
350	700	1400	2800
400	800	1600	3200
500	1000	**2000**	4000
600	1200	2400	4800
650	1300	2600	5200

Hence, the same metric hierarchy is temporally shifted, suggesting perceptual shifts of the reference level, as already illustrated by musical examples in section 2.2.2. This is supported by a comparison of the data in figures A.1 and A.2 with table 3.1 which reveals an increase of reference-level ambiguity in the transition zones between tempo-metrical types (dashed lines in table 3.1). The rhythms in figure A.1 (a) and in figure A.2

[60]London, 2012, pp. 94 ff.
[61]London, 2012, p. 94. Metrical types are further examined in section 4.2.
[62]ibid.

(c) provoke responses which are consistent with the 8 cycle. Fairly even distributions of tappings (thus, the ambiguity) between two adjacent metric levels in 2:1 ratio are found with N-cycle IOI close to the transition zones marked in table 3.1.[63]

The types of metric ambiguity, distinguished so far, can be summarized as follows.

- *period ambiguity*: different listeners focus on different periodicities which are in a non-integer ratio to each other (this parallels KREBS' type A dissonance)
- *phase ambiguity*: different listeners have different sensations on downbeat and upbeat locations (on the tactus level), that is, they follow parallel attentional patterns which are temporally shifted (corresponding to KREBS' type B dissonance)
- *reference-level ambiguity*: different listeners focus on different levels in a metric hierarchy, involving different perceptions of tempo (relating to LONDON's *tempo-metrical types*)

Metric dissonances, as well as tempo shifts, may be accompanied by dynamic attention shifts. They can be induced by external cues, that is, properties of a perceived rhythm, or by deliberate refocusing, that is, a conscious switch of attention. Performers can purposefully switch reference levels to perform tempo modulations, and a listener's voluntary switch of reference level can change his or her subjective feeling of tempo.[64]

As mentioned before, the indicated types of metric ambiguity can occur in combination. In figure A.2, the more complex patterns (d), (e), and (f) provoke a higher number of different tappings than the simpler cycles in figure A.1 (a) and (b), as well as in figure A.2 (c). Among the former, combinations of tapped pulses involve all specified types of ambiguity. Thus, a general tendency can be observed that metric ambiguity (and thus, metric malleability) grows proportional with the temporal complexity of a cyclic rhythmic pattern. This firstly raises the question, which kind of complexity may correlate with which type of ambiguity. Secondly, a multidimensional space of metric ambiguity involving different types of ambiguity, may serve as an appropriate conceptual framework for metric malleability. For instance, a numeric measure of malleability may increase with the number of involved types. Such issues are tackled in the sequel, particularly in chapter 5. In the remainder of the current chapter, the differentiation of the space of metric ambiguity is further developed in regard to specific aspects of meter.

[63] The following examples in figure A.1 (a) and figure A.2 (c) illustrate the reference-level ambiguities in the range of the transition zones proposed by London, 2012, pp. 95 f.:
Rhythm (a), event rate 6 (N-cycle IOI is 150 ms): 7 vs 10 responses at (2,-) vs (4,-), 77.3 % of all responses
Rhythm (c), event rate 4 (N-cycle IOI is 259 ms): 10 vs 11 resp. at (2,0) vs (4,0) and (4,2), 95.5 % of all resp.
Rhythm (a), event rate 4 (N-cycle IOI is 345 ms): 9 vs 12 responses at (1,-) vs (2,-), 95.5 % of all responses
Rhythm (a), event rate 3 (N-cycle IOI is 522 ms): 15 vs 7 responses at (1,-) vs (2,-), 100 % of the responses
Rhythm (c), event rate 2 (N-cycle IOI is 594 ms): 10 vs 9 responses at (1,-) vs (2,0), 86.4 % of all responses

[64] cf. for instance Benadon, 2004, and Jones, 1987b. See also sections 2.2.2 and 3.2.3.

3.2.1 Simple versus mixed metric frameworks

In section 1.3, a preliminary discrimination of metric types included the distinction between *simple* and *mixed* meters.[65] In some cases a rhythm may either be conceptualized in a simple or in a mixed metric framework. This introduces another type of metric ambiguity which will we discussed in this section.

In section 2.3, serial grouping was opposed to periodic grouping. The former is driven by a sequential bottom-up process and the latter involves top-down processing, as a predictive scheme is projected on the sequential process. Both principles interact and the relative influence and importance of each in the whole process of rhythmic interpretation depends on the listener.[66]

Grouping generally implies that we cognitively organize events hierarchically. At first, grouping adds an individual quality to each element, depending on the position within the group. For instance, the first and last event in a group may obtain different perceptual saliences than events in between, due to the group context and independent from their inherent perceptual qualities.[67] Secondly, when elementary groups form a larger group, we possibly organize groups of events hierarchically, introducing another level of depth to the event hierarchy. These two aspects of perception are involved in both serial and periodic grouping. Periodic groups are hierarchically structured according to a metric framework, whereas serial grouping implies a hierarchy of events only determined by their sequential temporal relations. Thus, serial grouping takes place independently and unaffected by the detection of periodicity.

Serial grouping is "additive" in the sense that events accumulate to form perceptual groups. Mixed (or additive) meter implies a recurring structure or succession of such accumulative groups, though, not of events but rather of isochronous, elementary time units. We can for instance notate a $\frac{3+3+2}{8}$ time signature to indicate a measure by accumulation of three elementary pulses, followed by an identical group and a shorter of two units. The measure establishes another metric level which is isochronous, unlike the lower mixed level.[68] Hence, from a cognitive point of view, a mixed metric framework involves serial grouping of time units in a metric context.

Alternatively, a metric hierarchy built up solely by uniform multiplication of an elementary time unit is obviously more compatible with periodic grouping. The same holds for the inverse establishment of a hierarchy by equal division of larger units. If

[65] These terms are favored by Gotham, 2015b to distinguish meters comprising different period durations in single levels from those consisting entirely of isochronous levels. Other terminologies exist in parallel, like *isochronous* versus *non-isochronous* meters (London, 2012 also uses the abbreviations I meter and NI meter), and *divisive* versus *additive* meters (cf. footnote 86 in section 1.3). Metric types are explored in depth in section 4.2.

[66] cf. Parncutt, 1994, p. 412

[67] cf. Povel and Essens, 1985, Povel and Okkerman, 1981, and Yu, Getz, and Kubovy, 2015

[68] In section 4.2.1, a more differentiated terminology is introduced to describe types of metric hierarchies. Isochronous *levels* are distinguished from non-isochronous *layers*.

we mentally divide evenly by four the metric period of eight units which arose from the addition $3+3+2$, we construct a metric framework which will differently guide the grouping processes on the smaller scale within that period. Such a temporal division embodies a cognitive top-down strategy in metric-processing behavior. Instead, cognitive accumulation (addition or multiplication) of temporal units may be associated with cognitive bottom-up processes.

Consequently, an ambiguity or variance of metric interpretation can occur in terms of a bottom-up or a top-down strategy. The interval pattern [3-3-2], for example, may be interpreted in a bottom-up fashion, that is, the rhythmic articulations fall together with the peaks of metric attention (see the next section 3.2.2). Alternatively, the second articulation may be interpreted as syncopated in an isochronous metric scheme.[69] Then, a metric accent, or a "virtual beat" is interpolated by the listener, as illustrated below.

[x . . x . . x .]	rhythmic pattern
• · · · • · · ·	metric reference pulse

Another hint for this – or at least a similar – type of ambiguity comes from a different perspective. An interval theory developed by LEWIN draws on the distinction that

> our perception of relation among musical durations is at least as much additive as multiplicative, particularly in the foreground of rhythmic textures. That is, given durations s and t, we often perceive t as so much longer (or shorter) than s by a certain *difference* $t - s$, rather than so many times as long (or as short) by a certain *factor* t/s.[70]

This relates the currently discussed type of ambiguity back to the interplay of serial and metric grouping because the same issue can be put into terms of perceptual grouping and *rhythmic contour*. These are "obtained by coding the change in the durations of two adjacent inter-onset intervals using 0, $+1$, and -1 to stand for equal, greater, and smaller, respectively."[71] TOUSSAINT compares the two rhythmic cycles [2-2-3-2-3] (*Fume-fume*) and [3-3-4-2-4] (*Son*) and observes that they have an identical rhythmic contour: [0, $+1$, -1, $+1$, -1].

> Rhythmic contours are relevant from the perceptual point of view because humans have an easier time perceiving qualitative relations such as "less than" or "greater than" or "equal to" than qualitative relations such as the second interval is four-thirds the duration of the first interval.[72]

[69] London, 2012, p. 153, notes that placing the [3-3-2] pattern "in the context of 4/4 implies that (1) performers should count and listeners should entrain under an I meter, and (2) one will hear the second note as syncopated. And indeed, this is often how this rhythm may be heard, especially by performers who are reading in 4/4. However, it is also possible to hear the three-beat pattern without syncopation; at times performers may use one metric framework while listeners use another."

[70] Lewin, 1980. p. 246

[71] Toussaint, 2013, p. 47

[72] ibid.

This is essentially the same assertion as LEWIN's above and also reminds us to the constraints of categorical perception, discussed in section 2.3.4. Therefore, the two rhythms, played at a fast speed, sound almost the same.[73] In other words, their perceptual grouping is the same: "one group of three onsets followed by a space, followed by one group of two onsets followed by another space."[74] This perspective interestingly contrasts to the perspective of metric grouping. If the two rhythms are interpreted within simple metric frameworks, they offer very different possibilities. Whereas the rhythmic cycle of the Son has 16 elementary pulses, the Fume-fume is based on a 12 cycle. Thus, the Son can only be cast in a binary meter like 4/4, but the Fume-fume is much more malleable to be interpreted either in 12/16, or 6/8, or 3/4, if we assume the sixteenth-note to represent the elementary time unit. This depends on the *cardinality* of its N cycle, as the 12 cycle entails three possible structures for simple meters, and the 16 cycle allows for only one.[75] This is thoroughly discussed at a formal level in section 4.2.

At this point, let us take a look at some structural properties of rhythmic cycles which affect metric ambiguity, particularly in regard to simple versus mixed metric interpretation.

- Rhythmic cycles with prime cardinalities > 3 strongly suggest mixed metric frameworks, as simple, divisive metric cycles entail non-prime cardinalities. Thus, they are less amenable to be interpreted within simple metric frameworks.

- The rotational position (see section 4.1.3), that is, the phase of the rhythmic cycle in regard to the *clasp* (section 3.2.5), the starting point or "downbeat" of the cycle, seems to be an important determinant, whether a specific rhythms are perceived as syncopated patterns in a simple meter, or as rhythmic manifestations of mixed metric beats (see also the next section 3.2.2). In other words, the

  > rotation of a beat cycle can influence its metric plausibility. Some rotations of a non-isochronous pattern are more likely to be construed as surface rhythms of an I meter than as congruent with the IOIs of the beat cycle of a NI meter, whereas others more naturally lead to NI meter construals. The presence or absence of a half-measure seems to be an important factor in this regard.[76]

[73]ibid.: "the two rhythms are so similar, they can easily be interchanged during the performance of a piece, as is done in the *Highlife* music of West Africa". This also implies shifting the meter from ternary to binary or vice versa (cf. the discussion of "polymeter" above).

[74]Toussaint, 2013, p. 298, substantiates this view by quoting experimental results reported by Handel: "'Two rhythms that had the same perceptual grouping were judged as being identical, even if the timing between the groups was different.'" (p. 497 of Handel, S. (1992). "The differentiation of rhythmic structure". In: *Perception and Psychophysics* 52, pp. 492–507.)

[75]cf. London, 2012, pp. 155 ff. and pp. 168 ff.

[76]London, 2012, p. 170

Thus, particular properties of rhythmic patterns may suggest mixed metric frameworks whereas others rather give rise to simple ones. However, depending on a musical context, many patterns can be heard either as syncopated within an I meter construal, or as matching non-isochronous beats in a mixed metric framework.[77] The model of malleability presented in chapter 5 suggests an approximative method to estimate probabilities of possible frameworks of both types, which may be evoked by a particular rhythmic cycle. According to the model, these probabilities depend on the accentual structure of a rhythm (see sections 2.3.2 and 5.2.1). Though, it is important to bear in mind that contextual factors like metric priming or complex component rhythms can have crucial and decisive influence on metric interpretation. In other words, the inherent structure of a rhythm is only one parameter beside others in the process of meter emergence.

The perception of mixed meter in terms of non-isochronous metric beats usually requires an explicit feeling of the subdivision level. This is an important aspect because temporal ratios between metric beats are categorically interpreted as isochronous or non-isochronous. Under certain conditions, adjacent-integer ratios like 4:3 or 3:2 can even be categorized as 1:1 (section 2.3.4). It can therefore be crucial to focus on a lower, categorically isochronous level to realize the categorically distinct durations of mixed-meter beats, like 3+2 or 3+2+2, and so forth. SETHARES expresses this by using the terms "additive" and "divisive rhythm", as well as "tatum" for the isochronous succession of elementary time units.

> In the additive perspective, there is a regular grid of short equidistant segments of time called the *tatum*. Notes are felt against the tatum by adding together the total number of tatum elements in the duration of the note. [...] The divisive perspective emphasizes the relationship between the notes of the rhythm and the perceptual beats.[78]

The latter perspective – in which the mentioned relationships occur as on-beat or off-beat notes – is regarded as the perceptual alternative. "Thus the 'same' rhythm can be thought of as either additive or divisive."[79] That is, both ways of metric interpretation may compete as they can be induced by the same rhythm. The correspondent type of ambiguity can therefore be related to the reference-level ambiguity, discussed in the previous section. To get back to the example above, we can think about other mental pulses and their impact on the perceptual gestalt of the rhythm.

[x . . x . . x .]	rhythmic pattern
• • • • • • • •	elementary pulse
• · • · • · • ·	interpretive pulse 1 (–> simple meter)
• · · • · · • ·	interpretive pulse 2 (–> mixed meter)

[77] cf. London, 2012, pp. 144 ff.
[78] Sethares, 2007, p. 58. See also section 4.1.2, footnote 33.
[79] ibid.

The upper two of the three pulses illustrated above, are consonant to each other and to the pulse in the previous example, constituting a full binary metric hierarchy altogether. Suppose, the focus initially lies on the elementary pulse. Pulse 1 may then emerge and the elementary pulse is felt as the subdivision level, that is, pulse 1 would become the (isochronous) reference level. The second onset of the rhythmic pattern would be felt "off-beat" in this case. Though, the initial focus on the elementary pulse can alternatively function as a reference and subdivision level for a non-isochronous categorization of metric beats which exactly correspond to the [3-3-2] pattern, that is, pulse 2. The metric reference level would correspond to the interpretive levels in both cases, but in the latter, the elementary pulse has to be explicitly maintained. Therefore, rhythmic cycles like [x.xx..xx] and [x.xx.xx.][80] may better support the mixed metric framework [3-3-2] than [x..x..x.]. They have correspondent accentual structures while providing a more pronounced rhythmic embodiment of the elementary pulse level.[81] Thus, the decision for metric interpretation resolving simple-versus-mixed-meter ambiguity may depend on different temporal focuses.[82]

As it was outlined before (for instance, in section 3.1.2), a number of studies suggest culture-specific preferences for particular metric schemes. These behaviors can also be specified in relation to simple-versus-mixed metric interpretation. "Western notions of measure and meter are often thought of as divisive",[83] that is, simple meters are the prevailing schemes among a larger number of possible metric frameworks.[84] In contrast,

> musics in various cultures – notably west African and Balkan – exhibit non-periodic metric and beat patterns. That is, there exist entire musical genres that are not based on tactus-level periodicity.[85]

HANNON and TREHUB have experimentally demonstrated that subjects enculturated with mixed meters of the Balkan region have a greater ability to detect a shift from 7/8

[80]This pattern is also a palindrome as it gets obvious by regarding its interval pattern [2-1-2-1-2].

[81]London, 2012, p. 129, notes the importance of the elementary level's rhythmic presence for the categorical perception of mixed metric frameworks and draws a relation to *maximal evenness* (cf. section 3.2.2), another requirement for mixed metric beat durations: "The presence of isochronous subdivision in an NI meter lets us know that the beats, although uneven, are still as evenly spaced as possible. Knowing how many elements of subdivision each beat contains, also lets us know which beats are categorically equivalent and which are not. [...] It is also perhaps for this reason, that in most music that involves an NI meter, the N cycle (i.e. the beat subdivision) tends to be phenomenally and continuously present, providing a constant underpinning for the uneven beats."

[82]See section 2.2.3: beside the discussed data of Parncutt, 1994 (see appendix A), empirical evidence is found that subjects with musical expertise tend to focus on higher/slower metric periods near the *preferred tempo* (section 2.2.2). See for instance Eck, 2001 and McAuley, 2010, both referring to McAuley, J. D. and Semple, P. (1999). "The effect of tempo and musical experience on perceived beat". In: *Australian Journal of Psychology* 51(3), pp.176–187.

[83]Sethares, 2007, p. 57

[84]In section 4.2.2, it is demonstrated that from the combinatorial point of view, mixed meters constitute the vast majority of all possible metric frameworks.

[85]Huron, 2006, p. 201. Although "nonperiodic" (equatable with "non-isochronous") tactus levels are essential for mixed meters, it should be added that mixed meters are categorically periodic on the elementary and the cycle level (see section 4.2.1).

to 4/4 meter in a rhythmic sequence than others.[86] Different sensitivities for mixed meter due to learned schemes may also cause differences in metric interpretation, in cases where both simple and mixed metric frameworks would be plausible. At least, the outlined perspective on mixed metric interpretation as a bottom-up process – like serial grouping focussing on the accumulation of elementary pulses – should be modified. Particular listeners may in fact activate top-down schemes of mixed meters when appropriate, while others do not.

Another instance of the equivocality between simple-meter and mixed-meter perspectives is found in the debate about African metric perception. TEMPERLEY, referring to NKETIA and others, characterizes the "richness" of African rhythm as "conflicts" between rhythmic events and "an (often implicit) framework of equally spaced beats".[87]

> But this does not necessarily imply that events on weak beats are to be understood as syncopations, that is, as displaced events which belong on some other beat. Some studies have pointed to quite different ways of understanding African rhythms; for example, in terms of additive rather than divisive structures [...].[88]

That is, if mixed meter arises, non-isochronous metric accents prevent the perception of syncopation. Experimental investigation has revealed the subtlety and flexibility involved in the cognitive organization of rhythm by African musicians.[89] The "asymmetric" structure of African timelines (section 3.2.2) may be an outgrowth of this cognitive practice, "which is context-dependent, and linked to specific training regimes."[90] In the context of evaluating the model of hierarchic meter by LERDAHL and JACKENDOFF[91] for its applicability to African music, some authors review the debate if African concepts of rhythm include hierarchic meter or not.[92] Given the diversity of opinions, it can at least be assumed that a variety of "preconditions" for metric interpretation exist among different cultures. More specific, the intercultural diversity of attentional strategies underlying metric interpretation add another dimension to the simple-versus-mixed aspect of metric ambiguity.

[86]Hannon and Trehub, 2005 also found that infants (6 to 7 months of age) show this ability as well. They conclude that the sensitivity for mixed metric structures gets either lost – because simple, prevailing schemes are learned – or amplified by enculturation processes.

[87]Temperley, 1999, pp. 38 f.

[88]ibid., referencing Nketia, J.H.K. (1974). *The music of Africa*. New York: W. W. Norton Co.

[89]cf. Magill and Pressing, 1997

[90]ibid., p. 189

[91]cf. Lerdahl and Jackendoff, 1983

[92]cf. Temperley, 2000, pp. 68 ff. and Toussaint, 2015, pp. 3 f.

3.2.2 Meter versus rhythm

> It is common in many musical traditions for metre to be defined by a repeated rhythmic pattern, and for the performers of such music, metre without rhythm is unlikely to be a meaningful concept. (This intimacy also causes much theoretical confusion, to the extent that in some musicological works the distinction between rhythm and metre is drawn loosely if at all.)[93]

It was already indicated in section 1.3 that in musicological literature, rhythm and meter are often described as oppositions. “Arguing in favor of merging the two concepts, Hasty [...] reviews the historical bifurcation of meter and rhythm that has characterized music theory studies of Western art music.”[94] This theoretical problem was resolved for the context of this study by characterizing meter as an attentional framework for the structural interpretation of rhythm (see sections 2.2.1 and 2.3). A rhythm, that is, a temporal pattern of articulations or onsets, can naturally coincide with a predictive attentional scheme or an induced pattern of beats. Then, “the rhythm is perceived as reflecting its underlying metric organization.”[95] Nevertheless, an underlying meter can be confused with – or is not clearly distinguished from – the rhythmic surface. Examples of mixing up these terms are occasionally found in the literature, as mentioned in section 1.3. A first guess about such motivation addresses the regularity of a rhythm. It was shown, for instance, by HANDEL, and surveyed in section 3.1.1, that polyrhythms (several isochronous rhythms in a non-integer temporal relation) are perceived both as temporal frameworks and rhythmic figures. In other words, their metric ambiguity is resolved in one or the other way, providing both a *figure* and a *ground*.[96]

> What is common to both roles is the ubiquity of regular meter rhythms. This may be due to the simplicity of meter rhythms or may indicate a basic rhythmic predisposition [...].[97]

Hence, a regular rhythm may be regarded as a *meter rhythm*. In contrast, TOUSSAINT, discussing the rhythmic pattern [3-3-2-2-2],[98] notes that “this irregular rhythm is also

[93]Clayton, 1997 (p. 15 of the translation)

[94]Locke, 2011, p. 52, referring to Hasty, 1997, pp. 3-21.

[95]Vuust and Witek, 2014, p. 2

[96]Clayton, 1997, for instance, adopts Kolinski’s notion of meter as “the ‘ground’ to rhythm’s ‘figure’”, and characterizes aspects of the Indian tal system as the “clearest possible endorsement of Kolinski’s view, in that Indian musicians clearly separate rhythm and metre conceptually in a manner analogous to figure and ground in Gestalt psychology.” (pp. 4 and 15 of the translation, referring to Dowling, W. J. and D.L. Harwood (1986). *Music cognition*. London: Academic Press; and Kolinski, M. (1973). “A cross-cultural approach to metro-rhythmic patterns”. In: *Ethnomusicology* 17/3, pp. 494–506.)

[97]Handel, 1984, p. 483, referring to Martin, J.G. (1972). “Rhythmic (hierarchical) versus serial structure in speech and other behavior”. In: *Psychological Review* 79, pp. 487-509

[98]This pattern plays a significant role in the further discussions throughout this study. Occurring in diverse musical contexts (see sections 3.2.4, 4.2.2, and 5.3), it exhibits some formal features which are examined in the following, and which support *hemiola* (Temperley, 2000, p. 80, see section 3.2.4) and

the meter of the guajira style of the flamenco music of southern spain."[99] This can be interpreted, as already quoted in section 1.3: "Well-known rhythms (such as dance patterns) act in a manner very similar to meters."[100] Well-known rhythms thus become predictive metric schemes and a "positively valenced predictive reward is misattributed to the [rhythmic] stimulus".[101] This cognitive amalgamation of rhythm and meter could be a reason for the mentioned theoretical confusions. As mentioned in section 3.1, LONDON disentangles both by paraphrasing GJERDINGEN: "if 'meter [is] a mode of attending,' then rhythm is that to which we attend."[102] Nevertheless, using again the metaphor above, a rhythmic figure can contrast with its metric ground, melt together with it, or something in between can happen. Different intensities of this specific figure-ground contrast may successively occur in the course of the perception of a musical development.

Polyrhythm perception was characterized as a particular instance of a bistable figure-ground relation (section 3.1.1), that is, metrical gestalt flip occurs when the rhythmic figure becomes the metric ground and vice versa. Therefore, confusion between rhythm and meter can also be interpreted as *figure-ground ambiguity* or *rhythm-meter ambiguity*. This may particularly be an aspect of mixed metric interpretation, as mixed metric beats (and their categorically distinct subdivision units) tend to be rhythmically articulated. As we saw in the previous section, this may be necessary to distinguish a mixed meter from a simple one which can be easier articulated by a syncopated rhythmic structure. However, meter embodiment by a rhythmic texture can be simple, that is, close to the metric structure like in a *meter rhythm*, or more complex by means of differentiation from the suggested metric framework.

> In some cases there is no distinction between pulses prescribed by the signature and beats – for example, in a 4/4 measure containing four quarter notes. Where there is a distinction, the difference between pulse and what I am calling beat is often regarded as the difference between meter and rhythm [...].[103]

For HASTY, a "beat" exclusively terms an articulated strike. A metrical beat, which is only felt, he calls a *virtual articulation*, which is "no less real (though less vivid) than a sonic articulation."[104] The "difference between rhythm and meter" is thus felt as a *rhythmic tension*, caused by the difference between sonic and virtual articulations.

metric malleability. To that effect, it will be identified as a prevalent instance of a malleable rhythmic necklace in section 5.3.1 (about *rhythmic necklaces* see sections 1.1 and 4.1.3).

[99] Toussaint, 2013, p. 100. See also pp. 266 f., and pp. 204 f. where [2-2-1-1-2] is called a metric pattern. However, this pattern can only be characterized as rhythmic, as it conflicts with *metric well-formedness constraint 3.4* in London, 2012, p. 92 (cf. section 4.2.1). From this point of view, the rhythm overdetermines a duple meter.

[100] Huron, 2006, p. 202

[101] ibid.

[102] See footnote 5 in section 3.1.

[103] Hasty, 1997, p. 129

[104] Hasty, 1997, p. 130

The more (or less) difference the more (or less) rhythmic tension may be felt. This notion may indicate a broader perspective on the issues of syncopation and *off-beatness* (see section 4.3.1). However, HASTY's concept of *metrical particularity* approaches the relation of rhythm and meter from another point of view. By typecasting meter,

> we run the risk of loosing sight of what is rhythmic about meter. If any measure is reducible to an instance of a type an thus to a typical organization of equal beats and if meter is equated with this "underlying" organization, the uniqueness or particularity of any actual measure will be viewed as a product of rhythm and not meter. The problem here is not the identification of type but the reification of the type or an identification of the type with a particular instance.[105]

From my point of view, this sequence puts HASTY's approach to "meter as rhythm" in a nutshell. His characterization of meter as projection largely corresponds to the notion of meter as evolution and maintenance of an attentional pattern by projecting metrical accents or attentional peaks that match a perceived rhythmic structure. Though, *metrical particularity* represents a conceptual alternative instead of the separation of rhythm and meter. It thus allows to resolve rhythm-meter ambiguity on this conceptual level but it also reflects the intricate relation between a rhythmic percept and a metric framework. "If meter is identified with projection, there will be no reason to identify meter with bar or to presuppose an invariant procession of equal beats. In this case, the rhythmic particularity of a bar will be inseparable from its metrical particularity."[106]

Formal representations of rhythmic and metric structures are similar and based on correspondent elements. Investigating rhythm on a pulse-based,[107] symbolic level already implies a basic form of metric interpretation by categorization or "metrification" of elementary time units. Being aware of the difference between rhythmic (articulatory) and metric (attentional) patterns, the symbolic patterns we use to notate rhythmic cycles can also be interpreted as meters. However, their structures have to fulfill constraints of metric well-formedness (section 4.2.1) like HANDEL's "meter rhythms" would do. Therefore, the [3-3-2-2-2] pattern mentioned above can in fact be interpreted as a meter. Basically all metric patterns are also rhythmic, as they can be rhythmically articulated. On the other hand, not all rhythmic patterns are metric on a higher level than the elementary level, as it is discussed in section 4.2.

Rhythm-meter ambiguity may be particularly invited in the context of mixed metric frameworks. 9 cycles, for instance, may naturally suggest a simple triple meter (three beats with ternary/compound subdivision respectively). Though, for instance, the rhythmic pattern [2-2-2-3] first establishes a feeling of simple (duple) meter which is

[105] Hasty, 1997, p. 148
[106] Hasty, 1997, p. 149
[107] Here, pulse does not mean a perception, but an isochronous elementary time unit.

thwarted by the last interval completing the cycle length.[108] It is therefore less likely that this rhythm will be interpreted in a [3-3-3] framework. It may rather induce a feeling of mixed meter by being structurally identified with the meter, that is, the rhythm exactly matches the induced metric beats.[109] This kind of identification is vital to establish a feeling of mixed meter, especially in contexts where simple metric frameworks are also plausible (see the previous section 3.2.1).

A related structural reason why rhythm and meter melt together in this example is the pattern's property of *maximal evenness*, which is also constitutive for metric well-formedness as amplified later. Thus, maximally even rhythms embody correspondent meters. Nevertheless, such rhythms do not prevent metric ambiguity. The contrary is true as we saw in the previous section by means of the [3-3-2] cycle. However, temporal regularity or evenness is a general feature of metric patterns. They naturally follow expectancy schemes consisting of temporally balanced attentional peaks.

> Our attentional expectations involve a tradeoff between when things have happened in the past and when we expect them to happen in the future.[110]

An emergent meter hence balances the projection or extrapolation of a perceived pattern with our ecological predisposition for an even temporal distribution of attentional energy. Aspects of evenness also belong to the prevalent topics of formal investigation into rhythm and meter. Due to isomorphic pattern structures (see section 4.1.1), those studies also benefit from explorations regarding the pitch domain. Indeed, the principle of "maximal evenness was developed in the study of scales and pitch-class sets"[111] by CLOUGH and DOUTHETT. If a rhythmic or metric pattern is represented by a set of points on a circle, as for instance by means of necklace notation (see section 1.1), a valuable measure for evenness is "the sum of all pairwise Euclidean distances between points on the circle".[112]

LONDON develops a sophisticated analysis of the relation of maximal evenness and metric well-formedness. Accordingly, the demand for maximal evenness of a single

[108] As already mentioned in sections 2.2.3 and 2.3.2, the rotational starting point of a rhythmic cycle plays an important role for metric interpretation. Thus, it would provoke a significant perceptual difference if this pattern would occur in its rotational form [3-2-2-2] (see also sections 3.2.5 and 4.1.3).

[109] Similarly, 11 cycles are of prime cardinality and thereby exclude simple meters. Basically, the possible mixed meters matching this cardinality can be represented by the two necklaces [2-2-2-2-3] and [2-3-3-3] (for details see section 4.2.2, particularly figure 4.4 and related discussion). Both structures, however they may be rhythmically embodied, roughly resemble or seem to suggest a simple meter, as they contain longer sequences of simple metric successions (2-2-2-2 and 3-3-3, respectively). These are thwarted once per metric cycle by a single, categorically different metric beat, similar to the 9-cycle example.

[110] London, 2012, p. 131

[111] London, 2012, p. 126, referring to Clough, John and Jack Douthett (1991). "Maximally Even Sets". In: *Journal of Music Theory* 35 (1 and 2), pp. 93–173.

[112] Demaine et al., 2009, p. 432, referring to Block, S. and J. Douthett (1994). "Vector products and intervallic weighting". In: *Journal of Music Theory* 38, pp. 21–41. They point out "that the mathematician Fejes-Toth showed in 1956 that a configuration of points on a circle maximizes this sum when the points are the vertices of a regular polygon." (ibid., referencing L. Fejes Toth (1956). "On the sum of distances determined by a pointset". In: *Acta Mathematica Hungarica* 7 (3/4), pp. 397–401.)

metric level can be relaxed in cases where that allows for a hierarchic manifestation of maximal evenness in metric structures.[113] For instance, the [3-3-2-2-2] pattern discussed above,[114] although not maximally even, suits these guidelines as it establishes an isochronous half-cycle level [6-6] which supports the metric stability of this pattern. More precise, as a metric framework, the pattern's rotations [3-3-2-2-2] and [2-2-2-3-3] may be preferred over the others because even half-measures would emerge that stabilize the feeling of the cycle downbeat.[115] The occurrence of even half-cycles may also have impact on simple-versus-mixed-meter ambiguity because they naturally support simple-meter construals.[116] On the contrary, rhythmic cycles which lack of temporal intervals marking half the cycle length strongly suggest mixed metric frameworks. This property – implying that the number of elementary pulses in a cycle is even – is known as *rhythmic oddity*, a term introduced by AROM.[117]

More general measures of evenness that are based on discrete mathematics were employed in computational and comparative studies about rhythm.[118] For the current context, it is instructive to distinguish between maximal evenness and perfect evenness. Consider an N cycle that contains M beats (either "articulatory" in case of a rhythm, or "attentional" in case of a meter). The beats can be only be distributed perfectly even on the cycle when N is a multiple of M.[119] Then, maximal evenness equals perfect evenness, and a perfectly even temporal pattern has an isochronous structure. In all other instances, a maximally even pattern exhibits some imperfection as the [2-2-2-3] pattern above which can be described in terms of $N = 9$ and $M = 4$. Thus, in simple meters, the periods of any metric level are perfectly even, whereas in mixed meters they may be maximally but not perfectly even.[120] An additional constraint for metric well-formedness is that $2M \leq N \leq 3M$, to guarantee that – on any level –metric groupings include two or three elements of the next lowest level. This constraint

[113] cf. London, 2012, pp. 125 ff. and pp. 155 ff. For a summary see p. 169. To define a correspondent and appropriate metric well-formedness constraint (WFC), London, 2012 states "WFC 4.2.2: If the beat cycle is NI, then either (1) it is maximally even or (2) the cycle above the beat cycle, in most cases the half-measure cycle, must be maximally even." (p. 129) On p. 158 he comments: "WFC 4.2.2 allows us to take a broader view and consider maximal evenness as it is hierarchically manifest, that is, not just maximal evenness on each individual subcycle but also for the meter as a whole."

[114] cf. footnote 98

[115] cf. section 3.2.5

[116] See quotation on page 107 (footnote 76).

[117] As further discussed in section 4.1.3, Chemillier and Truchet, 2003 (p. 261) and Toussaint, 2013 (pp. 85 ff.) list the patterns [3-3-3-2-3-3-2-3-2] and [3-2-2-2-2-3-2-2-2-2-2] (comprising 24 elementary pulses in a cycle), as well as the 12-pulse pattern [3-2-3-2-2] as examples of timelines used by the Aka Pygmies of Central Africa, satisfying the oddity-property. The latter two are also maximally even, and – in its rotational form [2-2-3-2-3] – the last pattern is further discussed below as one of the "standard" patterns in in sub-Saharan African rhythm. Hall and Klingsberg, 2004 have extended the notion of rhythmic oddity to other possible uniform partitions of a cycle, e.g. "third-cycle oddity", and so forth. Toussaint, 2013 (ibid.) proposes to calculate degrees of oddity, reflecting the number of structural violations in regard to this property – that is, the number of "equal bipartitions" in a rhythmic pattern – to estimate the amount of irregularity in that pattern.

[118] cf. for instance Gomez-Martin, Taslakian, and Toussaint, 2008, Taslakian, 2008, Toussaint, 2005

[119] cf. London, 2012, p. 129

[120] "Period" here denotes the duration of a metric unit on a particular metric level.

works together with maximal evenness to avoid structural ambiguities in a metric hierarchy.[121]

Much research about maximally even rhythms has concentrated on the condition that M and N are relatively prime, that is, their greatest common divisor equals 1. Some well-known patterns comply with this condition, for instance, the maximally even rhythmic timelines [2-2-3-2-3] ($M = 5$ and $N = 12$) and [2-2-1-2-2-2-1] ($M = 7$ and $N = 12$).

> The most illustrious ternary rhythm timelines in sub-Saharan Africa and the Caribbean have either five onsets or seven onsets, with durational patterns [2-2-3-2-3] and [2-2-1-2-2-2-1], respectively, in a cycle of 12 pulses. [...] Both rhythms have been called the *standard* pattern in the literature.[122]

These patterns play an important role in the following examinations.[123] Their interval structures are equivalent – or *isomorphic*, see section 4.1.1 – to the pentatonic scale and to the diatonic scale (in *Ionian* mode), respectively.

TOUSSAINT has introduced the notion of *Euclidian rhythms*, describing and generating maximally even rhythms by a procedure derived from the *Euclidian algorithm*.[124] Euclidian or maximally even rhythms that follow the rule $2M \leq N \leq 3M$ also fulfill LONDON's *metric well-formedness constraints*. For instance, the patterns [2-3-3], [2-2-3-2-3], and [2-2-2-3] are Euclidian rhythms and also well-formed, mixed metric structures. Accordingly, we can approach a more precise notion of *metric* rhythm, that is from our perspective, a rhythm which prevents the distinction of rhythmic articulation and metric accent. Basically, a metric rhythm is a precise articulation of the tactus level of the metric framework in that it is heard. The condition $2M \leq N \leq 3M$ approximately holds when a rhythm neither overdetermines ($M \leq N < 2M$) the meter, nor underdetermines it ($3M < N$), depending on the type of meter (duple, triple or mixed). These properties are also expressed by the mentioned metric well-formedness constraints, including a hierarchical manifestation of maximal evenness. These are further discussed in section 4.2. Nonetheless, a metric rhythm can be metrically overdetermined when its accentual structure clearly pronounces the metric beats, as for instance the pattern [2-1-2-1-3].[125]

[121] London, 2012, p. 92, expresses this in his *well-formedness constraint 3.4*: "Each subcycle must connect nonadjacent time points on the next lowest cycle." See also section 4.2.1.

[122] Toussaint, 2013, p. 45 (emphasis in source), referring to Agawu, K. (2006). "Structural analysis or cultural analysis? Competing perspectives on the 'standard pattern' of West African rhythm". In: *Journal of the American Musicological Society* 59.1, pp. 1–46; to p.51 of King, A. (1960). "Employments of the 'Standard Pattern' in Yoruba music". In: *African Music* 2.3, pp. 51–54; and to p. 53 of Kubik, G. (1999). *Africa and the Blues*. Jackson: University Press of Mississippi.

[123] cf. for instance sections 3.2.3, , 4.1.3, 5.2.4, and 5.3. See also footnote 194 in section 2.3.

[124] Toussaint, 2005

[125] Toussaint, 2013, p. 240, considers this pattern while discussing "odd rhythms": "The slip jig (also called hop jig) has a nine-pulse meter with durational pattern [2-1-2-1-3]." Again, the distinction between rhythm and meter is vague in this case.

As outlined before, metric rhythms are ubiquitously employed in the context of many musical idioms. For instance, *aksak* rhythms are commonly conceptualized as consisting exclusively of IOI comprising two and three elementary time units, "which Arom calls *binary* and *ternary* cells."[126] Therefore, they precisely represent mixed metric frameworks and are sometimes called aksak meters.[127] Though, Arom's classification of aksak rhythms incidentally tells us something about their metric ambiguity: *pseudo-aksak* and *quasi-aksak* rhythms, that fit in even and non-prime odd N cycles respectively, like [3-3-2] (pseudo) or [2-2-2-3] (quasi), may also give rise to regular, isochronous metric beats, as opposed to *authentic* aksak cycles of prime cardinality like [2-2-3]. Note that in the case of pseudo-aksak, a two-pulse beat (or a three-pulse beat in case of the quasi-aksak 9 cycle) is easier to perceive because it is consonant to the cycle length. In contrast, a two-pulse beat applied to [2-2-3] would only be consonant to two cycles, that is, the full metric cycle would include two rhythmic cycles. This corresponds to an interesting type of complex metric ambiguity which includes reference-level ambiguity (section 3.2.3). Either, the [2-2-3] pattern can naturally induce a beat coinciding with the first two intervals, which results in an oscillation of three on-beat articulations and three off-beat articulations, spread over two rhythmic cycles.

[x . x . x . . x . x . x . .]
• · • · • · • · • · • · • ·

Or, the pattern is alternatively interpreted in the mixed metric framework which is exactly embodied by the rhythmic strokes. If we think of the elementary pulse or the subdivision level as corresponding to the eighth-note level, the first interpretation can be conceptualized as a 7/4 meter. The second would match a 7/8 framework respectively. Both frameworks include mixed metric grouping, but the 7/8 clearly promotes a feeling of non-isochronous tactus whereas the 7/4 beats are even and the mixed metric structure is felt more subtly above the reference level.[128]

African and African-derived *timelines* and *claves* are as well cyclic rhythmic patterns which predominantly carry out a metric function.[129] That is, they are employed as a temporal framework for complex, polyphonic ensemble performance. Although not representing the meter itself, timeline designs tend to support orientation regarding the structure of the metric cycle. In other words, they provide phase information related to the location of the cycle downbeat, which may be conventionally known or

[126]Toussaint, 2013, p. 40, referring to Arom, Simha (2004). "L'aksak: Principes et typologie". In: *Cahiers de Musiques Traditionelles* 17, pp. 12–48.

[127]In section 2.2.1, it is noted that aksak "cells" could also be performed as a sort of "block timing" without categorizing the elementary pulses. This also involves considerable deviations from the nominal 3:2 ratio between the durations of the rhythmic cells.

[128]See also http://www.bradmehldau.com/rock-hemiolas. Last retrieved on 12 October 2020.

[129]cf. Colannino, Gómez, and Toussaint, 2009: "*Claves* are rhythmic patterns repeated throughout a piece whose main functions include rhythmic stabilization as well as the organization of phrasing".

initially derived from the rhythmic structure of the timeline. Consider again the "standard" pattern [2-2-1-2-2-2-1] discussed above. As this timeline features a low degree of symmetry,[130]

> each position [...] is unique with respect to its intervals with other positions [...]. Thus it is possible for one to orient oneself to the pattern simply from the intervals [...] presented. [...] This has possible relevance to [...] meter-finding. Once it is conventionally established that a certain position in the standard pattern is the "downbeat," then one could orient oneself metrically to whatever is going on simply by locating that position in the rhythm and considering it the downbeat.[131]

This has an interesting relation to the emergence of durational accents (section 2.3.2) in mixed metric structures, where different metric durations occur at a particular metric level.[132] The pattern [2-2-2-3] for instance, perceived as tactus, suggests a stronger accent on the longer beat, provided a relatively fast presentation rate that promotes a clear phenomenal difference between the two metric categories. To that effect, the distinctive rhythmic fabrics of timelines support recognition and metric orientation, specifically when the duration of a rhythmic cycle does not exceed that of the perceptual present. At the same time, typical timeline structures feature a high degree of metric ambiguity as there are different possible locations on the cycle to metrically "latch on to".[133] In respect of the cycle level, this rather leads to metric phase ambiguity (section 3.2.5) than to metric period ambiguity (section 3.2.4), as the cycle length is unequivocal due to a non-isochronous profile of the rhythmic intervals. Thus, we perceive "the same overall measure period and pattern of beat periods; we differ only with respect to the subjective sense of accent at each beat."[134] Though, regarding lower metric levels, the metric ambiguity of some timelines can cover both mentioned aspects. For instance "the standard pattern is almost maximally ambiguous, as it samples several different meters (12/8, 6/4, and 3/2) – and different phases of those meters – almost equally."[135]

These aspects bear different conditions for performers and listeners. In a genuine sense,

[130] In fact, some amount of symmetry can be verified for this pattern when regarded as cyclic. Then, it holds translational (or rotational) symmetry related to the cycle length and mirror symmetry when rotated so that it starts from the second interval, yielding [2-1-2-2-2-1-2]. For further discussion of symmetry in rhythmic and metric patterns see sections 4.1.3 and 4.2.2.

[131] Temperley, 2000, p. 82. Temperley, 2000 further notes that it is unclear "wether the standard pattern serves this function for African listeners [...]. If so, it suggests a factor in African meter perception which is quite unlike the other factors proposed by GTTM [Lerdahl and Jackendoff, 1983]: a *conventional* cue to meter, which relies simply on the listener's knowledge that a certain position in the pattern is conventionally metrically strong." (ibid.)

[132] cf. London, 2012, pp. 136 ff.

[133] London, 2012, p. 139

[134] ibid.

[135] Temperley, 2000, p. 81, referring to pp. 46 f. of Pressing, J. (1983). "Cognitive Isomorphisms between Pitch and Rhythm in World Musics: West Africa, the Balkans and Western Tonality". In: *Studies in Music* 17, pp. 38–61.

a timeline is used to coordinate independent voices, the players of which establish different metric orientations related to the timeline. Thus, it organizes different temporal layers which combine to iridescent polyphonic fabrics. Such metrically equivocal structures may be differently perceived by performers and listeners. Whereas musicians concentrate on individual voices, the recipient's perspective suggests multiple possible metric frameworks.[136] As indicated before, this can lead to inter-individual diversity of metric interpretation and to individually perceived phenomena of equivocality, such as gestalt flip.

Some pieces of Steve REICH specifically demonstrate such impact of timeline-derived textures on meter perception at a cross-cultural level. In *clapping music*, two instances of the rhythmic 12 cycle [xxx.xx.x.xx.] are presented simultaneously by two players. The piece is simply shaped by the consecutive rendition of all twelve phase relations between the two voices.[137] In this way, the line can be perceived as an underlying metric framework and as a contrasting rhythmic figure at the same time. Similar to a polyrhythm, the figure-ground relation becomes malleable. The overall structure and the phase shifts invite metrical shifts as the different constellations of the two canonic voices keep the the following aspects in abeyance: (1) if the meter, at a specific moment, will be derived from one voice or the other, or from a cross rhythm emerging from both; (2) if a chosen metric framework or tactus can be maintained or is forced to be reconfigured after the phase constellation, and thus the rhythmic texture, has changed; (3) wether spontaneous perceptual gestalt flip will occur due to the malleable pattern structure.

COLANNINO et al.[138] conducted a *phylogenetic analysis* of the different rotations of REICH's pattern in terms of the differences which occur between the two rhythmic instances in each phase relation.[139] In this study, the measure of *directed-swap distance*

[136]cf. Locke, 2011, p. 54: "African musicians [...] feel beats in relation to sounded patterns, such as a time-keeping ostinato, or in Nketia's [...] widely adopted terminology, 'time line'." (referring to Nketia, J.H.K. (1963). *African Music in Ghana*. Evanston: Northwestern University Press.) Patel, 2008, p. 98, summarizes that in Ghanaian drumming, (1) "the basic rhythmic reference is a repeating, non-isochronous time pattern played on a set of hand bells", and ensemble drummers "keep their rhythmic orientation by hearing their parts in relation to the bell, rather than by focussing on an isochronous beat"; (2) the beginning of a cycle does not coincide with a strong metric beat, rather, "the most salient beat comes at the end of the cycle" (referring Temperley, 2000); and (3) the resulting music "emphasizes diversity" in metric interpretation. The polyrhythmic texture provides "alternative perceptual possibilities depending on the rhythmic layers and relationships one chooses to attend to" (referring to Locke, David (1982). "Principles of Offbeat Timing and Cross-Rhythm in Southern Ewe Dance Drumming". In: *Ethnomusicology* 26(2), pp. 217–246; and Pressing, Jeff (2002). "Black Atlantic rhythm: its computational and transcultural foundations". In: *Music Perception* 19, pp. 285–310).

[137]Steve Reich, *clapping music for two performers* (1972), Universal Edition, London. The twelve phase relations are based on the coincidence of the twelve elementary pulses.

[138]Colannino, Gómez, and Toussaint, 2009

[139]cf. Toussaint, 2013, p. 269: "Phylogenetic analysis techniques have been used for quite some time now to study the evolution of cultural objects." They include the generation "of *proximity graphs* that provide not only an effective method for visualizing the distance matrix [a table of quantified differences], but yield additional structures such as clustering relationships and information for the reconstruction of possible ancestral rhythms."

was employed to compute transformational distances, which basically sum the necessary onset movements on the elementary-pulse grid to transform one pattern into the other. It turned out that the hypothetic "ancestral" rhythm that minimizes its maximum distance to any rotation of the pattern equals [2-1-2-1-2-1-2-1].

> This fundamental rhythmic pattern is none other than a group of trochees. A trochee is a rhythmic grouping consisting of a long note followed by a short note. This "ancestral" rhythm has a strong metric time-keeping character. The trochee, expressed in box notation as [x.x], is a common Afro-Cuban drum pattern, also found in disparate areas of the globe.[140]

This finding supports the idea that the rhythmic textures – though highly malleable – are metrically balanced over the course of the piece. As mentioned, timelines indirectly represent and are suggestive about meter, and in this case, a hypothetical meter may be suggested throughout a piece.

In sum, timelines are illustrative to show that, under certain conditions, non-isochronous rhythmic patterns can be considered as "metric". This is particularly the case when a well-known pattern evokes a conventional metric scheme. Though, by no means this prevents from metric ambiguity, as many common timelines include equivocal metric cues.

3.2.3 Reference-level ambiguity

This and the next two sections further illustrate and differentiate the three types of metric ambiguity which were preliminarily distinguished in section 3.2. We start here with *reference-level* ambiguity, followed by *period* ambiguity and *phase* ambiguity.

It is demonstrated in section 3.2 that a rhythm which evokes several metric levels, can involve inter-individual diversity in choosing the tactus or reference level. The individual resolution of this ambiguity affects tempo sensation. The close relation of preferred tempo and pulse salience (which approaches a maximum at a resonance period around 600 ms) was discussed in section 2.2.3. We concluded that tactus periods are preferred in a range between 400 and 900 ms.

> However, in many individual pieces the tempo is ambiguous, allowing several metric levels to serve as the beat. Quite often a majority of the listeners judges a metric level far from the resonance frequency most salient. [...] In quite many cases, the tempo will even be set at a value that is a multiple or divisor of a tempo in the preferred range.[141]

[140]Colannino, Gómez, and Toussaint, 2009, p. 123
[141]Moelants and McKinney, 2004, p. 558

Hence, there may be instances where periodicity is present in the preferred tempo range, but the accentual structure rather suggests a lower or higher level to serve as the metric reference. MOELANTS and MCKINNEY note that this ambiguity occurs both with music and artificial stimuli and test the predictability of the resonance model of van NOORDEN and MOELANTS[142] regarding the variance of tempo responses to a corpus of individual pieces of music. In terms of ambiguity, they conclude that if "the tempos of adjacent metrical levels straddle the resonant tempo, the perception across listeners would tend to be split evenly across those tempos and be more ambiguous."[143] This corresponds to the observation of PARNCUTT and GOTHAM that the overall metric salience will be maximized in this case (see sections 2.3.1 and 5.2.3). At the same time, the possible tactus period lies on either side of the most salient pulse tempo. However, a specific distribution of phenomenal accents in music can profoundly distract tempo perception from the resonance principle.[144]

The tactus level is also not meant to be defined in advance by LONDON's tempo-metrical types, discussed in section 3.2. In fact,

> performers and listeners have some flexibility as to which level they may hear as a tactus in a sufficiently rich metric context, especially at moderate tempos.[145]

Recall that, in the context of mixed meters, reference-level ambiguity implies that the tactus can be either construed as isochronous or non-isochronous (see the previous section 3.2.2). In a moderately fast 5/8-meter, the tactus may be felt on the isochronous measure level, or a non-isochronous 2+3/8 or 3+2/8 duple beat may emerge. Recall as well that the latter choice depends on a sufficient articulation of the eight-note level, as otherwise a categorical distinction of two beat classes is problematic. A slower 5/4-meter may again suggest an isochronous beat at the quarter level, resulting in non-isochronous super-tactus groups on the half-measure level. For HASTY, additive meters imply a challenge for metric projection. For instance, "a second measure of 5/4 demands considerable reinterpretation",[146] due to an "ambiguity in the projective 'hierarchy'".[147] This may introduce another specific impact on the choice of tactus in addition to the absolute durations of the projected intervals and the rhythmic texture.

Reference-level ambiguity can facilitate dynamic perceptual shifts of the tactus across metric levels. In addition to inter-individual variance of tactus perception, this implies an individual experience of contrasting successive frameworks which can also be forced as a musical effect of tempo modulation (section 4.4). Corresponding practices

[142] Noorden and Moelants, 1999, cf. sections 2.2.3 and 5.1.3.
[143] Moelants and McKinney, 2004, p. 561
[144] Moelants and McKinney, 2004, p. 562
[145] London, 2012, p. 96, referring to Meyer, Rosalee K. and Caroline Palmer (2001). *Rate and Tactus Effects in Music Performance* (Unpublished MS).
[146] Hasty, 1997, p. 145
[147] ibid.

are reported across cultures. LOCKE addresses examples of those by the term *tactus augmentation/diminution*, that is, "the rate at which beats are present to musical attention varies by factors of two, a process akin to 'double time' or 'cut time' in African diasporic music like jazz."[148] A musically contrasting example of the same priciple is found in common Hindustani rhythmic practices.

> Indian percussionists are well aware that perceptual shifts are sometimes needed when playing the "same" *tala* at widely different tempos. [...] At high speeds it may be necessary to conceptually shift the tempo downwards by factors of two until a more moderate tempo is achieved.[149]

This relates to figure 2.9 in section 2.2.2, where the flexibility of tempo perception in the face of dynamic variations of a rhythmic texture is already discussed.

Tactus ambiguity may also basically account for diverging opinions about reference levels in African-originated musical styles. TEMPERLEY[150] reviews analyses of Ewe drum ensemble music, based on the "standard" pattern [2-2-1-2-2-2-1], already discussed in the previous section (3.2.2). In this music, the pattern functions as the timeline which is played on a bell. Three different levels are assumed as the beat level by different authors: the elementary pulse ("eighth-note level", twelve beats per cycle), the quarter-cycle comprising three pulses per beat ("dotted-quarter-note level", four beats per cycle), and the half-cycle comprising six pulses per beat ("dotted-half-note level", two beats per cycle). Although "the strongest support is given to the dotted-quarter level"[151], the tempo indications of the discussed authors vary from 80 to 170 beats per minute for this level. This suggests at least two levels as possible references with a salient (moderate) pulse rate. A "slower" reference level with beat class 6 may introduce additional metric ambiguity on the subdivision level (see the next section 3.2.4). The opinion that the elementary-pulse level corresponds to the reference level is specifically motivated by the assumption that it is "the *only* level of meter in drum ensemble music."[152] This notion also contributes to the debate if African rhythm can actually be understood in terms of hierarchical meter (see section 3.2.1).

3.2.4 Period ambiguity

In section 3.2, *period ambiguity* was distinguished from reference-level ambiguity by the type of ratio between different pulses to be considered as the tactus. To that effect,

[148] Locke, 2011, p. 55
[149] Sethares, 2007, p. 67
[150] Temperley, 2000, pp. 69 ff.
[151] ibid.
[152] Temperley, 2000, p. 70 (emphasis in source), referring to Koetting, James (1970). "Analysis and Notation of West African Drum Ensemble Music". In: *Selected Reports in Ethnomusicology* 1(3), pp. 116–146; Pantaleoni, Hewitt (1972). "Three Principles of Timing in Anlo Dance Drumming". In: *African Music* 5(2), pp. 50–63; and Pantaleoni, Hewitt (1972). "Towards Understanding the Play of Sogo in Atsia". In: *Ethnomusicology* 16(1), pp. 1–37.

it was associated with inter-individual differences of listeners' attentional focuses on periodicities which stand in non-integer ratios to each other. Thus, period ambiguity parallels metric type A dissonance, defined by KREBS (section 2.3.3). In contrast, reference-level ambiguity precisely applies to consonant metric levels, the periods of which are integrally related. Moreover, a notion of period ambiguity only makes sense when the tactuses in question are categorically isochronous. In fact, mixed beat durations inhibit the determination of metrical consonance and dissonance, and of temporal ratios between their oscillating periodicities. In a way, they are in themselves metrically dissonant, or include non-integer period ratios.

According to COHN, the term *hemiola* "refers to any successive or simultaneous conflict between a bisection and trisection of a single time-span. Thus a hemiola can be said to arise whenever pulses in a 3:2 ratio are perceived to conflict."[153] This concept may be generalized to conflicts of pulses in any non-integer ratio to include different variants of polyrhythm (cf. section 3.1.1). Indeed, the concepts of hemiola and polyrhythm can be regarded as structurally identical while focusing on different phenomenal aspects. Whereas the notion of polyrhythm is primarily analytic concerning a particular type of temporal complexity, hemiola rather describes the synthetic perceptual effect of conflicting metrical accents. In section 3.1.1 it was illustrated that the polyrhythmic pattern [x.xxx.] may be metrically casted in a two-beat or in a three-beat measure. Consider now the more complex case that two cycles of the pattern are combined in a single metric 12 cycle. Then, three possible simple meters could emerge which correspond to COHN's graphical notion of a double hemiola, depicted in figure 3.1.

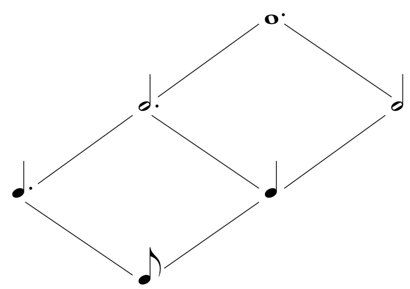


FIGURE 3.1: Double hemiola diagramed in the form of a *ski-hill graph* (Cohn, 2001, Ex. 6, p. 303).

The three meters are represented in this so-called *ski-hill graph* by the possible "routes" which can be taken "to ski down the hill." Or in terms of a map, moving southwestward implies a duple subdivision of an initial time span, whereas moving southeastward is associated with a triple subdivision. For instance, moving SW-SE-SW from the initial dotted-whole-note span corresponds to a 6/4 meter (2 x 3 x 2). The graph thus illustrates the potential of metric conflict – or in parallel, of period ambiguity – regarding possible simple metric hierarchies in a 12 cycle. A 12/8 meter (2 x 2 x 3) is most

[153] Cohn, 2001, p. 295

dissonant to a 3/2 meter (3 x 2 x 2) as the two "routes" on the graph are different on two levels. They only meet at the start (cycle span) and the end (elementary pulse). These meters conflict at both the half-note and the quarter-note level and thus constitute a double hemiola. The other metric relations correspond to simple hemiolas because the 6/4 meter mediates or minimizes metric dissonance.

The [x.xxx.] cycle completely articulates all four periodicities which are shown in figure 3.1 in between those of the 12 cycle (the whole-note pulse) and the elementary (eighth-note) pulse, that is, the periods 2, 3, 4, and 6, expressed as multipliers of the elementary pulse. Three of the six pairs in this set are metrically dissonant: 2 versus 3, 3 versus 4, and 4 versus 6. These pairs are also instances of period ambiguity, as the rhythmic pattern equally affords all four pulses to be perceived as the tactus period.[154] More precisely, it can be argued that their saliences very much depend on absolute pulse rates (pulse-period salience) when all pulses are rhythmically articulated (see pulse-match salience, in sections 2.2.3 and 5.2.1). Interestingly, the 3-versus-4-period ambiguity is closely related to the double hemiola, whereas the other dissonant pairs can also occur in the context of simple hemiolas. To that effect, a tactus shift from the dotted-quarter period to the half-note period – the most "distant" periods in figure 3.1 which also hold the most complex ratio – has the most contrasting perceptual effect on the rhythmic gestalt of the [x.xxx.] pattern. More generally, if period ambiguity or a metrical conflict implied by a hemiola involve tactus shifts by non-integer ratios, the perceived tempo gets affected, as it was also noted in the previous section (3.2.3) for the context of reference-level ambiguity. In terms of KREBS' taxonomy, such tactus shifts correspond to "indirect", that is, successive, metric dissonances of type A.[155]

TEMPERLEY reviews notions of hemiola as a "centrally important aspect of African rhythm, [...] an implied shifting between two meters, most often 3/4 and 6/8 [...]. This is often reflected in a simple alternation between quarters and dotted-quarters".[156] As an interesting coincidence, the pattern [3-3-2-2-2], already introduced in section 3.2.2, can serve as a basic example of hemiola ("the 'hemiola' pattern").[157] In the context of a simple isochronous meter it can be interpreted as metrically ambiguous at the tactus level, if the two-pulse and three-pulse intervals lie in an appropriate temporal dimension to compete as the reference pulse. From a mixed-meter perspective, the rhythmic pattern gets *metric* (see section 3.2.2), precisely embodying the mixed metric beat. This would imply less metric ambiguity, as the pattern stands in a one-to-one coincidence to the non-isochronous tactus and higher-level pulses corresponding to the half-cycle and the cycle would naturally emerge.

[154]The other, metrically consonant pairs of the set represent possible instances of reference-level ambiguity (2 vs. 4, 2 vs. 6, and 3 vs. 6).
[155]cf. Krebs, 1987, p. 108, and Cohn, 2001, p. 296
[156]Temperley, 2000, p. 80
[157]ibid. In African contexts, this "pattern is a common accompaniment pattern for songs (often expressed in clapping or work-related actions)." (ibid.)

Also in rhythmic 16 cycles, hemiolas are common "in numerous musical contexts, from Irish reel to Brazilian bossa nova."[158] For instance, the rhythmic pattern [3-3-3-3-2-2] "is relatively easy to hear [...] as a series of syncopations or hemiolas against"[159] a pure binary metric hierarchy, the only possible simple metric hierarchy corresponding to a 16 cycle (see section 3.2.1). In this case, the rhythmic figure "metrically 'rights itself' at the end of each measure."[160] This kind of adjustment is inevitable as the 16 cycle does not allow for complete polyrhythmic cycles like the ones descibed in section 3.1.1. Superpositions of isochronous sequences in non-integer ratios necessarily entail other common-cycle cardinalities when they are synchronized according to an elementary-pulse grid, because 16 never occurs as a common multiple of any pair of integers which constitute a non-integer ratio. Thus, within metric 16 cycles, hemiolas either do not cover the whole cycle, like in the example above, or get disrupted. Alternatively, in the context of mixed metric frameworks the hemiola can disappear, for instance when the [3-3-3-3-2-2] rhythm exactly coincides with the metric beat pattern (see section 3.2.2). However, metric interpretations like in figure 3.2 may not be far-fetched. That is, the first part of the [3-3-3-3-2-2] pattern may feel like a "two-against-three" hemiola when primed by a mixed meter based on another elementary level (eight-note) than that of the rhythmic pattern (sixtheenth-note). Thus, different types of metric ambiguity may combine in the context of hemiola, as the discussed examples potentially involve simple-versus-mixed ambiguity and reference-level ambiguity beside period ambiguity.

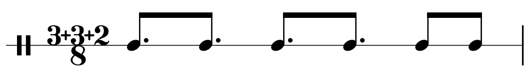


FIGURE 3.2: Hemiola in a mixed metric context

Finally, the notion of period ambiguity particularly recalls the question of polymeter perception (see sections 3.1.1 and 3.1.2) as the related discourse concentrates on simultaneous cues for pulses in non-integer ratios. For instance, TEMPERLEY refers to JONES who

> suggests that, in pieces with pervasive cross-rhythms and hemiolas, African listeners are able to perceive both duple and triple meters simultaneously [...]. I am doubtful about this. In my own experience, one of the striking things about metrical perception is the difficulty of perceiving multiple interpretations at once. It is relevant to consider, however, how *different* the

[158] London, 2012, p. 164
[159] ibid.
[160] ibid.

> conflicting metrical structures are in such cases. [For the [2-2-2-3-3] pattern],[161] the two different possible metrical structures (3/4 and 6/8) share both the eighth-note and dotted-half-note levels; only the intermediate level is different. In such a case, then, it seems plausible that both structures might be entertained simultaneously, or at least that it would be relatively easy to switch back and forth between them.[162]

3.2.5 Phase ambiguity

To determine the *phases* of a rhythmic cycle, an orientation point has to be perceptually identified according to which they can be related. The temporal category coinciding with this moment corresponds to the *cycle downbeat* which indicates the metric period of the cycle level. In other words, the temporal span of the cycle is marked by a perceptual *clasp*[163] (a "zero phase"), provided that the span can be recognized within the duration of the psychological present (see sections 2.2.2 and 2.2.3). The clasp is often determined by the first attack of a rhythmic pattern which is either perceived as the downbeat or as (part of) an anacrusis. Beat-induction models define criteria or suggest rules for the inference of meter from such initial rhythmic cues (section 5.1). The cognitive mechanisms which are involved to establish period and phase of a metric beat work very fast.[164] As already discussed, an important criterion are the relative durations of the IOI, present in a rhythmic structure. Relatively longer IOI have more impact on phenomenal accentuation of their initial onsets when the shorter IOI are below 250 ms (see section 2.3.2).[165]

Though, how do listeners determine a subjective starting point when a pattern is presented cyclically without a perceptible beginning?[166] As an aspect of grouping, this issue was addressed in experimental studies and theories of *auditory parsing*. Yu et al.

[161] Temperley, 2000 refers to this pattern as three quarter notes followed by two dotted-quarter notes. The parenthesis is added to relate to our context.

[162] Temperley, 2000, p. 81, referring to p. 102 of Jones, A. M. (1959). *Studies in African Music*. London: Oxford University Press. Note 14 (p. 94) adds further information: "See also Locke [...]. Agawu takes issue with Jones, arguing that the 6/8 meter is primary in such cases" (referring to p. 223 of Locke, David (1982). "Principles of Offbeat Timing and Cross-Rhythm in Southern Ewe Dance Drumming". In: *Ethnomusicology* 26.2, pp. 217–246; and pp. 189-191, 193 of Agawu, Kofi (1995). *African Rhythm*. Cambridge, UK: Cambridge University Press.

[163] cf. Yu, Getz, and Kubovy, 2015

[164] In section 2.2.1, it was mentioned that the phase of an established metric beat can be adapted more rapidly than its period when the rhythmic input is perturbed. To reflect this behavior in a model of beat induction, it is appropriate to differentiate decisions of metric reinterpretation into period and phase aspects (see Lee, 1991, pp. 77 ff.).

[165] cf. for instance Parncutt, 1994, p. 419 (see also figure A.2 (c))

[166] Under experimental conditions, this may be realized by a fade in or by a deceleration (as applied by Yu, Getz, and Kubovy, 2015). Other methods to eliminate possible effects of meter induction by starting points of cyclic rhythms are mentioned in sections 2.2.3 and 2.3.2 (cf. the experimental setups of Parncutt, 1994 and Povel and Okkerman, 1981).

provide a recent review and an extended theory, further developing the classic principles found by GARNER and colleagues:[167]

> (a) the *run principle*, according to which the clasp is perceived as the first note of the longest run, and (b) the *gap principle*, according to which the clasp is perceived as the first note following the longest gap. For example, the run principle predicts that 11100110 will be heard as 11100110 whereas the gap principle predicts that it will be heard as 11011100.[168]

The binary words indicate cycles of attacks (1) and rests (0) with isochronous temporal spacing. YU et al. call them *auditory necklaces* (ANs) and focus on bistable or ambiguous necklace patterns, where the two principles compete with each other, like in the example cited above.[169] Indeed, as discussed further in section 4.1.3, a necklace representation provides the appropriate level of abstraction to formalize this type of ambiguity. As an equivalence class of cyclic rhythms it does not define the starting point. The cycle is thus open for interpretation and reinterpretation, that is, metric phase modulation.

Metric phase ambiguity can be observed in the context of different musical styles and cultures. For instance, according to data collected by VOS et al., recorded performances of preludes by J.S. BACH evoked a variety of downbeat responses in regard to phase. "Some listeners tapped in synchrony with the notated barlines, while others tapped at the same rate but out of phase with the barlines."[170] PARNCUTT relates this finding to his own observation, displayed in figure A.2 (c), that at slow to moderate note rates, tactus phases tend to be more ambiguous than at fast rates. One possible reason is the enhanced effect of durational accents at faster rates, when some IOI get shorter than 250 ms, as mentioned above (see also section 2.3.2).

Metrical phase ambiguity relates to, and allows for metrical phase switching. In Western classic music, this effect is commonly employed at a *hypermetric* level of two measures.[171] A known example of an indirect hypermetric phase dissonance is found in HAYDN, *Symphony no. 104* (see figure 3.3). At bar 16, the end of the opening theme

[167] Garner, W.R. (1974). *The Processing of Information and Structure*. Potomac: Erlbaum; Preusser, D., W.R. Garner, and R.L. Gottwald (1970). "Perceptual organization of two-element temporal patterns as a function of their component one-element patterns". In: *The American Journal of Psychology* 83(2), pp. 151–170; Royer, F.L. and W.R. Garner (1966). "Response uncertainty and perceptual difficulty of auditory temporal patterns". In: *Perception & Psychophysics* 1, pp. 41–47; and Royer, F.L. and W.R. Garner (1970). "Perceptual organization of nine-element auditory temporal patterns". *Perception & Psychophysics* 7(2), pp. 115–120.

[168] Yu, Getz, and Kubovy, 2015, p. 2729

[169] "Garner and colleagues conjectured that if the two principles are in agreement, the clasp is stable and emerges readily, but if they disagree, the clasp is ambiguous and takes longer to emerge." (Yu, Getz, and Kubovy, 2015, p. 2730)

[170] Parncutt, 1994, p. 419, referring to Vos, P. G., R. F. Collard, and E. L. Leeuwenberg (1981). "What melody tells about meter in music". In: *Zeitschrift für Psychologie* 189, pp. 25-33. See also Fraisse, 1982, p. 173.

[171] It is rarely used at a level of three bars (cf. Temperley, 2008, pp. 305 f.).

FIGURE 3.3: HAYDN, *Symphony no. 104*, Allegro, *measures 9–19* (cf. Temperley, 2008, p. 307)

overlaps with the beginning of a new section. The two-bar hypermetric pulse established by the theme is thwarted, as the new section starts with a strong measure, initiating another two-bar hyper-measure. The new hypermetric pulse thus inverts or "flips" the strong-weak pattern including two measures.

> This situation—in which a phrase overlap coincides with a hypermetrical shift—is known as "metrical reinterpretation" and has been discussed by a number of theorists.[172]

In Western common-practice music, the metric phases of the measure level and lower metric levels are perceived as very stable due to clearly communicated metric hierarchies. Though, within musical settings which do not trigger top-down metrical schemes, bistable rhythmic structures can facilitate deliberate or spontaneous gestalt flips in the listeners' mind. This can also happen without any change in the physical signal (section 2.3.1). In sub-Saharan African styles, it is common to create musical cues for perceptual phase switches. LOCKE, for instance, outlines corresponding techniques to "creatively manipulate onbeats".[173] Performing recurring accents on metrical offbeat positions can suggest a displacement of the beat. "In displacement, beats may be shifted from their onbeat positions to the several offbeat time-points contained within

[172]Temperley, 2008, p. 307. As emphasized by Temperley, 2008, the Haydn-example is a "sudden shift *par exellence*" (ibid.) because it is not prefigured by any musical cue. Temperley, 2008 distinguishes this from gradual shifts or *hypermetrical transitions*, "involving a smoother, more incremental realignment of the musical evidence, sometimes over quite a lengthy span of music." (p. 306)
[173]Locke, 2011, p. 54

the beat's duration."[174] This can also be described in terms of syncopation, as a means to undermine the stability of an already established downbeat, and thus, to introduce phase ambiguity. As VUUST and WITEK put it, syncopations "can also be thought of as phase-shifts, where the rhythmic onset, rather than occurring in phase with its metric reference point, has a negative lag and occurs before it."[175]

The *gankogui* bell pattern in *Gahu* drum music (of the Ewe of Ghana), [3-4-4-2-3], is a 16-cycle commonly supposed to evoke a metric subcycle with four beats [4-4-4-4]. It has a particular rhythmic drive, inviting phase shifts of the beat. According to LOCKE, fife rhythmic modes are conceivable by placing the downbeat on either of the fife rhythmic attacks of the cycle. These modes correspond to shifts of the cycle downbeat to either attack. If the cycle is conceptualized as indicated above, "the second, third and fourth strokes all lie one timepoint [before] the beat. If these become accentuated, they may cause the beat to 'turn around' and rotate one timepoint counterclockwise."[176]

The above-mentioned discussion about the African conception of meter, and in particular, about the reference level, also indirectly concerns metric phase ambiguity. If there is only a "single primary pulse level"[177] and no metric hierarchy, there are also no downbeats, respectively higher levels, to give phase orientation. "Since there is only one pulse level, no hierarchy of beats and no 'beat one', this organisation cannot be described as metre."[178] Consequently, if there is no "beat one", or any higher metric framework, bottom-up processes triggered by phenomenal accents may play a more important role. In other words, grouping the elementary beats according to rhythmic figures makes the process more ambiguous in regard to where the clasp of a figure is perceived.

[174] ibid. See also pp. 50 f.

[175] Vuust and Witek, 2014, p. 6. See also the "Syncopation Shift Rule" in Temperley, 1999, stated for syncopated melodic structures in rock music: "In inferring the deep structure of a melody from the surface structure, any event may be shifted by one beat at a low metrical level." (pp. 26 and 31)

[176] Sethares, 2007, p. 60

[177] Clayton, 1997, p. 4 of the translation

[178] ibid., referring to Arom, Simha (1991). *African polyphony and polyrhythm*. Cambridge: Cambridge University Press.

Chapter 4

Formal considerations about rhythmic and metric cycles

This chapter examines formal basics and quantitative analytic systems of rhythm and meter. They are evaluated in order to provide an appropriate descriptive basis for a quantitative heuristics of metric malleability. The primary tool is a structural model of hierarchic meter (section 4.2) which formally integrates the metric types, already described and discussed in the contexts of chapter 3. It will be employed for the development and evaluation of quantitative models of metric accent (section 4.3) and of relationships between different meters (section 4.4). The later are, for their part, important components for the approach to malleability suggested in chapter 5.

Rhythmic and metric cycles can be abstracted to mathematically defined entities which can be examined with geometric or algebraic means. The formal properties of these abstractions – including transformational relations between different patterns – are isomorphic to cyclic structures occurring in other domains which can be studied independently from time-specific conditions like irreversibility (see section 4.1.1). Xenakis, for instance, has described these features in terms of rules or laws which by definition imply "recurrence in time, or symmetry in realms outside time (hors temps)."[1] In section 1.2.3, the reduction of rhythm to structural information was characterized both as a means of analysis and of perception. Common music notation, for instance, basically represents rhythmic categories rather than the actual timing of rhythmic intervals. To that effect, categorical rhythm perception involves constraints on the recognition of rhythmic structure which can be systematically analyzed (section 2.3.4). From our perspective, the subjectivity of metric interpretation results from the interaction of such automated perceptual activities of abstraction with cognitive decision processes in metric interpretation.[2]

[1] Xenakis, 1992, p. 258

[2] In section 2.3.4, it was noted that the interaction of categorical rhythm perception and meter induction has not yet been fully understood (cf. Desain and Honing, 2003, p. 363). By abstracting from specific timings – which can communicate and therefore disambiguate meter – categorical rhythm perception may also support metric ambiguity and malleability. That is, by simplifying temporal ratios, categorical perception may give rise to diverse and competing implications about meter.

How do formal and perceptual properties of rhythm and meter correlate? More specifically, it is of interest which formal properties have particular impact on metric ambiguity. It may thus be helpful to introduce a formal, analytical system of representation which provides an adequate basis to study correlations between structural features of rhythm and meter, and perceptual properties such as the metric malleability of rhythm. The psychological research reviewed in chapter 2 has identified many aspects of rhythm perception which are generalizable on an inter-individual level. However, statistic analyses do not enable predictions about individual behaviors of rhythm perception. TOUSSAINT addresses the related limitation to capture perceptual particularities of rhythms with mathematical analyses. Since "the perception of rhythm is in part intrinsically subjective and constructed in our minds, a purely mathematical solution to this characterization problem is not attainable."[3] Though, it may be sensible to study mathematical structures of rhythm and meter as a foundation to be combined with "higher-level quantitative musicological knowledge[4] [...] in order to determine the 'correlation between external stimuli and internal structures'".[5]

4.1 Analytic representation of cyclic patterns

This section primarily examines the formal backgrounds of the notations employed in this study. They represent formal properties of patterns which can be analyzed at different levels of abstraction. Some analytic tools for the classification of these properties are outlined, as particular classes of patterns may correlate, or show a coherent behavior in regard to metric ambiguity and malleability.

Since we focus on rhythmic cycles that afford pulse sensation (section 2.2), the investigations in this study are based on a common, yet qualificatory type of representation. Rhythm is rationalized or "digitized", corresponding to an isochronous succession of elementary pulses or rhythmic categories (see section 2.3.4) to which the rhythmic articulations are associated.[6] This elementary periodicity is perceived as a pulse sensation if its period range implies rhythmic movement (section 2.2.2), and if it is articulated by sufficient occurrences of rhythmic events. From this perspective, cyclic rhythms are periodic on two levels, as the cycle period emerges at a higher level comprising a group of elementary periodic units. In other words, a simple meter is established that corresponds to YESTON's definition, quoted in section 1.3. A correspondent mathematical definition of rationalized periodic rhythm by VUZA – as cited by ANDREATTA et al.[7] – states that a

[3]Toussaint, 2013, p. 309

[4]Toussaint supposes this knowledge to be generated by "the neurobiology of rhythm." (ibid.)

[5]ibid. The quotation is from Goldenberg et al. (2001). "Structures of the mind and universal music". In: *Science* 292(5526), p. 2433.

[6]This regular temporal grid can be regarded as a basic, elementary form of meter (cf. section 3.2.2).

[7]Andreatta, Agon, and Amiot, 2002, p. 157

> periodic rhythm is a periodic locally finite[8] subset R of the set Q of rational numbers, i.e.:
> 1. It exists a positive rational number t such that $t + R = R$ (periodicity)
> [...]
> The least positive rational number satisfying condition 1. is called the *period* of R whereas the greatest positive rational number dividing all differences $r_1 - r_2$ with r_i belonging to R is called *minimal division* of R.[9]

This definition of periodic rhythm as well implies at least two metric levels. First, the minimal division, dividing all temporal intervals within the rhythm corresponds to the temporal unit of the elementary pulse. Secondly, the period of R, the time interval t (positive rational number) in which the rhythm can be temporally translated to itself ($t + R = R$ represents the temporal translational symmetry, or periodicity of the rhythm) corresponds to a second metric level.

For an alternative mathematical approach, HALL and KLINGSBERG define rhythm as a function.

> A rhythm pattern can be represented as a function $f : Z \to \{0, 1\}$, where $f(x) = 1$ if there is a note onset on pulse x and $f(x) = 0$ otherwise. The function f represents a periodic rhythm of period p if $f(x) = f(x+p)$ for all $x \in Z$; thus, f can be identified with a function with domain Z/pZ or Z_p. A *rhythm cycle* is defined to be an equivalence class of p-periodic functions modulo the shift $(s \cdot f)(x) = f(x - 1)$.[10]

Here the pulse level is associated with the integer elements of Z. The function returns either the value 0 or 1 for all $x \in Z$ and x is the index of a pulse. This represents either rhythmic onsets or rests (respectively continuations), distributed over a certain number of pulses. Periodicity on a higher metric level than the pulse level only occurs if $f(x) = f(x + p)$ for all $x \in Z$ where p is the number of pulses in a rhythmic period (a rhythmic pattern which is repeated).

VUZA's rational numbers in a subset R of Q represent rhythm as a collection of quantified time points associated with rhythmic onsets. His *minimal division* is a rather abstract value for the interval between two elementary pulses.[11] Thus, the function $f : Z \to \{0, 1\}$ which returns onset values (yes or no) for every pulse index $x \in Z$ represents the categorical nature and the temporal latitude of elementary pulses more accurately.

[8]"For a, b in Q with $a < b$, the set $R \cap [a, b[$ is finite, where $[a, b[= \{x \in Q : a \leq x < b\}$." (ibid.)
[9]ibid.
[10]Hall and Klingsberg, 2004, p. 190
[11]The temporal interval between two onset times t_1 and t_2 can be calculated as their difference $t_1 - t_2$.

To relate such definitions to the real world of musical rhythms, their limitations have to be taken into account. The possible rhythms that can be specified by means like mentioned above are all one-dimensional, homophonic rhythms. To approach polyphony, for instance, more than one set R or function $f : Z \to \{0, 1\}$ have to be defined and to be "synchronized" by the numeric values or indexes. However, conceiving rhythmic patterns as periodic or cyclic facilitates the analysis of specific properties, although in practice they usually may be dynamically transformed. The variation of rhythmic patterns in music either results in free rhythmic flux or in quasi-periodic patterning of rhythms. From a composer's point of view, the properties of rhythmic patterns affect the various possibilities to rearrange and transform rhythmic structure in the course of time. As an example, consider the pattern [x..x..x..x.xx.xx.xx.xxxx]. It can be interpreted as dynamic variation of the elementary pattern [x..x..] which receives one rhythmic onset more per repetition: [x..x..] $\to$ [x..x.x] $\to$ [x.xx.x] $\to$ [x.xxxx]. This kind of variation can be also described in terms of a (reversed) *shelling* operation, as described in the following section 4.1.1.

4.1.1 Isomorphisms between rhythm, scale, and meter

Temporal intervals can be measured accurately by periodic physical motion like pendular or cyclic movements found in nature. Such reference movements should be most regular, stable and reliable in terms of their recurrence or repetition.[12] Any other movement can be set in a temporal relation to a reference period and this relation is known as the duration of the movement.[13] Pitch perception and pulse sensation also depend on periodic physical movement.[14] We are able to judge the regularity of this movement by the perceived stability of the pitch or pulse over a certain time. The frequency of a pitch or a pulse corresponds to the number of motion periods within a time span related to a reference period.[15] Musical tempo, measured in beats per minute, also relates to frequency. In section 2.2, tempo perception is characterized as referencing to a principal pulse sensation (the *tactus*). There we argued that tempo (as frequency) is only perceived if it is entrained by regular movement. It can be a reference for complex rhythmic movement like a tonic or central pitch can be for harmonic movements. Hence, both pitch and tactus (or a more complex, hierarchic meter) actually emerge from the perceptual detection of periodicity in a succession of physically articulated

[12] Astronomic and atomic movements are the current references because they seem to be the most stable and regular movements we know. Nevertheless, their periods of motion are either too long or too short to be directly perceived by our senses and to function as reference periods as such. We are more sensitive to an intermediate time scale in between these two extremes. Intervals like a second or a minute are common measures within this scale, moderating multiples of periods of atomic movements and divisions of astronomic periods.

[13] cf. Roads, 2001, chapter 1 ("Time Scales of Music")

[14] The complex relation of pulse sensation to a physical stimulus is discussed in section 2.2.

[15] Thus, frequency is reciprocal to period duration ($f = 1/T$).

temporal intervals. Pitch intervals are thus relations or ratios of different frequencies, for their part constituted by the temporal regularity of vibrations.

To establish an *isomorphism* of rhythm and pitch scale on the level of interval sets, an analogy can be drawn between the elementary pulses of a rhythm and the elementary steps of an equal-tempered pitch scale, which are logarithmic in terms of frequency. Then, any structure in both domains can be represented by an ordered set of interval magnitudes measured in elementary steps or pulses. Isomorphic sets are interchangeable between domains and thus, the pattern structure of rhythms and pitch scales can be studied and compared independently from their respective domains. The term *isomorphism* is used in several scientific disciplines with slightly different meanings.[16] It can be applied to rhythms and scales in a mathematical sense, that is, equivalent interval sets correspond to equivalent properties of the *mathematical group* which acts on them, a topic widely referenced in music theory and related domains.[17]

> One of the main contributions by mathematician Dan Vuza [...] was the precise algebraic formalization of this correspondence between pitch domain and rhythmic domain (Vuza 1985). [...] his model of periodic rhythm [...] enables the transfer of the algebraic structure of the equal tempered system into the rhythmic domain by canonically associating an intervallic (temporal) structure to a given subset of a cyclic group.[18]

The definition of "periodic" (or cyclic) rhythm, quoted in the previous section, is part of this model. If a scale or a rhythm is periodic, there exists a modulus interval of n units which is equal to the sum of the intervals in the ordered set. This structure is isomorphic to a correspondent subset of a finite cyclic group, the additive group of $\mathbb{Z}/n\mathbb{Z}$ (integers modulo n). Thus, formal properties of cyclic rhythms can be studied on a more general level. Any insight in – or even a mathematical proof of – a particular formal property established on the level of generic interval structures is applicable to rhythm.[19]

A number of comparative studies have shown that particular interval structures exert a pull across domains, as they are frequently used both as rhythms and scales. The two "standard" patterns in sub-Saharan African music, [2-2-1-2-2-2-1] and [2-2-3-2-3] were introduced in section 3.2.2, indicating that their structures are isomorphic to the diatonic, and respectively, the pentatonic scale. These patterns can be *generated*, that is, they "can be represented as $\{0, m, 2m, ..., (k-1)m\}_n$ for some integer m, where all

[16] See for instance https://en.wiktionary.org/wiki/isomorphism. Last retrieved on 12 October 2020.

[17] See for instance Andreatta, 2004, Andreatta et al., 2006, Benson, 2007, Hall and Klingsberg, 2004, and Jedrzejewski, 2006. Group-theoretical aspects are further discussed in the following and in section 4.1.3.

[18] Andreatta, 2011, pp. 37 f., referring to Vuza, Dan Tudor (1985). "Sur le rythme périodique". In: *Revue Roumaine de Linguistique-Cahiers de linguistique Théorique et Appliquée* 23/1, pp. 73–103.

[19] For instance, Xenakis, 1992, pp. 268 ff., demonstrates the application of *sieve theory* to both domains: logical concatenations of set operations lead to generic interval structures which can then be interpreted e.g. as rhythms or scales. Furthermore, the mathematical structure could also be applied geometrically to other domains like architecture.

arithmetic is modulo n."[20] This corresponds to the generation of these scales by iterative steps ($k = 7$ or $k = 5$) on the circle of fifths ($m = 7$) or fourths ($m = 5$), transposed into one octave ($n = 12$).[21] The attraction of these well-known examples is also determinated by the maximal evenness of both patterns,[22] and the relations of the number of intervals (7 and 5) to the number of elementary units (12) in the patterns: both are relatively prime. These properties are already mentioned and related to the notion of Euclidian rhythms in section 3.2.2. Note as well that relative primeness is also required between the order n of a finite cyclic group and its generator g (see also section 4.1.3), and that the generative steps of size m in the example above follow this principle as well (7 and 5 are relatively prime in relation to 12). This also implies that both patterns are *deep*, that is, they contain a unique number of each occurring interval category when regarded as cyclic, which can be shown by their *interval content histograms*.[23] The correspondent histogram of the diatonic scale is [2, 5, 4, 3, 6, 1], meaning two intervals of size 1 (minor second), fife intervals of size 2 (major second), and so on.[24] For the pentatonic scale it is [0, 3, 2, 1, 4, 0]. Deep patterns can be universally generated exactly with the method shown above, provided the relative primeness of m and n.[25] The deepness property is preserved at each step, as generated patterns "have an interesting property called *shellability*."[26]

> A rhythm admits a *shelling*, with respect to some property P if there exists a sequence of insertions or deletions of onsets so that, after each insertion or deletion, the rhythm thus obtained continues to have property P.[27]

A consideration on isomorphism in the context of this study should also include metric structures, or in other words, attentional patterns. In section 3.2.2, some general properties of metric patterns like the just mentioned maximal evenness were examined in the context of rhythm-meter ambiguity. The close relation of rhythm and meter is thus reflected on the formal level by the occurrence of isomorphic pattern structures. However, the benefit of emphasizing isomorphisms between rhythm, scale, and meter

[20] Taslakian, 2008, p. 18

[21] Iterative steps of fifths or fourths modulo the octave generate different rotations (modes) of these patterns (scales). The former results in [2-2-2-1-2-2-1] and [2-2-3-2-3], whereas the latter yields [1-2-2-1-2-2-2] and [3-2-3-2-2].

[22] The two patterns are also complementary. Several authors have mathematically proven that the complement of a maximally even pattern is also maximally even (cf. Gómez-Martín, Taslakian, and Toussaint, 2009, p. 19).

[23] cf. for instance Taslakian, 2008, pp. 16 ff. and Toussaint, 2013, pp. 178 ff. The following histograms include all interval sizes measured by the shortest distances on a cycle. According to this scheme, the pitch intervals within the chromatic scale vary between minor seconds and tritones (excluding their complementary intervals in the octave and the prime). Regarding a rhythmic cycle, every "pair of onsets determines an interonset geodesic distance (or interval)" (Taslakian, 2008, p. 140), which can sorted in a histogram by the interval size as a sum of elementary pulses.

[24] As indicated in the previous footnote, complementary intervals (of size > 6) in the cycle are not considered in regard to the deepness property.

[25] ibid.

[26] Taslakian, 2008, p. 19

[27] Toussaint, 2013, p. 188. The "terms 'insertions' (or deletions) refer to replacements of a silent (or sounded) pulse with a sounded (or silent) pulse." (ibid.)

seems to be limited from the cognitive perspective. On this level, formal equivalences are less significant than similarities in the cognitive processing of rhythm and pitch. Nevertheless, HURON points out that

> the similarity between scale-degree expectation and metric expectation is not merely metaphorical or informal. Both scale degree and metric position are perceived categorically.[28] Like scale-degree pitches, metric positions provide convenient "bins" for expected stimuli. The metric hierarchy is truly homologous to a scale or scale hierarchy. [...] Both are schematic frameworks that allow us to hear events "as" something. We hear the pitch C as the tonic. We hear a particular moment as the downbeat. In each schema, events receive mental labels that reflect their relative probability of occurrence in some commonly experienced context.[29]

While HURON focuses on the cognitive commonality that both scale degrees and metric positions emerge as categories for schematic expectations in the sequel, LONDON compares structural constraints of meter and scale for the establishment of plausible cognitive schemes.

> In borrowing concepts such as maximal evenness from mathematics and scale theory, we must be cognizant of both the similarities and differences between the pitch and time domains. [...] Metric beats are not perfect analogs to scale steps, [...] meters are subject to additional constraints that keep subdivision and beat levels distinct.[30]

Such issues are further treated in section 4.2 by means of an integrated description and a classification of hierarchic metric types. An important criterion for the reliability of a metric framework, in contrast to the operation of a scale, are distinct cues for interpolation. In the words of LONDON, a successfully established "meter keeps cycling through and articulating each beat wether or not the beats are articulated by the unfolding rhythmic surface. We readily interpolate beats into a long surface duration; we do not analogously interpolate missing tones into a melodic skip."[31]

[28]Huron, 2006 here (p. 179) quotes Clarke, 1987. See also section 2.3.4.

[29]Huron, 2006, p. 179 and p. 200

[30]London, 2012, pp. 140 f. The metric well-formedness constraints, proposed by London, 2012 and further discussed in section 4.2.1, "also draw on work in the theory of musical scales and combinatorics, although they differ from scalar and mathematical well-formedness most importantly by the metric prohibition against including N-cycle units on higher levels, such that a subdivision can never function as a beat. [...] in the case of tonal/scalar relationships a half-step that functions as the integer unit of the tonal space can have both a chromatic and a diatonic identity. Even so, maximal evenness [...] is important to both scalar and metrical contexts." (p. 193)

[31]London, 2012, p. 141

4.1.2 Notation

As categorical rhythm perception distinguishes structural from accidental and expressive timing, most musical notations similarly abstract rhythm at a symbolic level which represents the normative, structural framework for rhythmic timing.[32] Many of these systems are also configured by the periodicity of an elementary pulse, sometimes also called the *tatum*, the temporal quantum of rhythmic structure, or "the shortest possible unit of subdivision".[33] Rhythm and meter are thus represented according to a common basis: the temporal locations indicated by elementary pulses, provide a structure for the placement of metrically distributed attentional peaks, and for the temporal arrangement of rhythmic onsets.

A music may be notated according to a meter which differs from that which perceptually emerges, when it is performed. This relates to the distinction of *conceptual* versus *emergent* tempo,[34] discussed in section 2.2.2. LONDON notes that perceptually emergent "meter – the felt pattern of beats (and other levels) – remains invisible" in common music notation, and surveys "various kinds of analytic notation, that is, attempts to make meter visible", developed during a "long history" of metric analysis.[35] However, notation is suggestive about meter, but we have to distinguish "between notated and expressed meters – that is, between what is written and what we hear".[36] In other words, a notated meter may be perceptually valid or not, and this uncertainty can be regarded as an aspect of metric ambiguity. Moreover, different "notational contexts" can be "formally equivalent".[37] Figure 4.1 shows an apropos example found by GOTHAM, examining equivocal aspects of meter representation in music notation.[38] In this context, he also addresses related ambiguities in analytic notation systems for meter, such as the "dot-wise system" of LERDAHL and JACKENDOFF.[39] These issues are further discussed in section 4.3.1 in the context of theories of *metric weight*.

[32] See Sethares, 2007 for a survey and a comparison of "symbolic notations" (pp. 24 ff.) and "literal notations" (pp. 41 ff.).

[33] London, 2012, p. 201 (note 4). The term was proposed in honor of the pianist Art Tatum (ibid., referring to Bilmes J.A. (1993). "Timing is of the Essence: Perceptual and Computational Techniques for Representing, Learning and Reproducing Expressive Timing in Percussive Rhythm". MA Thesis. MIT.). See also Sethares, 2007 (cf. section 3.2.1).

[34] Benadon, 2004

[35] London, 2012, p. 77

[36] London, 2012, p. 72 (referring to Caplin, William (1981). "Theories of Harmonic-Metric Relationships from Rameau to Riemann". PhD thesis. University of Chicago). See also Vuust and Witek, 2014, p. 2: "the time signature of a given piece can often be notated in more than one way, and the subjective experience of its meter may be at odds with its formal time signature"; and Huron, 2006, p. 371: "a person who believes that subjective perception is the same as objective reality is called a naive realist. In music theory, naive realism is evident in two assumptions: that the structures we see in the notated music are the ones we experience, and that the structures we experience can be seen in the notation. Those who aim to describe the music as it is can do so only by abandoning the path of naive realism".

[37] Gotham, 2015a, p. 24

[38] ibid.

[39] cf. Lerdahl and Jackendoff, 1983

FIGURE 4.1: Gotham, 2015a, p. 25 (figure 2): a formally equivalent musical structure in two different notational contexts; "third movement of Martinu's *Les Fresques de Piero della Francesca*. Left: the opening of the movement (in 6/8); right: the recapitulation of that opening material six bars before rehearsal mark 36 (in 3/4 at almost exactly twice the tempo)." (ibid.)

Music notation draws upon "syntactic aspects of rhythm" which "can be formally described by considering rhythm to be the result of a metrical grammar".[40] In other words, "generating music notation for a piece requires identification of its meter", and "the metrical structure of a piece essentially provides the information required to rhythmically notate it."[41] If we apply this terminology to metric ambiguity, it can be described as the possible variety of syntactic interpretation on the "phonological" basis of rhythm. This notion is instructive, though, it does not distinguish serial and metric grouping (see section 2.3). Rhythms which do not induce pulse sensation can also be syntactically ambiguous, as their perceptual grouping can vary.[42] However, while such rhythms can involve grouping ambiguity, they are not metrically malleable in the sense considered in this study.

Notation and meter theory developed under mutual influence. HASTY claims that "systematic theories of meter draw upon a conceptual framework grounded in the technology of metric notation".[43] The traditional dichotomy of meter and rhythm may stem from performance issues, as metrical notation facilitates disposition and counting of musical time. Counting generates recurring temporal identities, classes of temporal "places" – "we return to or at least pass through this place even if it is not articulated with an attack".[44] This notion also implies periodicity at two levels. The temporal "places" we pass through correspond to the elementary metric categories, and if we regularly return to a place of temporal identity, periodicity on a higher level or a metric cycle is established.

[40] Desain and Honing, 2003, p. 345, referencing Longuet-Higgins H.C. (1978). "The grammar of music". In: *Interdisciplinary Science Reviews* 3, pp. 148–156 (reprinted in Longuet-Higgins H.C. (1987). Mental Processes. Cambridge, MA: MIT Press). Desain and Honing, 2003 amplify that such "a grammar describes a rhythm as accommodated by a hierarchical tree of duple or triple subdivisions. The resulting metrical tree specifies an important recurrent time interval, the *bar*, and the way it is subdivided recursively." (pp. 345 f., emphasis in source)

[41] Temperley, 2004, p. 28

[42] cf. Desain and Honing, 2003, p. 346: "metre is just one structuring factor in rhythm (cf London 2001), and unfortunately a complete grammar (or formal theory) of rhythm does not exist." (referring to London, J. (2001) Rhythm. In: The New Grove Dictionary of Music and Musicians (2nd edition). London: Macmillan.)

[43] Hasty, 1997, p. 6

[44] Hasty, 1997, p. 5

Thus, common forms of music notation basically entail particular constraints on the representation of rhythm. A second-level periodicity which groups an elementary pulse can be represented either by means of a linear or a circular notation. Both forms are used throughout this study. As an example of the latter, BENADON uses a circle to represent a "beat" which can be differentiated in terms of phases associated with different angles in the circle.[45] More general, a cyclic rhythmic structure is adequately represented by some sort of clock-face diagram. One of the oldest forms of circular notation was used by SAFÎ AL-DIN AL-URMAWÎ in the 13. century in Baghdad.[46] A circular plot facilitates instant insight into metric and geometric properties of rhythm, some of which are hard to detect by hearing or in common music notation.[47] This was outlined in section 1.1 and will be further illustrated the following section 4.1.3 by means of *necklace* notation. LONDON also uses a cyclic notation for metric types, which is as well an appropriate means to represent the aspect of continuity.[48] A similar cyclic plot of metric types is used in section 5.3 to illustrate alternative metric interpretations of rhythmic necklaces as concrete examples of metric malleability.

4.1.3 Combinatorial enumeration

> Indeed, one might say that the application of combinatorics to music shifts the creative process from the ineffable domain of invention to the more worldly mechanical realm of choice.[49]

Motivations to enumerate specific sets of rhythms – or more general, of interval patterns – are found throughout music theory and analysis, as well as musical composition.[50] The enumeration of combinatorial sets which are "subject to various constraints on their structure"[51] has a wide applicability to different musical structures which are isomorphic to each other (see section 4.1.1). Initially, abstract patterns are enumerated and then interpreted as musical entities. Constraints on the mathematical structure of these patterns can be related to specific properties of rhythm, for instance rhythmic "asymmetry" or "oddity".[52] Hence, constraint enumeration leads to defined sets of rhythm or specified *rhythmic spaces*.

[45]cf. Benadon, 2007

[46]cf. Sethares, 2007, p. 28 (referring to Safî al-Din al-Urmawî, *Kitâb al-Adwâr* (1252), trans. by Erlanger R. (1938). In: *La Musique arabe*. Paris: Paul Geuthner), and Toussaint, 2013, p. 33

[47]cf. Toussaint, 2013

[48]London, 2012, pp. 83 ff.

[49]Toussaint, 2013, p. 277

[50]cf. for instance Benson, 2007 (chapter 9), Chemillier, 2004, Fripertinger, 1991, Hook, 2007, Jedrzejewski, 2014

[51]Hook, 2007 (abstract)

[52]These notions are related in the approaches of Chemillier and Truchet, 2003, Chemillier, 2004, and Hall and Klingsberg, 2004. Rhythmic *oddity*, already mentioned in section 3.2.2 (see footnote 117), is further discussed at the end of this section.

"Combinatorial word theory studies strings of symbols, usually drawn from a finite alphabet".[53] Strings can as well represent rhythmic pattern structure, as they are patterns themselves. Strings are classified by their length n and the size of their alphabet k. They can refer to a particular interval pattern in the following ways, for instance

2 3 4 1 3 $(n = 5, k = 4)$, or
1 0 1 0 0 1 0 0 0 1 1 0 0 $(n = 13, k = 2)$

The former can represent a set of intervals as integer magnitudes, as in the interval-based notation [2-3-4-1-3], employed in this study. The latter is a binary string $(k = 2)$ and corresponds to box notation, as "1" can be interpreted as "onset" whereas "0" stands for "rest" or "continuation": [x.x..x...xx..]. On the level of strings we encounter some important possibilities of generic representation of rhythmic cycles. Types of strings which are of particular interest, are *necklaces*, *bracelets* and *Lyndon words* with a binary alphabet $(k = 2)$.

Necklaces in mathematics and computer science are strings which are invariant under rotation. They can represent interval cycles without differentiating a phase or starting point. Thus, all rotations of a circular rhythmic pattern are represented in one necklace. In other words, they are *equivalence classes* under rotation,[54] for instance, the strings abc, cab and bca are all rotations of the same necklace. It is common to plot necklaces with "beads" in different colors. As an example, all necklaces with $n = 3$, $k = 2$ (binary) are displayed in figure 4.2. The correspondent strings underneath are read clockwise from the necklaces, starting at "twelve o'clock".

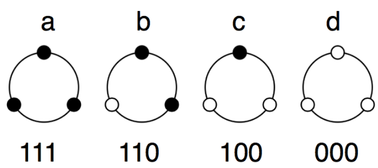


FIGURE 4.2: All necklaces and correspondent strings with $n = 3$, $k = 2$

To apply the structure of a binary necklace to the rhythmic domain it can be read like the function by HALL and KLINGSBERG, introduced in section 4.1. Then, a black pearl represents the value 1 or a rhythmic onset and a white pearl is associated with the value 0, a rest or continuation in the sense that there is no rhythmic impulse. Each pearl thus represents an elementary pulse on the minimal division level of the rhythmic cycle. In the mentioned function, p is the number of pulses in a rhythmic period. Then p

[53] Clampitt and Noll, 2011, section [3]. "Word theory and mathematical music theory have proceeded in almost total ignorance of each other [...]. Notable exceptions include the rhythmic studies of Chemillier and Truchet, 2003 and Chemillier, 2004." (ibid.) In contrast, Clampitt and Noll, 2011 "explore transdisciplinary mappings between word theory and scale theory. It is not merely a matter of translation of existing music-theoretical conceptions into a new terminology (although that in itself is non-trivial), but of mapping word-theoretical constructions into music theory to create new music-theoretical understandings." (section [4])

[54] cf. Hall and Klingsberg, 2004

equals n in terms of necklaces because a necklace represents the periodicity through its cyclic form. A binary necklace also corresponds to a rhythmic cycle expressed in box notation, if the cycle is thought of as invariant under rotation. Necklaces initially represent rhythmic cycles independently from temporal or metric context. Figure 4.3 illustrates that the necklace representation can be flexibly related to metric duration.

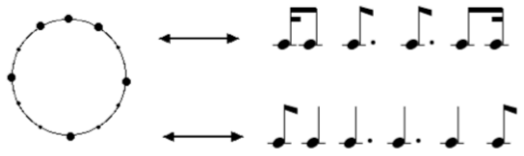


FIGURE 4.3: Flexible "mapping from the circular representation to the rhythmic domain" (quotation and figure from Andreatta, 2011, p. 37).

The possible rotations of a necklace are isomorphic to a finite cyclic group acting on the necklace. Let us assume, a rotation by one unit ($+1$) is a left-to-right rotation of a string ($abc \rightarrow cab$) by one element or a clockwise rotation of a necklace around its center by $360/n$ degrees. The group action or law ($\bullet$) is then the concatenation of rotations and n rotations by d units generate the cyclic group Z_n of all rotations of a necklace with n beads. In the case of the necklaces in figure 4.2, d (the group generator) can be 1, 2, -1 or -2. The rotations can be associated with temporal translation or phase shift of the rhythm in regard to a metric cycle by a certain number of elementary pulses. In the following this is called *metric rotation* or *metric phase shift*. For instance, the necklaces in Figure 4.2 represent all cycles of three metric units or pulses. To enumerate all metric rotations of $n = 3$ we can list all correspondent rotations of binary strings which are different from each other. In sum there are seven metric rotations with rhythmic impulses and one without, a metric rest (Table 4.1).

TABLE 4.1: Possible necklace rotations

Necklace (figure 4.2)	Strings
a	111
b	110, 011, 101
c	100, 010, 001
d	000

This set corresponds to the combinatorial set $n = 3$, $k = 2$ but it is ordered in a way we can make musical sense of, by realizing that a metric rotation of a rhythm consists of a rhythmic figure and its phase in regard to its metric ground.[55] In a metric N cycle (section 3.2) of n elementary pulses, the same rhythmic cycle can occur in n different phases. Thus, if we project the rhythmic motif represented by necklace b in figure 4.2

[55]For a discussion of the figure-ground relation of rhythm and meter see section 3.2.2.

to a metric 6 cycle we obtain six different metric rotations (Figure 4.4). This holds for all cases where the rhythmic necklace is *aperiodic*, as discussed later.

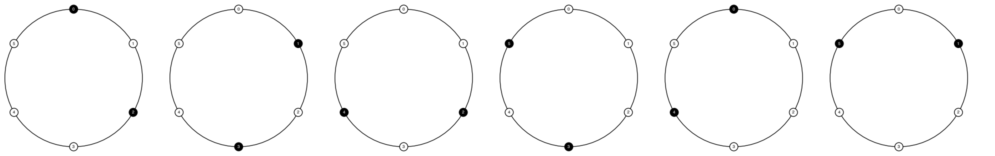


FIGURE 4.4: Projection of necklace b in figure 4.2 to a metric 6 cycle, yielding six metric rotations.

The rhythmic cycle represented by necklace b is also characterized by an IOI ratio of 2/1 and can be written as [1-2] or [2-1] in interval notation. The projection of this simple rhythm into a more differentiated metric context can be associated with a more hierarchic metric interpretation. A metric 6 cycle suggests a third metric period on a level between the elementary pulse and the cycle period which results either in a 3 x 2 or in a 2 x 3 grouping of the six elementary pulses, corresponding to 3/4 and 6/8 meters with elementary pulses of 1/8. The [1-2] or [2-1] pattern can thus be interpreted as metrically straight (in 3/4) or syncopated (in 6/8).

Two variants of necklaces, *bracelets* and *Lyndon words*, reveal additional relational structure which can be verified for rhythmic cycles by their enumeration. A *bracelet* is a string which is invariant under rotation and reflection (inversion). Hence, a set of necklaces contains a subset of bracelets, and the possible rotations and reflections of a bracelet with n beads are equivalent to the dihedral group D_n acting on the bracelet.[56] To illustrate this by means of an interesting case of a constrained set, all binary necklaces of $n = 7$ with a *fixed content*[57] of four "black-bead" characters and three "white-bead" characters are listed in figure 4.5. All correspondent rhythmic cycles would thus contain seven elementary time units and four rhythmic onsets. Three of the necklaces exhibit mirror-symmetry (see the grey symmetry axes). Therefore, the same transformation can be represented by either a particular rotation or a particular reflection.[58]

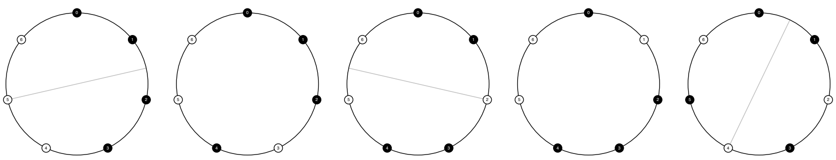


FIGURE 4.5: All necklaces $n = 7$, $k = 2$ with fixed content $n_0 = 3$, $n_1 = 4$

[56] Jedrzejewski and Johnson, 2013, for instance, "consider set classes up to transposition and inversion, that is under the action of the dihedral group." (p. 128)

[57] cf. Karim et al., 2013, p. 103: "If the number of occurrences of each character i is given by n_i where $n_0 + n_1 + \cdots + n_{k-1} = n$, then such strings are said to have *fixed content*."

[58] In group-theoretical terminology, these necklaces represent equivalent *orbits* under the dihedral and the cyclic group.

The other two necklaces contain no symmetry but are *chiral*.[59] This basically means that they do not possess inherent mirror-symmetry, but are mirror-symmetric to each other, like for instance left and right hand.[60] They thus belong to the same bracelet. The constrained set of strings $n = 7$, $k = 2$ with fixed content $n_0 = 3$, $n_1 = 4$ hence contains five necklaces and four bracelets.

The major benefit of necklaces and bracelets concerns the integral representation of properties which apply to all rhythmic cycles belonging to a necklace or a bracelet. Many features are thus related to these equivalence classes. As shown in figure 4.4, the metric rotations of a rhythmic cycle result in different musical rhythms due to their different accentual structures. However, “the interval content of a rhythm is invariant to its rotations.”[61] The same holds for reflections or inversions of rhythmic cycles. Thus, the discussion about interval histograms in section 4.1.1 relates to properties of bracelets. Other features like evenness and deepness, mentioned there and in section 3.2.2, are also properties of bracelets. These are not just abstract attributes, they also have perceptual and musical relevance. TOUSSAINT evaluates such necklaces (or bracelets) which occur as different metric rotations (or reflections) of rhythmic cycles within musical practice.

> One way to measure the robustness of the effectiveness of a necklace as a template for the design of rhythm timelines is by the number of its rotations that are actually used in practice. If many rotations are used, it suggests that the effectiveness of the rhythms it generates does not depend crucially on the starting onset, even though the result may sound quite different. A rhythm necklace that has the property that all its onset rotations are used as timelines in practice will be called a *robust* rhythm necklace.[62]

“Onset rotations” are metric rotations with an onset “at pulse zero”,[63] that is, the cycle downbeat, clasp (section 3.2.5), or metric starting point of the cycle. Metric onset rotations may be more common, though, *anacrustic* rhythms with a *loud rest* at the cycle downbeat are easily found in practice as instances of metric *offset rotations* of rhythm necklaces. To illustrate the musical robustness of particular rhythm necklaces, TOUSSAINT lists examples of both. Different onset rotations of [3-3-2], [2-1-2-1-2], and [2-2-2-3], as well as the two “standard” patterns discussed in sections 3.2.2 and 4.1.1, occur in many musical cultures all over the globe. Some of them are also found as offset rotations, for instance [2-1-2-1-2] occurs as [.x.xx.xx] in a Rumanian dance, and the standard pattern [2-2-1-2-2-2-1] is employed in Cuban *rumba* as [.x.x.xx.x.xx] (*palitos* rhythm).[64]

[59]cf. Petitjean, 2010

[60]“The word *chirality* is derived from the Greek $\chi\epsilon\iota\rho$ (kheir), ‘hand,’ a familiar chiral object.” (https://en.wikipedia.org/wiki/Chirality. Last retrieved on 12 October 2020.)

[61]Toussaint, 2013, p. 73

[62]Toussaint, 2013, pp. 74 f.

[63]Toussaint, 2013, p. 81

[64]Toussaint, 2013, pp. 75 ff.

Necklaces and bracelets are of crucial importance for the conceptualization of metric malleability. As it was outlined by the example pattern [1-2-3] in section 1.1 (see figures 1.1 and 1.2), metric rotation is one of the aspects of malleability, which corresponds to phase ambiguity and type B dissonance (sections 2.3.3 and 3.2.5). The aspect of meter and periodicity of rhythm can additionally be reflected at the level of strings by *Lyndon words*. "A Lyndon word is an aperiodic notation for representing a necklace."[65] More precisely, it "is an aperiodic word that is the lexicographically smallest of its rotations."[66] Constrained enumerations of Lyndon words thus yield more compact listings of *prime* (aperiodic) necklaces, which can then be repeated to generate periodic structures like ...101010... or ...100100... and so forth. If we enumerate the lexicographically smallest rotations of the necklaces in figure 4.5, we see that they all correspond to Lyndon words, because they are all aperiodic: 0001111, 0011101, 0011011, 0010111, 0101011. But as soon as constrained enumeration enables periodic structures, the differences become obvious. The set of necklaces $n = 4$, $k = 2$ is {0000, 0001, 0011, 0101, 0111, 1111}, but the correspondent set of Lyndon words is only {0001, 0011, 0111}. The other necklaces are repetitions of the more elementary Lyndon words {0, 01, 1}.

Periodicity in rhythmic structures generally supports the perception of metric structure. *Metric rhythms* were discussed in section 3.2.2 in terms of *rhythm-meter ambiguity*. The formal differentiation of necklaces, bracelets and Lyndon words provides an analytic tool to classify rhythmic equivalence classes under rotation, for instance, in terms of features relating to periodicity and symmetry. For instance, CHEMILLIER and TRUCHET have proposed a computational method involving Lyndon words, to enumerate words (or strings) which represent *rhythmic oddity*, a property of a particular class of asymmetric rhythmic patterns already mentioned in section 3.2.2.[67]

> A rhythm with an *even* number of pulses in its cycle has this property if no two of its onsets divide the rhythmic cycle into two half-cycles, that is, two segments of equal duration.[68]

This observation goes back to AROM, who studied the rhythmic timelines of the Aka Pygmies of Central Africa. These cycles "have the additional restriction that all note onsets are spaced by 2 or 3 units, and that the period is $4n$, thus ensuring that the rhythm splits into two patterns of length $2n - 1$ and $2n + 1$."[69] Well-known examples

[65] http://mathworld.wolfram.com/LyndonWord.html. Last retrieved on 12 October 2020.

[66] https://github.com/dvberkel/debruijn/wiki/Lyndon-Words. Last retrieved on 12 October 2020.

[67] Chemillier and Truchet, 2003 and Chemillier, 2004 employ Lyndon words to enumerate equivalence classes (up to a cyclic shift) of strings, that is, necklaces which feature rhythmic oddity. Hall and Klingsberg, 2004 use their rhythm function discussed in section 4.1, and *Burnside's Lemma* to solve the same problem, and generalize rhythmic oddity for all other integer fractions of the N cycle in addition to the half-cycle. Their method also allows for the construction of *tiling canons* (cf. for instance Andreatta, 2011 and Andreatta, Agon, and Amiot, 2002).

[68] Toussaint, 2013, p. 85

[69] Hall and Klingsberg, 2004, p. 194. These features correspond to the *pseudo-aksak* rhythms discussed in section 3.2.2.

featuring rhythmic oddity are the 24 cycles [3-2-2-2-2-3-2-2-2-2-2] and [3-3-3-2-3-3-2-3-2], as well as the 12-pulse standard pattern [2-2-3-2-3].[70] These patterns exhibit mixed beat durations in between the periodic levels of the elementary pulse and the cycle, and do not articulate isochronous subcycles. Nonetheless, the first and the last pattern also metrically well-formed in the sense of LONDON as they are maximally even (see section 3.2.2). Consequently, periodicity and "metricity" relate to each other in a complex way. The coincidence of rhythm with metric structure may rise with periodicity, but the examples above and in section 3.2.1 show that this is not generalizable in the context of mixed meters.

4.2 An integrated formal description of meter

In chapter 2, meter was characterized as a cognitive framework which correlates a perceived rhythm with an attentional pattern. Conversely, rhythm is the vehicle by which meter is embodied.[71] The descriptive system proposed in the present chapter applies to this sensuous and psychological basis. It is nevertheless thought of as an ideal projection of metric space. Hence, it contributes to a line of research integrating meter both as a formal structure and a cognitive device. The specifications about meter made so far will now be translated into a computational account of meter which will serve as a background for specifying the quantitative model of metric malleability. It integrates different types of meter into a common numeric scheme.

The models of LEHRDAHL and JACKENDOFF, and LONDON define basic cognitive and formal constraints of meter by means of metric *well-formedness* rules and metric *preference* rules.[72] The logical framework of well-formedness allows to construe plausible forms of meter which satisfy both types of constraints. In addition, the metric preference rules by LEHRDAHL and JACKENDOFF give estimates about the likelihood of which specific meter will perceptually emerge among several well-formed candidates.

A first sketch of the approach proposed in this section was published in 2014.[73] Shortly after, GOTHAM rendered a similar system of description, focusing on "'mixed' metrical structures that systematically use more than one type of duration on the same level, such as both 'short' and 'long' beats (which may or may not correspond to 2- and 3-unit groupings)."[74] GOTHAM's illuminating analysis and classification of formal relations between different metric types turned out to be helpful for improving my system. To update my approach in this light, the crucial features of the mentioned models will be

[70] see footnote 117 in section 3.2.2.
[71] cf. for instance Hasty, 1997, Petersen, 2010
[72] Lerdahl and Jackendoff, 1983 establish both types of rules. London, 2012 proposes a set of metric *well-formedness constraints* (pp. 91 ff.) and discusses these by means of analyses of musical instances.
[73] Härpfer, 2014
[74] Gotham, 2015b, see the *abstract* and the corresponding footnote, advocating the term "mixed meter" to be used in this sense despite possible confusion.

discussed along with the following account, and where appropriate, incorporated in the proposed quantitative framework.

In sections 1.3 and 3.2.1, metric forms employed in practice are preliminarily distinguished into two types. First, metric hierarchies arising from nested, isochronous pulse sensations are thought of as multiplicative, divisive or simple. They are prevalent in Western musical traditions. Second, if non-isochronous pulses emerge on a metric level below the measure,[75] corresponding meters are called additive, non-isochronous, or mixed. Such structures are prevalent in many musical traditions, for instance in Balkan *aksak* styles or in Hindustani *taals*. Both variants of meter are generally found all over the globe.

The complex interaction between rhythmic "surface" and metric interpretation is distinguished into several aspects in chapter 2. These factors concern temporal properties and relations within the rhythmic structure (rhythmic interval structure and event rate), cognitive bottom-up processes (phenomenal accents), and top-down influences (metric priming and learned schemes). In the following, the formal properties and constraints of possible attentional patterns, to which a rhythm can be matched by the process of metric interpretation, are subject to a generic description and a formal analysis. Both capture essential structural properties of metric types which are represented in terms of hierarchical numeric structures, integrating simple and mixed forms of meter.

4.2.1 Metric levels and metric layers

We start the examination of metric types on the basis of the general notion of meter as a cyclic sequence of metric periods. The mentioned rules of metric well-formedness primarily concern formal constraints on hierarchic groupings of these periods. A simple well-formed metric ordering is represented in figure 4.6 by means of isochronous rhythmic strata.

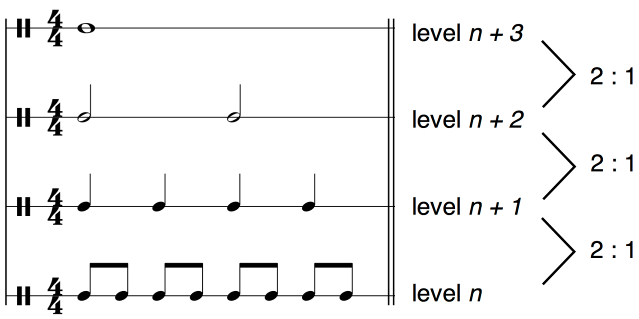


FIGURE 4.6: Metric levels and hierarchic grouping structure (see Härpfer, 2014, p. 1024)

[75]Non-isochronous structures on the measure level, that is, changes of time signatures (involving different pulse- or beat cardinalities), and on higher (hypermetric) levels are more common in Western music.

Uniform successions of metric periods constitute metric levels which are superposed and nested according to fixed integral ratios (2:1 between all adjacent levels in the case displayed in figure 4.6).[76] This formal framework thus represents meter as an ordered set of levels relating to each other by integral ratios. A theoretically infinite metric hierarchy is in fact limited by the temporal window for meter (see section 2.2.2). Thus, level 0 denotes the elementary pulse level which cannot be further subdivided into metrically functioning units. The duration of the metric cycle is conversely limited by the cognitive integration period, that is, the psychological present.[77] Metric grouping – as opposed to serial grouping (see section 2.3) – in the sense of LERDAHL and JACKENDOFF follows from the temporal coincidence of nested levels.[78] The metric grouping structure of a level is thus determined by the upper levels, the periods of which entail the groupings of lower periods. LERDAHL and JACKENDOFF as well as LONDON demand temporal coincidence of nested levels as it is implied in some of their metric well-formedness constraints.

> **MWFR 2** (**revised**) Every beat at a given level must also be a beat at all smaller levels present at that point in the piece.[79]

> WFC 3.3: The N cycle and all subcycles must begin and end at the same temporal location; that is, they must all be in phase.[80]
> WFC 3.4: Each subcycle must connect nonadjacent time points on the next lowest cycle.[81]

LONDON's N cycle corresponds to the elementary pulse level 0 which is supposed to be the lowest, categorically distinct metric level. The subcycles correspond to higher metric levels which hierarchically group elementary pulses. These constraints also allow for non-isochronous pulses or groupings which would for now contradict the hierarchical framework, outlined in figure 4.6. As discussed in section 3.2.1, we can hear a pattern like [3-3-2] either as syncopated in an isochronous framework or as a metric pattern with mixed beat lengths. Hence, if we contextualize this pattern, assuming for instance that the framework above was established before by metric entrainment to previous rhythms, it may be perceived as a syncopated variant of [4-2-2] (see figure 4.7). In any case, the second attack disarranges the framework and makes it ambiguous.

If we assume the other hearing, facilitated for instance by former entrainment to the subdivision level n, the accurate notation is shown in figure 4.8. The time signature

[76] The notion of meter as a multiple stratification of metric levels corresponds to the models of meter proposed by Lerdahl and Jackendoff, 1983, London, 2012, Yeston, 1976, and others.

[77] cf. London, 2012, well-formedness constraint 1.4: "The maximum duration for any or all [metric] cycles is ≈ 5,000 ms" (p. 92). See also sections 2.2.2 and 2.2.3.

[78] See also the notion of Yeston, 1976, mentioned in section 1.3 (cf. footnote 85), characterizing meter as an interaction of nested metric strata.

[79] Lerdahl and Jackendoff, 1983, p. 72

[80] London, 2012, p. 92

[81] ibid.

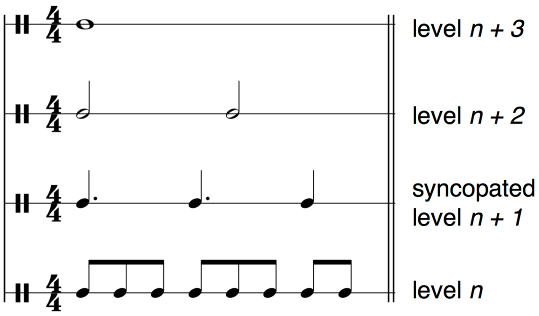


FIGURE 4.7: Rhythmic contradiction of metric grouping (see Härpfer, 2014, p. 1025)

here reflects the non-isochronous grouping structure. In the following, non-isochronous strata in a metric hierarchy will be called metric *layers* to distinguish them from metric *levels*, the periods of which are supposed to be isochronous by definition. In contrast, the metric periods of layers are regularly changing.

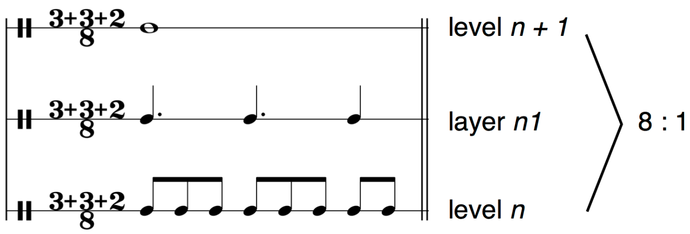


FIGURE 4.8: Layer structure and time signature (see Härpfer, 2014, p. 1025)

Figure 4.8 displays three metric strata.[82] Layer $n1$ imposes an "idiomatic" grouping structure [3-3-2] on level n. Consequently, level n stands in an eight-to-one relation with its next adjacent level $n + 1$. Due to this terminology, "the ratio of two levels is always fixed but there is no fixed ratio between a level and a layer."[83]

As stated in section 2.2.3, the tactus (or main beat) is the most salient pulse, felt within a metric hierarchy. According to the present distinction, the tactus can coincide either with a level or with a layer, and in the latter case, the tactus consists of non-isochronous beats. The notation in figure 4.8 suggests such a non-isochronous tactus, because the eight-note level is the counting level. This matches the assumption that the subdivision level has to be cognitively present in order to distinguish the two different beat types, corresponding to simple and mixed meter respectively (see section 3.2.1). It also implies that the pulse rate of layer $n1$ is moderate, that is, its average period lies around the center of the *existence region of pulse sensation* discussed in section 2.2.3. This is both

[82] The more generic term *stratum* is employed in this study as the superordinate concept of *level* and *layer* (cf. Yeston, 1976)

[83] Härpfer, 2014, p. 1025

important for a proper temporal discrimination of the subdivision level, and to allow for perceptual salience of the tactus.

Most formal studies of rhythm assume categorical equivalence (nominal isochrony) between elementary-pulse intervals. This conflicts with some evidence concerning Balkan (aksak) and West African (timeline) practices. In performances of some aksak styles musicians apparently do not feel subdivisions of non-isochronous beats.[84] Also, categorical distinctions of non-isochronous subdivisions of isochronous beats, occurring for instance in particular African music, contrast the general assumption of an isochronous N cycle.[85] Consequently, GOTHAM relaxes the demand for isochrony of the "common fast pulse (CFP) [...] to include any other system of two beat-types (such as a 'Long' versus 'Short' pairing) that is excluded by the CFP."[86] To summarize, non-isochronous beats may be categorized by (acoustic or virtual) isochronous subdivisions or by internalized temporal intervals, the durations of which stand in a complex relation. Usually, only two beat-types are distinguished.[87] Complex NI categories can also be found on the subdivision level. For the purpose of this study, rhythm is however defined on a grid of nominally isochronous elementary pulses. Hence, the metric forms discussed in the following, are only based on such elementary pulses and all other mentioned cases will be kept aside.

On this basis it will be assumed that metrical groupings contain either two or three elements. This is suggested as a general principle by perceptual and musical experience. Metrical patterns tend to perceptually decompose into sub-patterns when they comprise more than three elements.[88] This is also supported from the perspective of metric

[84] As already quoted in section 2.2.1, Baraldi, Bigand, and Pozzo, 2015 suggest that, at least in some aksak practices, "musicians conceive two 'blocks' of durations (S and L), which are largely independent one from the other, that is, they do not rely on a common underlying pulse" (p. 275, see footnote 98 in section 2.2.1 for a more extensive quotation). Timeline-oriented African performance timing was for instance explored by Magill and Pressing, 1997, who found that their experimental data suggest a mental model based on an asymmetric *timeline-ground* (TLG), rather than a pulse-based representation.

[85] Polak and London, 2014 have explored specific examples of Mande drumming from Mali, asserting "discrete categories of beat subdivision (Long vs. Short) as well as evidence of expressive variations within each category." Thus, instead of different beat categories, non-isochronous metric categories on the sub-beat level are found. Furthermore, Polak, 2010 studied jembe (drum ensemble) music from West Africa in 4-beat meter, involving ternary beat subdivision (12-pulse cycle). For a particular piece, he showed "that the ternary beat subdivision forms a repeated sequence of unequal (short, flexible, and long) subpulses. This stable rhythmic feel pattern, SFL, is unmistakable and non-interchangeable with a second ternary pattern, which is characterized by long, flexible, flexible subpulses (LFF) and occurs in other pieces of jembe music." Referring to the *Many Meters Hypothesis* in London, 2012 (see section 3.1.2), he claims that "these timing patterns distinguish individual meters."

[86] Gotham, 2015b, section 2.3. The notion of CFP is equivalent to that of the elementary pulse e, used in this thesis. Gotham, 2015b utilizes the instructive term *beat-type* "to avoid the term 'beat class', which is used in the music theory literature as a rhythmic analogy to 'pitch class', which denotes a metrical position" (footnote 11).

[87] London, 2012, pp. 132 ff., provides a complex discussion of perceptual and combinatoric limits for the distinction of different beat types, asserting that one "rarely finds metric patterns with more than two specific beat classes." (p. 135)

[88] cf. for instance Fraisse, 1982, p. 168, and Krebs, 1987, p. 120 (note 15): "Levels whose cardinalities are multiples of two are normally heard as breaking down into faster duple levels, even if no division of this kind is explicitly presented in the music."

projection. HASTY asserts that the “realization” of higher-order metrical “projections” (metric groupings) cannot be deferred “to inequalities of more than three beats.”[89] Subjective grouping (depending on event rates, see section 2.3.1) is another phenomenal argument for the decomposition of metric projection, that is, projected metric periods fall apart into smaller groups, and an intermediate metric stratum is established.[90] LERDAHL and JACKENDOFF consequently include a correspondent rule in their metric well-formedness criteria, which is based on those phenomena.

> **MWFR 3** At each metrical level, strong beats are spaced either two or three beats apart.[91]

In contrast to the other MWFRs, this rule is not extensively discussed or further refined in the course of contradicting arguments. LERDAHL and JACKENDOFF only mention the problem of quintuplet subdivision, which is characterized as belonging to “extra-metrical” categories as grace notes, trills, and the like.[92] Though, MWFR 3 is classified as idiom-specific, taking into account musical forms exhibiting “a much more loosely structured metrical idiom such as *recitativo*”.[93] However, LONDON regards this rule as a “universal aspect of metric well-formedness; this constraint helps to preserve the stability of each hierarchic level.”[94] More specifically, it “prevents adjacent accented beats”[95] and “excludes the use of a single-element span as a metrical duration on more than one structural level.”[96]

In mixed meters, metric grouping “into twos and/or threes [...] coexist within the same level”.[97] Perceptual decomposition similarly takes place in mixed metric contexts, for instance, “if we are given a group of five (at a moderate tempo) we will hear this as a composite of duple and triple groups.”[98] This implies a fundamental equivocality arising with mixed metric types. Due to the commutability of [3-2] and [2-3] groupings there are two perceptual variants for the subdivision of a metric level with five elements. In practice, other contextual aspects may guide a particular metric projection.

While discussing perceptual constraints and musical application of mixed meter, LONDON asserts a prevalence of 3:2 ratios between non-isochronous beats.[99] Several perceptual and combinatoric determinants have to be taken into account here. Part of the explanation is that the durations of non-isochronous beats have to be similar enough to represent the same metric level, but distinct enough to be categorically differentiated.

[89] Hasty, 1997, p. 140
[90] Hasty, 1997, pp. 131 ff.
[91] Lerdahl and Jackendoff, 1983, p. 69
[92] Lerdahl and Jackendoff, 1983, p. 72
[93] Lerdahl and Jackendoff, 1983, p. 97
[94] London, 2012, p. 91
[95] ibid.
[96] Gotham, 2015b, section 2.3, note 10. Gotham, 2015b nevertheless mentions some contradictory examples of Western music notation such as in Stravinsky’s *Dumbarton Oaks*.
[97] ibid. (section 2.3)
[98] Hasty, 1997, p. 132
[99] London, 2012, pp. 132 ff.

In other words, they should neither be perceptually augmented in the direction of 2:1, nor be assimilated as 1:1 (see section 2.3.4). The other crucial aspect is the tendency for metrical decomposition of more complex pulse ratios to simpler ones. This essentially corresponds to the general restriction for metrical groupings including two or three elements, outlined above and formally expressed by MWFR 3. First, it prevents from metrical inconsistencies, as for instance, a rhythmic pattern like [2-3-4] can be casted in the non-isochronous metric pattern [2-3-2-2], interpolating a beat within the last rhythmic interval. Secondly, decomposition of complex ratios leads to intervening levels within certain temporal constraints. At a fast rate, adjacent-integer ratios of 3:4, 4:5, or higher may be assimilated as nominally isochronous. A 4:5 ratio

> may be apparent only if every subdivision is phenomenally present in the musical texture. And if the tempo is slow enough to make the 4:5 ratio salient, for example where IOIs on the N cycle are ≥ 150 ms, the listener is [...] apt to interpolate an intervening level of metric articulations. creating a 2-2-2-3 or 2-2-3-2 pattern, decomposing the fours and fives into twos and threes.[100]

The limits of beat subdivision and beat duration provide additional evidence for the prevalence of 3:2 ratios between non-isochronous beats, as the absolute temporal intervals affect the exemplified perceptual tendencies.[101] Accordingly, the assumption that metrical groupings contain either two or three elements could be applied to mixed meters as well. Though, some pitfalls have to be taken into account. For example, at fast rates of elementary pulses, the subdivision level can form an intermediate metric level between the elementary and the beat level.[102] A repeating rhythmic pattern based on seven very short elementary pulses like [2-2-2-1] renders some uncertainty about how to construe its metric hierarchy. We may feel the beat level to integrate each two attacks, thus, it would comprise two beats in 4:3 ratio. But in this case, a subdivision level would be hierarchically ambiguous. If [4-3] will perceptually decompose into [2-2-3], the last interval will be both part of the beat level and the subdivision level, violating LONDON's WFC 3.4 (see above). If the subdivision level is construed parallel to the rhythmic pattern [2-2-2-1], the last interval will be both part of the subdivision level and the elementary level, violating again WFC 3.4. Hence, a subdivision level cannot be established and the pattern would be either have three beats [2-2-3] or two [4-3]. In terms of evenness one may favor [4-3] as the beat layer, though, as discussed in section 3.2.2, LONDON's WFC 4.2.2[103] about maximal evenness in non-isochronous metric patterns allows some flexibility, and thus, the three-beat interpretation [2-2-3]

[100] London, 2012, p. 134
[101] In this context, London, 2012 proposes the sextuplet as the limit of beat-subdivision and 1200 ms as the upper limit for the period the beat level.
[102] An isochronous sequence with IOIs about 150 ms may evoke a subjective tactus at 600 ms and a faster pulse sensation at the double rate (300 ms). See sections 2.2.3 and 2.3.1.
[103] London, 2012, p. 129

would be plausible at all but the fastest rates. The same holds from the perspective of LERDAHL and JACKENDOFF's MWFR 3.

Hence, the assumption made above, corresponding to MWFR 3 and to the descriptive system of GOTHAM (see above), may be generally valid. Exceptions may occur at particular temporal conditions like extremely fast elementary pulses. The other specific metric formations mentioned above, the lack of isochrony on sub-beat levels and non-isochronous beats, characterized as metrical atoms without a subdivision structure, need an adapted theory.[104] The latter is usually described by means of beat-type categories like S (short) and L (long). As already mentioned in footnote 87, metric patterns with more than two beat types, for instance S (short), M (medium), and L (long), are rare. Moreover, "establishing a meter with three specific beat classes is a perceptually difficult task, because various constraints lead one to reduce the three beat classes to two."[105] These cases may thus be handled in parallel by first assuming that metrical elements can be S or L, and secondly inferring that hierarchic groups contain either two or three of these elements. GOTHAM describes tactus levels of mixed meters "by an ordered succession of duple and triple beat groupings [...] 22223 therefore refers to a 5-beat, 11-pulse meter."[106] Additionally, he introduces a generic notation by defining a *meter vector*, representing the beat level as an unordered set of beats. It just contains two dimensions "to designate the number of duple and triple beats respectively. For instance, (22223) would be represented as {4,1} as it has four duple beats, and one triple."[107] In both notations, the twos and threes may be replaced by S and L elements.

4.2.2 Classification of hierarchic meters

In the following, the combinatorial possibilities of hierarchically organized metric cycles up to 24 elementary pulses will be discussed. The limit of 24 elementary pulses is reasonable in terms of the temporal limits for meter, discussed in section 2.2.2. If the "metric floor",[108] the lower temporal limit of metrically meaningful periods is assumed around 100 ms, a full cycle would then have a length around 2400 ms. The upper temporal limit of 3 to 5 s for periods perceived as metrical would be reached soon when the elementary pulses are between 100 and 200 ms.

Recursive groupings in longer and more complex meters may become phenomenally or perceptually salient by additional pulse sensations at slower rates. Higher metric strata may be rhythmically embodied by accentual structures, facilitating or enhancing entrainment at slow pulse rates which offer little pulse-period salience (see

[104] See footnotes 84 and 85.
[105] London, 2012, p. 135
[106] Gotham, 2015b, section 2.5
[107] Gotham, 2015b, section 2.6. Simple meters are represented by vectors in which one of the two dimensions takes the value 0.
[108] London, 2012, p. 134

section 2.2.3). Expanding the example above, GOTHAM uses nested brackets "to differentiate between different higher-level groupings of the same beat succession, such as ((2,2)(2,2,3)) and ((2,2,2)(2,3))."[109] I will utilize the forms of notation proposed by GOTHAM and a pulse-based code for hierarchic groupings for the purpose of the following discussion. According to that system, (223) would for instance be coded by 1010100, that is, a metric stratum of order s is represented in that string by the characters $\geq s$. Instead of using nested brackets, like ((2,2,2)(2,3)), higher-order groups are associated with higher cardinalities (20101020100).[110] This corresponds to LERDAHL and JACKENDOFF's dot notation for metric hierarchies: raised by one, the cardinalities would reflect the number of dots, assigned to each metrical position.[111]

The following examination emphasizes two aspects. First, the possibilities for higher-level groups unfold by generalizing the premise for either duple or triple grouping to all metric levels and layers. Secondly, the taxonomy of levels and layers, introduced in the previous section, leads to a disposition of hierarchic meters into three different classes. In contrast, the binary categorization into simple and mixed meters by GOTHAM is primarily related to the beat or tactus level. Though, it is desirable to expand the classification to account for differences in higher-level groupings. For instance, ((2,2),(2,2)) and ((2,2),(2,2,2)) are both simple on the beat level, though the second meter features a mixed metric structure at a higher level whereas the first meter is simple at all levels. The latter I call a *hybrid* meter as it features both simple and mixed structures related to different metric strata.[112]

Thus, recapitulating and adjusting my earlier approach,[113] I will differentiate simple, mixed, and hybrid meters. These classes better characterize the overall structure of a metric hierarchy independent from the reference level, that is, the stratum (level or layer) which is perceived as tactus. As the tactus or main beat subjectively emerges, it is basically equivocal (see section 2.2.3). Therefore, a structural description should be valid regardless of individual perception. According to the suggested taxonomy, simple meters exclusively comprise isochronous metric strata (levels). In contrast, mixed meters exhibit isochronous periods at the elementary and at the cycle level, whereas all intermediate strata are non-isochronous (layers). Finally, hybrid meters feature both levels and layers in between the elementary and the cycle periods.

It is also instructive to relate this disposition to the common notions of multiplicative (or divisive) meters as opposed to additive meters. Although it is better to avoid these

[109] Gotham, 2015b, section 2.5

[110] If these strings are regarded as numeric output series of a function applied to each elementary pulse, a pulse n belongs to a particular stratum of order s if $f(n) \geq s$ (or, $n \in s$ if $f(n) \geq s$).

[111] See figure 2.7 in section 2.2.1. In the context of rhythm analysis, Lerdahl and Jackendoff, 1983 write an extra dot to indicate the cycle level or cycle downbeat. The present examples could thus as well be notated as 2010100 and 30101020100. Though, for an abstract discussion of metric structure, this additional phase information may not be necessary. In this context, see the following discussion about *rotational equivalence*.

[112] Härpfer, 2014, p. 1027

[113] Härpfer, 2014

terms, as they also associate issues of musical notation,[114] the underlying mathematical concepts should be reflected. Metric cycles can be constructed by the iterative and/or combined application of three mathematical operations on temporal spans: addition, multiplication, and division. According to the assumption above, all operations may utilize only 2 or 3 as summand, factor, or divisor. Multiplication takes effect as an "addition" of a higher metric level, division "adds" a lower metric level, and addition extends a specific metric level by a group of two or three unit spans. Accordingly, simple meters are constructed exclusively by multiplication or division, mixed meters are made up by addition, and hybrid metric structures result from the combined application of both operations. Table B.1 (in appendix B) exhaustively lists the simple meters up to 24 elementary pulses per full metric cycle. In the next section 4.2.3, a generic notation for stratified meters is introduced which reflects the analysis given in this section and the mathematical operations underlying particular metric structures.

According to the definition above, layers are formed by non-isochronous sequences of metric groups. They can be understood as a generalization of mixed metric structure as they can basically occur at any metric level. Hence, the number of periods n included by a layer structure corresponds to the ratio between two adjacent metric levels (see figure 4.8).

TABLE 4.2: Layer necklaces and correspondent meter vectors for $n \leq 10$

periods	layer necklace	meter vector
5	(2-3)	{1,1}
7	(2-2-3)	{2,1}
8	(2-3-3)	{1,2}
9	(2-2-2-3)	{3,1}
10	(2-2-3-3)	{2,2}

Table 4.2 gives an overview on layer structures of $n \leq 10$. All mixed metric cycles of a particular n are subsumed here by two kinds of equivalence classes, a single meter vector (see the previous section 4.2.1) and a single aperiodic necklace (an equivalence class up to rotation, see section 4.1.3), henceforth called a *layer necklace*. For instance, $n = 7$ only allows layer necklace (2-2-3), and its three rotations 2-2-3, 3-2-2, and 2-3-2 yield all possible mixed metric structures of that period cardinality. "Rotational equivalence (R)" belongs to GOTHAM's "vector-identical order relations".[115] More precise, rotational equivalence R entails that an identical meter vector is shared between R-related

[114]In this context, London, 2012, p. 89, makes a good point: "Changing notational license is not only motivated by changing musical practice – in order to write down a new rhythmic idea, the notational system is forced to evolve – but also invites further changes in musical practice. The use of additive time signatures is a case in point; once they were used for indicating the metric construal of certain rhythmic figures, it became apparent that they could be used for others as well." Elsewhere, he infers that the terms *additive* and *multiplicative* "apply more to systems of rhythmic notation than to the structure of patterns of entrainment." (pp. 124 f.)

[115]Gotham, 2015b, section 6. "Imagining the meter as a cycle, different rotations of the same beat succession may be achieved by starting at a different position on that cycle [...]. Indeed, for aurally transmitted music, it may be obvious that a pattern is repeating, but much more problematic to assert a particular

metrical structures.

> Any meter with only a single entry for either duple or triple beats will have rotational equivalence between all of its orderings. In those cases, the rotational equivalence between forms is guaranteed by the vector identity that holds between them. For meter vectors with two or more of both beat types, vector identity does not automatically entail rotational equivalence.[116]

The latter is the case for meter vector {2,2}, listed in table 4.2. Matching $n = 10$, two layer necklaces, (2-2-3-3) and (2-3-2-3), correspond to this vector. However, (2-3-2-3) is a periodic necklace and therefore it repeats the structure of $n = 5$. It is thus not listed.

Expanding the premise about metrical groupings, layer structures with $n \geq 9$ may induce higher-order groupings of two or three elements. Such additional metric layers can feature more complex ratios, for instance $4{:}5$ and $5{:}4$ among the rotations of layer necklace $n = 9$.[117]

(2,2)(2,3)	(2,2)(3,2)	(2,3)(2,2)	(3,2)(2,2)
| |	| |	| |	| |
4 5	4 5	5 4	5 4

In mixed metric cycles with non-prime $n \geq 10$, second-order groupings can have equal durations and thus can "form an isochronous stratum which has to be distinguished from a layer, shaping non-isochronous groups by definition."[118] Two rotations of layer necklace $n = 10$ exemplify this, as second-order groupings of 2-3-3-2 and 3-2-2-3 would comprise five periods each (5-5). In those cases, mirror symmetry (inversion) holds between the structures of each second-order group (2-3 inverts 3-2 and 3-2 inverts 2-3). In terms of musical notation this structure could be interpreted as alternating time signatures, for instance $\frac{2+3}{8}$ and $\frac{3+2}{8}$. Thought of as a repeated metrical structure,

> it would make more sense to combine the structure into one time signature $\frac{2+3}{8} + \frac{3+2}{8}$ [...]. The five-period-level would then be the half-measure level. Generally a *measure* in this context can be defined as a period of a level n containing a metric structure with at least one lower level which is exactly repeated in the next period of level n. In other words metrical groupings within different periods of a measure are always translational symmetric to each other. The mentioned half-measure level can be regarded as a case of an *intermediate level* whose successive periods contain different, alternating metric structures.[119]

starting point. Rotational equivalence reflects this." (section 6.2) Thus, rotational equivalence also reflects and corresponds to metric phase ambiguity (see section 3.2.5).

[116] Gotham, 2015b, section 6.3

[117] cf. Härpfer, 2014, p. 1026

[118] ibid.

[119] ibid.

Such metric structures share reflection symmetry (orthogonal to the time axis) with simple meters, constructed exclusively by multiplication or division (see above, and table B.1), as well as with palindromes and MESSIAEN's *non-retrogradable rhythms*.[120] Mixed metric structures exhibiting this property "are not the only cases where intermediate levels occur, as we will see later."[121] Table 4.3 lists all layer structures resulting from necklace $n = 10$ and implied second-order groupings.

TABLE 4.3: Layer structures of $n = 10$

layer necklace	(2-2-3-3)
layer structures without intermediate level	2 0 1 0 2 0 0 1 0 0 2 0 0 1 0 0 2 0 1 0
layer structures with intermediate level	2 0 1 0 0 2 0 0 1 0 2 0 0 1 0 2 0 1 0 0

For growing $n > 10$, an increasing number of associated layer necklaces, of rotational variants thereof, and of alternative second-order groupings lead to an "exponential growth of the number of possible layer structures."[122] Eleven periods for instance imply two different layer necklaces and 14 possible layer structures (see table 4.4). The ratios of their second-order groupings are all complex (4:7 or 5:6), as no intermediate levels occur when n is prime.[123]

TABLE 4.4: Layer structures of $n = 11$

layer necklaces	layer structures
(2-2-2-2-3)	2 0 1 0 2 0 1 0 1 0 0
(2-3-3-3)	2 0 1 0 1 0 2 0 1 0 0
	2 0 1 0 2 0 1 0 0 1 0
	2 0 1 0 1 0 2 0 0 1 0
	2 0 1 0 2 0 0 1 0 1 0
	2 0 1 0 1 0 0 2 0 1 0
	2 0 1 0 0 2 0 1 0 1 0
	2 0 1 0 0 1 0 2 0 1 0
	2 0 0 1 0 2 0 1 0 1 0
	2 0 0 1 0 1 0 2 0 1 0
	2 0 1 0 0 2 0 0 1 0 0
	2 0 0 1 0 2 0 0 1 0 0
	2 0 0 1 0 0 2 0 1 0 0
	2 0 0 1 0 0 2 0 0 1 0

As pronounced, intermediate levels do not only result from mirror symmetry occurring in mixed metric structures. Two layer structures in the context of $n = 12$ feature an intermediate level due to a periodic pattern of first-order groupings (see table 4.5). The number of first-order groups "alternates in both cases between two and three every period of the intermediate level. The resulting meters could be notated as $\frac{3}{4} + \frac{6}{8}$ and

[120] cf. for instance Duţică, 2010, Papadopoulos, 2014, Toussaint, 2013, p. 233
[121] Härpfer, 2014, p. 1026
[122] ibid.
[123] cf. section 3.2.2, footnote 109, for a short discussion of perceptual properties and rhythmic embodiment of 11 cycles.

$\frac{6}{8} + \frac{3}{4}$."[124] These meters are well known and often discussed in the context of hemiola (section 3.2.4). The related rhythmic pattern [3-3-2-2-2] has been identified in various contexts.[125] TOUSSAINT for instance mentions the *guajira*, a flamenco meter emerging from this pattern, and its many occurrences, from the *abakkabuk* of the Tuareg to BERNSTEIN's *America* in *West Side Story*.[126]

TABLE 4.5: Layer structures of $n = 12$

layer structures with intermediate level	2 0 1 0 1 0 2 0 0 1 0 0
	2 0 0 1 0 0 2 0 1 0 1 0
layer structures violating WFC 5	2 0 1 0 2 0 1 0 0 1 0 0
	2 0 1 0 2 0 0 1 0 1 0 0
	2 0 1 0 2 0 0 1 0 0 1 0
	2 0 1 0 0 1 0 0 2 0 1 0
	2 0 0 1 0 1 0 0 2 0 1 0
	2 0 0 1 0 0 1 0 2 0 1 0

The durations of the second-order groupings suggested by the layer structures shown in the lower part of table 4.5, stand in 1:2 ratio. This may violate the discussed criteria of metric well-formedness, in particular WFC 3.4 (see the previous section 4.2.1). This constraint was deployed by LONDON in order to prevent contradictions within a metric hierarchy. If "a short beat is half the duration of a long beat" it may happen that "each short is equivalent to a unit of subdivision for the long."[127] Although this is not the case here because the long parts are subdivided by three, it is very unlikely to feel two strong beats in a 1:2 ratio, unless the indicated metric positions receive strong phenomenal accents. The structures listed in the lower part of table 4.5 can therefore be removed from the list of well-formed layer structures. Then, the remaining second-order ratios are either simple (1:1, resulting in intermediate levels), or complex (5:7, for instance in 201010020010).

"Generally intermediate levels divide and shape oscillating metric substructures which are combined"[128] in a layer structure. Table 4.6 lists the layer structures of $n = 15$ which feature intermediate levels comprising three periods in each metric cycle. These fife-element periods oscillate between 2-3 and 3-2 internal (first-order) groupings. All possible combinations with three elements of two types occur among the possible layer structures (a corresponds to 2-3 and b corresponds to 3-2).

The combinatorial space of larger mixed metric cycles provide a rich source of musical possibilities. Exceedingly long and complex mixed metric cycles are for instance

[124] Härpfer, 2014, p. 1027
[125] cf. sections 3.2.2 (footnote 98), 3.2.4, and 5.3.
[126] Toussaint, 2013, pp. 266 f. and pp. 271 f. See also London, 2012, pp. 156 ff.
[127] London, 2012, p. 94
[128] Härpfer, 2014, p. 1027

TABLE 4.6: Layer structures of $n = 15$ featuring intermediate levels

layer structures with intermediate level	combinations of first-order groupings ($a \mathrel{\widehat{=}}$ 2-3, $b \mathrel{\widehat{=}}$ 3-2)
2 0 1 0 0 2 0 1 0 0 2 0 0 1 0	a a b
2 0 1 0 0 2 0 0 1 0 2 0 1 0 0	a b a
2 0 1 0 0 2 0 0 1 0 2 0 0 1 0	a b b
2 0 0 1 0 2 0 1 0 0 2 0 1 0 0	b a a
2 0 0 1 0 2 0 1 0 0 2 0 0 1 0	b a b
2 0 0 1 0 2 0 0 1 0 2 0 1 0 0	b b a

employed by MESSIAEN (inspired by Indian tala),[129] in improvised Bulgarian *mega-meters*,[130] and in complex *math rock* grooves.[131] The mentioned exponential growth of this space with a linear progression of n is illustrated in table B.2 for cycles for $13 \leq n \leq 24$.[132]

Layers can be embedded in metric cycles which comprise more than the two basic levels formed by the periods of the cycle and of the elementary pulses. Consider again the above mentioned example by GOTHAM: the meter ((2,2),(2,2,2)), related to meter vector {5,0},[133] exhibits a 2-3 layer which groups the isochronous beat level. The elementary level corresponds to the binary subdivisions of the beats. I suggested to address metric cycles including simple and mixed metric structures at different metric strata generically as *hybrid meters*.[134] However, a mixed metric structure can be hierarchically extended by simple or multiplicative grouping in two directions within the metric hierarchy. It is thus necessary to clarify and differentiate the perceptual, cognitive, and musical implications of such operations. At the end of the following section it is shown that more complex hybrid meters can be conceived if a full cycle includes more than 24 elementary pulses.

4.2.3 A generic notation for stratified meters

The formal analysis and enumeration of hierarchic meter developed so far, yields a homogenous space of musical meter which will serve as background for an analysis of the metric malleability of rhythmic cycles. Basically, rhythmic cycles can be mapped onto this space to calculate goodness-of-fit values between particular rhythmic and metric cycles. The distribution of these values across the space of hierarchic meters will tell us something about metric malleability (see section 5.2, and in particular section 5.2.2).

[129] cf. Papadopoulos, 2014
[130] cf. Kirilov, 2012
[131] cf. Osborn, 2010
[132] cf. Härpfer, 2014, p. 1028
[133] cf. Gotham, 2015b, figure 2
[134] Table B.3 enumerates hybrid metric structures for cycles containing up to 24 elementary pulses (taken from Härpfer, 2014)

To complete this basis, the notation for the integrated description of the distinguished types of meter will now be standardized as a generic notation for stratified meters (GNSM in the following). This analytic description is based on the distinctions made so far and displays a metric cycle independent from musical time signature. It parallels the pulse-based data format already introduced in the previous section, which is designed to be easily implemented and processed in applications. Examples a) and b) recall the analogy between this format (in the middle) and interval notation (leftmost). The third notational variant (rightmost) corresponds to the mathematical term which describes the operation to construct the metric cycle from elementary pulses.

a) 2-3 ≘ 1 0 1 0 0 ≘ 2 + 3
b) 2-2-3-3 ≘ 2 0 1 0 2 0 0 1 0 0 ≘ 2 + 2 + 3 + 3

In b) a second layer is supposed to arise, that is, the first-order groups become elements of a second-order grouping making up a higher metric stratum. Example b) is unambiguous, as four elements can only be metrically structured in two groups of two. Though, for more complex series an explicit notation is necessary to keep the structure unique, that is, unambiguous. This can be realized with brackets, like in examples c) and d).[135]

c) 2 0 1 0 0 2 0 1 0 1 0 0 ≘ (2 + 3) + (2 + 2 + 3)
d) 3 0 1 0 2 0 1 0 0 3 0 1 0 0 1 0 2 0 1 0 ≘ (2 + 2) + (2 + 3) + (2 + 3 + 2) + (2 + 2)

These examples again demonstrate that the ratios between lengths of higher-order groups in mixed meters can be very complex (see the previous section 4.2.2). In c) the second-order ratio is 5:7 and in d) the ratios at progressive layers are 4:5:7:4 and 9:11.[136] In contrast, all metric strata which constitute simple or multiplicative meters are isochronous. Nevertheless, these meters can be notated in the same fashion as well, for instance

e) 2 0 1 0 2 0 1 0 ≘ 2 x 2 x 2
f) 2 0 1 0 1 0 2 0 1 0 1 0 ≘ 2 x 3 x 2.

The notations on the right side of examples c) to f) can be combined to notate any stratified, hierachically organized metric cycle. The correspondent GNSM data format on the left side is designed to show, which metric strata are represented by an elementary pulse or a metric position. Consider a musical situation where several subdivision levels of a non-isochronous beat are successively embodied by a developing rhythm. This may serve as an instructive illustration for a combined usage of the notational variants already introduced: if the elementary pulse level in example b) would be notated as $\frac{1}{8}$, an accurate time signature for the metric cycle would be $\frac{4}{8} + \frac{6}{8}$. If we assume the

[135] Note that this is similar to the notation of higher-level groupings suggested by Gotham, 2015b, section 2.5 (see section 4.2.2).
[136] Two third-order groups uniquely result from the four second-order groups in d). As soon as there are more than four second-order groups, nested brackets become necessary for disambiguation.

elementary pulse level of this meter to be $\frac{1}{16}$ or at a deeper binary division, we could also notate the cycle as (2 + 2 + 3 + 3) x 2 or (2 + 2 + 3 + 3) x 2 x 2. The data code for the sixteenth-note cycle and the thirty-second-note cycle respectively would be

g) 3 0 1 0 2 0 1 0 3 0 1 0 1 0 2 0 1 0 1 0

h) 4 0 1 0 2 0 1 0 3 0 1 0 2 0 1 0 4 0 1 0 2 0 1 0 2 0 1 0 3 0 1 0 2 0 1 0 2 0 1 0.

However, from the perspective of GNSM, the tactus – that is, the primary metric reference or counting level – is not defined a priori. Example h) could thus be interpreted differently, for instance as a hypermeter which groups a series of ten $\frac{2}{4}$ measures. Then, the stratum $s = 1$ (the pulses ≥ 1, see section 4.2.2) would represent the quarters or the counting level. This scope of interpretation corresponds to GOTHAM's *equivalence* relation (E).[137] Metric equivalence simply assumes a parallel metric grouping structure related to different strata or temporal levels of different meters.

> For instance, all meters with exactly one level of ternary grouping may be represented by the metrical structure $\langle 3 \rangle$ (assuming binary grouping on all other levels). If 3/8 follows 3/4 with the eighth-note pulse level held constant, then the two meters are still equivalent, but the ternary grouping occurs on a different metrical level and is therefore a considerably different entity to the counting musician.[138]

Usually, due to the temporal window for meter (section 2.2.2), mixed groupings or nested layer structures can only emerge across one level in a metric hierarchy. Within the range of 24 elementary pulses per metric cycle it is simply not possible to embed two layers in a metric cycle which are intermediated by a level. From 25 units per cycle, such structures can be conceived, for instance (2 + 3) x (3 + 2) = 25 elementary pulses (see example i) below), or (3 + 2 + 2) x (3 + 2) = 35 elementary units. However, hybrid metric structures of this length and complexity may cross the border from rhythmic movement to musical form. In other words, the metric cycle as such lies beyond the levels (or layers) which convey metric pulse. By the way, these considerations show that GNSM implies the number of elementary pulses in a metric cycle by interpreting a metric structure as a mathematical term, for instance

h) $(2 + 2 + 3 + 3) \times 2 = 10 \times 2 = 20$ ($\widehat{=}$ 2 0 1 0 2 0 1 0 2 0 1 0 1 0 2 0 1 0 1 0)

i) $(2 + 3) \times (3 + 2) = 5 \times 5 = 25$ ($\widehat{=}$ 3 0 0 1 0 2 0 0 1 0 3 0 0 1 0 2 0 0 1 0 2 0 0 1 0).

4.3 Models of metric accentuation

Chapter 2 represents a complex notion of metrical accent which developed from largely separate, and thus diverging musical and scientific discourses. Cognitive processes

[137] Gotham, 2015b, section 3

[138] ibid. (section 3.2)

(entrainment, pulse sensation, grouping and priming mechanisms, as well as categorical perception) and their interaction, leading to metric accentuation, were reviewed to establish a basis for the notion of metric ambiguity and malleability. From this perspective, metric malleability correlates with the potential of a rhythm to be perceived according to different metric accentuation patterns. As soon a rhythm gets casted into a particular metric framework, its inherent malleability gets resolved. However, this potential can become evident again if metric conflicts arise (section 2.3.3), for example, when figural grouping leads to an accentual pattern which stands in contrast to a metric top-down scheme.[139] As a limited instance of metric conflict, syncopation exclusively arises in the context of a previously induced pattern of metric accents or beats.

> This process [of beat induction], for example, facilitates the percept of syncopation, i.e., to "hear" a beat that is not carried by an event.[140]

Numerous quantitative models of metric accent and syncopation have been developed for different purposes. They will be shortly overviewed in the following section. Such models are closely related to the hierarchic, formal structures described in the previous section 4.2.2 and to other aspects examined before, like categorical perception and simultaneous pulse sensation at several metric levels. In this context, the notion of metric weight has been put forward to further differentiate the perceptual salience or the attentional status of temporal categories related to a metric framework. From the perspective of musical schemes, different weights or metric strengths of successive temporal categories represent an ideal order or hierarchy for the metric placement of rhythmic onsets. In parallel, metric weight corresponds to the strength of attentional peaks which correlate with those metric categories or "positions" in time.[141] Rhythmic tension, as well as syncopation can then be conceptualized as rhythmic opposition or antagonism in regard to this framework.[142] This happens when rhythmic onsets fall in categories of relatively low metric weight and neighboring categories of higher metric weight coincide with rests or continuations of rhythmic events.

4.3.1 Metric weight and syncopation

Computational models of meter perception apply the concept of metric weight or metrical accent strength by quantifying the relative perceptual saliences of the temporal categories defined by a meter. Correspondent approaches are based on both empirical backgrounds – as reviewed for instance in section 2.2.3 and later in this section – and formal models of meter, such as discussed in section 4.2. In particular, the concepts of

[139] cf. section 2.3: Lerdahl and Jackendoff, 1983 regard metrical accent and (figural) grouping mechanisms as independent aspects of cognitive rhythm processing. See also section 3.2.1 for a discussion of top-down and bottom-up processing.
[140] Desain and Honing, 1999, p. 29 (square brackets with supplement added by the author).
[141] cf. London, 2012, pp. 80 ff.
[142] See for instance Huron, 2006, Longuet-Higgins and Lee, 1984, Sioros and Guedes, 2014, Temperley, 1999, and Witek et al., 2014.

meter by YESTON (section 1.3), as well as by LERDAHL and JACKENDOFF, are frequently employed templates for the estimation of metric weight.[143] Commonly, the weight of a metric position, that is, a temporal category or an elementary pulse in a metrical framework, is thought of as proportional "to the metrical level that each metrical position [...] initiates. The slower the metrical level the higher the metrical weight."[144] In terms of pulse sensation, the metric weight of a position rises with the number and the saliences of pulses constituting different metric levels, which coincide on that position.[145] LERDAHL and JACKENDOFF "correlate metrical strength with hierarchical persistence and depth: 'if a beat is felt to be strong at a particular level, it is also a beat at the next larger level'".[146] More formally, they regard every pulse or beat on a metric level L which coincides with a pulse on the next higher level $L + 1$ as a strong beat on level L.[147] This results in an alternation of *downbeat* and *upbeat* which emerge from the interaction of two metric levels. A rhythmic event can therefore occur *on the beat* or *off the beat*, depending on which metric level we focus (see section 3.2.3). For instance, upbeats of level L would feel off-beat while focusing on higher levels than L. Thus, the notions of downbeat and upbeat are interwoven in the notion of the "importance of a metrical position" – its metric weight – which "is summarised by the number of metrical streams which coincide at a given point."[148]

However, LERDAHL and JACKENDOFF's graphic representation of metrical structure as a grid of dots (see figure 2.7 in section 2.2.1) implies "that the various metrical levels are equally salient and important to the metrical experience."[149] Furthermore, it complicates the identification of corresponding levels among different metric hierarchies because "it leads to strange correspondences of this kind in even the simplest examples. For instance, it attributes the same value to the bar-level in 3/4 as to the half-bar-level value in 4/4".[150] These weaknesses are shared by other models due to a common neglect to account for the influence of the levels' absolute period durations on metrical salience. This is comprehensively discussed in section 2.2 and integrated in the considerations in section 4.2. To develop a heuristics of the metric malleability of cyclic rhythms in section 5.2, the possible interaction of contrasting patterns of metrical accent will therefore be modeled on the basis of existing approaches, which are sensitive to temporal variation.[151] It is nevertheless worthwhile to examine models which are insensitive for absolute durations, but feature some other useful differentiations.

[143]See for instance Gotham, 2015a, London, 2012, Toussaint, 2013, and Volk, 2008.
[144]Sioros and Guedes, 2014, p. 637
[145]cf. Parncutt, 1994 and Gotham, 2015a.
[146]London, 2012, p. 91, quoting Lerdahl and Jackendoff, 1983, p. 19
[147]See also Lerdahl and Jackendoff, 1981, pp. 487 ff.
[148]Gotham, 2015a, p. 23. Coincidence of metrical streams relies on metrical consonance, providing that the metrical streams are nested, or "in phase" (see section 2.3.3).
[149]Gotham, 2015a, p. 25
[150](ibid.) For instance, when the eighth-note level is represented by one dot, the six resulting positions of 3/4 are indicated by 3,1,2,1,2,1 dots, whereas the eight resulting positions of 4/4 receive 4,1,2,1,3,1,2,1 dots.
[151]In particular Parncutt, 1994, Gotham, 2015a, and Flanagan, 2008.

The concept of metric *indispensability* proposed by BARLOW,[152] will be surveyed and advanced in the following sections, as it is connected to the idea of *metric coherence*. To this effect, metric coherence measures the amount of similarity between two meters in terms of their weighting structures. Sections 4.4.1 and 4.4.2 will establish the role of metric coherence, or in fact, of its counterpart *metric contrast*, for metric malleability.

Empirical arguments for the estimation of metric weights can be found in *position usage*, that is, statistical occurrences of rhythmic onsets related to metric categories within particular musical idioms.[153] For instance, HURON discusses correlations between metrical expectations based on learning processes and position usage.

> Some metric positions are more likely to coincide with note onsets than others. [...] What musicians call the "strength" of a metric position is correlated with the likelihood of a tone onset.[154]

As mentioned before, metric behavior is culturally influenced by the internalization of musical experience. Categorical distributions of rhythmic onsets in the context of specific meters in some way define and affirm the metric hierarchy as a template for expectations. Those learned patterns of rhythmic distribution within a metric scheme interact with bottom-up processes, forming emergent metric patterns. Metric bottom-up processing also involves the metric interpretation of phenomenal accents (section 2.3). From this perspective, the described notion of metric weight can be considered as rather static, and therefore insufficient to picture the dynamic process of metric emergence. It is thus important to recall in particular the dynamic interplay of phenomenal and metric accentuation.[155] Some models of metric weight therefore ignore static top-down schemes and act on the assumption that metric weight is formed and projected entirely from categorized rhythmic structures. According to HASTY's theory of projection and theories of expectancy (see section 2.2.1), these models extrapolate metric accentuation entirely from the given IOIs in a rhythm.[156]

Consequently, syncopation can generally by conceptualized as rhythmic tension which arises from differences between concurrent patterns of phenomenal and metrical accentuation. Indeed, syncopation necessarily relates to concepts of metrical accent, and thus, quantitative models of syncopation rely on quantified patterns of metric weight, as exemplified later. This corresponds to the fact that syncopation only occurs in the context of an established metric framework, as mentioned above. This is interesting from the perspective of metric malleability, as it can be concluded that the emergence

[152] Barlow, 2012
[153] cf. for instance Gotham, 2015a, Huron, 2006, and Palmer and Krumhansl, 1990
[154] Huron, 2006, p. 179
[155] Recall as well that pulse sensation can be interpreted as an extrapolation of phenomenal accent, and that the notion of different simultaneous pulse sensations is similar to the hierarchic model of Lerdahl and Jackendoff, 1983.
[156] See for instance, Desain, 1992 and Volk, 2004.

of syncopation depends on the resolution of metric ambiguity by the process of metric interpretation.

TOUSSAINT broadly illustrates the formal imprecision of traditional definitions of syncopation which he characterizes as a "slippery human perceptual skill."[157] Although a "fuzzy concept" complicates quantitative modeling, he argues that it may be "possible to construct unambiguous mathematical definitions of notions that may be used as [...] models that replace the traditional concept of syncopation."[158] Indeed, a number of formal metrics were developed which encourage this view (see for instance the measures of GÓMEZ et al.,[159] KEITH,[160] LONGUET-HIGGINS and LEE,[161] as well as SIOROS and GUEDES).[162] Though, many of these approaches operate on the premise of static and limited metric frameworks which are formalized as patterns of metric weight. From the perspective of metric malleability, this seems inappropriate, specifically if we assume that a complex rhythmic necklace which contrasts to any possible metric framework is also highly malleable (see section 5.3). Furthermore, possible effects of pulse rate or the absolute durations of rhythmic intervals on the particular feel of syncopation or generally on metric preference[163] are generally ignored by these measures.

An exception is FLANAGAN's proposal, who is particularly interested in "the amenability of a rhythm to gestalt flip, in which a listener reevaluates the location of the beat."[164] His method to quantify syncopation accounts for malleability, as it includes the calculation of saliences for all pulses – FLANAGAN calls them "meters" – the periods of which are factors of the rhythmic cycle length, and which could occur in any phase relation to it. Thus, the rhythmic cycle is regarded as a necklace,[165] which is amenable for different positioning of the metric starting point of the cycle. In this way, he adopts PARNCUTT's model[166] to estimate "induction strengths across a population of candidate meters".[167] Syncopation is then defined in relation to a particular meter. The amount of syncopation rises (and approaches a maximum of 1) when the relative strength of that meter

[157] Toussaint, 2013, pp. 67 f. "There must be more than 50 traditional definitions of syncopation adorning the pages of dictionaries, books, and Internet sites [...], most have their own particularities. However, to a mathematician, they all have one thing in common: vagueness." (p. 67)

[158] Toussaint, 2013, p. 68

[159] cf. Colannino, Gómez, and Toussaint, 2009. The *weighted note-to-beat distance* measure (WNBD) was originally proposed in Gómez, F., A. Melvin, D. Rappaport, and G. Toussaint (2005). "Mathematical Measures of Syncopation". In: *Proceedings of BRIDGES: Mathematical Connections in Art, Music, and Science*. Banff, Canada, pp. 73–84.

[160] Toussaint, 2013, pp. 70 ff., provides an illustration of Keith's measure of syncopation, proposed in Keith, M. (1991). *From Polychords to Pólya: Adventures in Music Combinatorics*. Princeton: Vinculum Press.

[161] Longuet-Higgins and Lee, 1984

[162] See Sioros, Holzapfel, and Guedes, 2012 and Sioros and Guedes, 2014.

[163] That is, the emergence of a subjective metric framework.

[164] Flanagan, 2008, p. 635

[165] Flanagan, 2008 uses a different term, relating the necklace perspective to the set-theoretic notion of *beat-class set type*.

[166] See sections 2.2.3 and 5.2.1

[167] Flanagan, 2008, p. 636

within in the range of the induction strengths (or saliences) gets lower and the overall dispersion of the saliences gets more balanced at the same time.[168] We will further examine this approach, and in particular, the aspect of pulse-salience distribution in sections 5.2.1 and 5.2.2, employing some of these ideas in modeling metric malleability.

Another example of a syncopation model which makes use of a flexible and context-dependent approach to meter is the one developed by VOLK. By *inner metric analysis*, "a quantitative method of creating weight profiles based on pulse descriptions [...] a metric hierarchy is induced"[169] which reflects the actual rhythmic structure. Similar to PARNCUTT's pulse-match salience (section 2.2.3), the pulse weights are deduced from the number of onsets in a musical texture, matching with a pulse which is defined by a certain period and phase at a symbolic, notational level. This radical bottom-up approach aims at estimating differences between perceived ("inner") meter and notated ("outer") meter. The amount of divergence is conceptualized by the notion of *metric coherence*,[170] that is, the more divergent the inner and the outer meter, the lower the level of metric coherence. VOLK establishes that syncopation consistently affects inner meter to deviate from notated meter. "Hence, syncopation prevents metric coherence."[171]

Perspectives on syncopation like those of FLANAGAN and VOLK are promising to further explore the relation between syncopation and malleability. Though, another questionable aspect is the generalizability of the concept of syncopation.

> Syncopation is very much a Western concept, and for some types of music, new mathematical substitutes for syncopation, which are not culturally dependent, may be more appropriate and useful.[172]

The variety of musical engagement and listening behaviors across cultures, discussed for instance in section 3.1.2, indeed affects the way how the interplay between metric and "contrametric" cues in rhythm is processed. As TEMPERLEY pointed out, "Western perception involves shifting the metrical structure in order to better match the phenomenal accents, while the African perception favors maintaining a regular structure even if it means a high degree of syncopation."[173] This level of subjectivity has to be taken

[168] cf. Flanagan, 2008, p. 639, equation (6)
[169] Volk, 2008, p. 271
[170] This term is employed as well by Barlow, 2012, but in a slightly different sense. His measure of metric coherence is examined in section 4.4.1.
[171] Volk, 2008, p. 271
[172] Toussaint, 2013, p. 68
[173] Temperley, 2000, p. 79, amplifying that syncopation is a form of rhythmic complexity which leads to "conflicts with the underlying metrical framework; and African listening requires a greater ability to maintain a steady beat despite conflicting accents. [...] However, the difference between African and Western rhythm is not simply a matter of complexity. Viewed in another way, the greater tendency of Western listeners to shift their metrical structures in response to phenomenal accents might be seen as a greater sensitivity to metric shift in the music."

into account, as it has influence on aesthetic judgements about rhythmic tension.[174] Thus, the above mentioned concept of inner meter – and its "emancipation" from outer meter by syncopation – can not properly represent those aspects of subjective metric interpretation.

To this effect, quantitative research on rhythm has substantially addressed a more general view on rhythmic complexity. However, the diversity of rhythmic complexity measures is rather confusing at first sight.[175] Aspects of complexity which relate to meter may be more specifically understood as more precise alternatives of the syncopation concept. Some complexity measures explicitly estimate the amount of cognitive effort to process a rhythmic structure according to an implied metric framework.[176] TOUSSAINT brought forward a "mathematical definition of syncopation" called *metrical complexity*, which involves predefined metric weights according to LERDAHL and JACKENDOFF's system for the distribution of accentual strengths over a metrical grid.[177] In other words, such meter-related complexity measures also assume a prevailing meter or fixed metric expectations.[178] The perceptual complexity of a particular rhythm may however vary to a great extend in the context of different metric frameworks. In contrast, the notion of *off-beatness* at least accounts for the possibility that a metric cycle can include different numbers of tactus beats.[179] Figure 4.9 illustrates that all four possible regular subcycles of a metric 12 cycle (comprising the elementary pulses indexed by {0, 2, 4, 6, 8, 10}, {0, 3, 6, 9}, {0, 4, 8}, and {0, 6} respectively) omit the elementary pulses 1, 5, 7, and 11. Thus, these positions are off-beat for any regular beat that coincides with pulse 0 at every cycle, and that spans at least two elementary pulses. The off-beatness of a rhythm, related to the 12 cycle, can accordingly be defined by "the number of onsets that the rhythm contains at these four distinguished locations."[180] This simple measure can be differentiated by calculating "an off-beatness weight for every pulse in the cycle

[174] Burns, 2010 for instance explores "rhythmic archetypes", commonly used in African Music (including the Diaspora) as prototypes for the rhythmic organization on various levels. "Performers draw on these expected rhythmic devices to create compositions and variations that are aesthetically satisfying in their local contexts, not to create tension." (section [89]). Other related phenemena may be found in traditional dance forms, as for instance the Norwegian *gangar*, "in which the variable meter (6/8 - 3/4) of the music not always coincides with the movements of the related popular dance. However the dance naturally follows the overall metric feeling, the internal movement of the music and is not regarded as particularly syncopated by the performers." (Moelants, 1997, p. 270, referring to Blom, J. and T. Kvifte (1986). "On the problem of inferential ambivalence in musical meter". In: *Ethnomusicology* 30, pp. 491–517.)

[175] See for instance Thul and Toussaint, 2008 for a comparative study in which 32(!) measures of rhythmic complexity are evaluated.

[176] For instance Pressing, 1997 and Shmulevich and Povel, 2000.

[177] Toussaint, 2013, pp. 69 f.

[178] From the perspective of models of temporal expectancy (cf. sections 2.2.1 and 5.1.3), syncopation and metric complexity are "easy to formalize [as] the violation of a maximum in expectancy by the absence of an event at that time" (Desain, 1992, p. 450).

[179] cf. Toussaint, 2013, pp. 100 ff.

[180] Toussaint, 2013, p. 101. Colannino, Gómez, and Toussaint, 2009, p. 125, mention that this measure is "based on the underlying polyrhythmic group-theoretic structure of the meter." One important feature of this structure is that the off-beat positions coincide with the possible generators of the cyclic group Z_n acting on the N cycle (cf. sections 4.1.1 and 4.1.3, as well as Benson, 2007, pp. 317 ff., and Toussaint, 2013, p. 102).

that depends on how many meters render the pulse off-beat."[181] For all non-prime N cycles, off-beat positions can be specified according to the same principle, resulting for example in the off-beat positions {1, 3, 5, 7, 9, 11, 13, 15} for the 16 cycle and {1, 5, 7, 11, 13, 17, 19, 23} for the 24 cycle.

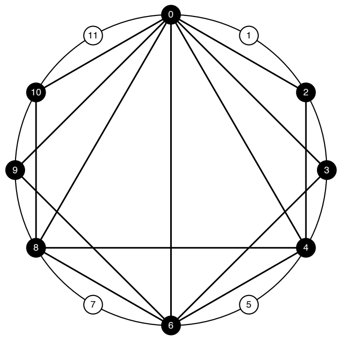


FIGURE 4.9: Off-beat pulses in a 12-cycle (Toussaint, 2013, p. 101).

Although off-beatness considers the variety of possible tactus periods, it has – like the other mentioned measures – less significance in the context of rhythmic necklaces. The off-beatness of a rhythm can radically change when shifted in relation to a metric cycle. Figure 4.10 shows two metric rotations of the rhythmic necklace [2-4-2-4]. In the context of the metric scheme of figure 4.9, the left rhythm does not articulate off-beat positions (corresponding to a low off-beat value 2.25 according to the weighted measure described in footnote 181). In contrast, a counterclockwise rotation by one elementary pulse (see the rhythm to the right) exactly matches the four off-beat positions and thus has maximal off-beatness (value 5.0).

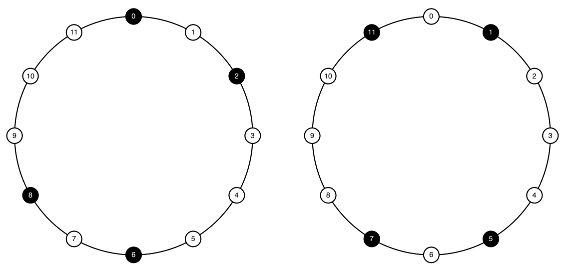


FIGURE 4.10: Two metric rotations of the rhythmic necklace [2-4-2-4] resulting in contrasting off-beatnesses.

Thus, the off-beatness measure does not directly indicate, to which extend a rhythm deviates from what we have called a metric rhythm in section 3.2.2.[182] In particular,

[181] ibid: "Thus, for a 12-pulse cycle with meters [12], [6-6], [4-4-4], [3-3-3-3], and [2-2-2-2-2-2], we would obtain, for pulses 0 to 11, the weights {0, 5, 4, 4, 3, 5, 2, 5, 3, 4, 4, 5}. The generalized off-beatness value would then be calculated by summing these weights for all pulses that have an onset, and normalizing by dividing the sum by the number of attacks in the rhythm."

[182] This may however be estimated by measuring off-beatness for all metric rotations of a rhythmic necklace. It would be interesting to explore, if and how the resulting dispersions of these values correlate with

isochronous rhythms are completely off-beat when shifted in between isochronous metric beats.[183] In other words, off-beatness relates to a predefined metric phase, whereas the period of the beat can be assumed as variable. This has some parallels to perceptual features of polyrhythms (section 3.1.1). Finally, possible metric frameworks with mixed beat durations are ignored by this measure.

To summarize, measures of rhythmic complexity which are related to meter, as for instance the off-beatness, make "sense only when the rhythm is viewed in the context of an underlying regulative beat structure."[184] In this context, computational measures of syncopation have gained success in modeling the cognitive complexity of rhythms in the context of metric frameworks.[185] In contrast, for an exploration of metric malleability, the metric interpretation of a rhythmic cycle cannot be preconceived. A rhythmic cycle should thus be related to possible candidates of metric frameworks, and for all plausible constellations, the measuring of syncopation or metric complexity should be performed. The evaluation of metric ambiguity may then be based on a comparison of the obtained values. Similar values for different constellations may characterize metric malleability.[186]

4.3.2 Metric indispensability

As a prevalent feature, computational measures of metric weight correlate the strength of a metric category with the number of metric levels it constitutes.[187] The GNSM code corresponds to that principle as it distinguishes metric strata by different digits. For instance, 20102010 represents a simple binary meter. This can also be interpreted as a weighting scheme indicated by integer values.[188] Accordingly, the same weight value is commonly assigned to all pulses of a particular metric level which are not part of higher levels, in the following called *residual pulses*.[189] However, it is arguable if this is consistent in regard to musical intuition and metric position usage, as described

other measures of rhythmic irregularity and complexity. Lower differences among the values may indicate a higher irregularity. This may also correlate to the measures of metric ambiguity and malleability, proposed in section 5.2.

[183]cf. Toussaint, 2013, p. 283, figure 37.2

[184]cf. Toussaint, 2013, p. 282

[185]Witek et al., 2014, p. 2, comment on the study of Thul and Toussaint, 2008 (see footnote 175) that "measures of syncopation outperformed other measures in explaining the behavioural data from four separate studies. The data comprised of judgements regarding perceptual, metric and performance complexity of rhythmic patterns. It was found that models of syncopation better explained the variability in these judgements, compared to for example standard deviation and entropy [...]. Syncopation therefore appears to be a more appropriate predictor of perceived rhythmic complexity." (For other aspects relating to *information entropy* which are relevant for this study, see sections 2.3.4 and 5.2.2.)

[186]It would also be interesting to explore correlations between this approach and measures of metric ambiguity (see section 5.2).

[187]See Gotham, 2015a for a recent discussion.

[188]In the case of the GNSM code, a digit stands for an elementary pulse and expresses the number of coinciding metric strata at that pulse minus one.

[189]Though the off-beatness measures the opposite or negative of metric weight related to several metric beats at once, there is also no differentiation in terms of weight among the beats of a particular beat.

in the previous section 4.3.1. From that perspective, one may consider the hierarchic relations of adjacent metric positions. Exploring statistical position usage in the context of metrical schemes, HURON distinguishes between "zeroth-order probabilities that favor particular beats" and "first-order temporal probabilities."[190] From corpus-based analyses he concluded that both are decisive for particular musical sources. "When a tone appears at a certain point in the metric frame, its appearance makes some ensuing moments more likely."[191] That is, first-order probabilities correspond to distributions of intervals between successive rhythmic onsets within a metric scheme.

> Onset moments tend to cluster together. We might refer to this tendency as *metric proximity*. Thus, we can see that two phenomena seem to dominate the *when* of music: the simple zeroth-order probability of events where some metric positions (like the downbeat) are more likely to occur, and the first-order metric proximity where an event tends to increase the likelihood of a neighboring ensuing event. Said another way, event successions are shaped by *metric hierarchy* and *metric proximity*.[192]

Metric proximity is for example reflected by anacruses, or "tendency moments"[193] suggested by a metric framework. That is, onsets at positions of less metric relevance or expectancy (coinciding with lower metric levels) tend to be ensued by onsets at the following strong metric position. This tendency causes schematic expectations, and syncopation is exactly the result of a violation of such expectations. Thus, a metric hierarchy is affirmed by the usage of anacrustic onsets at weaker metric positions, directly followed by onsets at stronger positions (higher levels). In contrast, a metric framework is challenged when onsets at weaker positions are not followed by metrically stronger onsets.

These findings may be abstracted to a more general assumption: as an effect of metric proximity, the metric relevance of a position is influenced by the weights of neighboring positions which belong to higher levels. BARLOW's computational measure of *metric indispensability*[194] reflects this by estimating individual weights for all distinguished metric positions. This differentiation also affects the relative weights of the residual pulses of a particular metric level. Their individual linear contexts within the metric cycle are factors involved in the quantification of their metric relevance. The resulting values of metric strength are associated with the amount of support, a pulse contributes to a metric framework by its rhythmic articulation. Expressed in BARLOW's terminology, the more weight a pulse holds according to a particular meter, the more indispensable is its articulation for the meter's rhythmic representation.

[190] Huron, 2006, p. 243. "Beats" here rather mean elementary pulses, that is, all temporal positions which are distinguished by an analytical perspective.
[191] ibid. The discussed data are "based on an analysis of 544 German folk songs in 4/4 meter." (ibid.)
[192] Huron, 2006, p. 244
[193] Huron, 2006, p. 245
[194] Barlow, 2012, Part I, p. 45

As outlined in the previous section 4.3.1, metric hierarchies with a certain depth imply embedded alternations of stronger and weaker metric periods which correspond to the succession of pulses associated with different metric strata. From a grouping perspective, the periods of higher metric strata contain groups of lower-level periods. For instance, a metric 4 cycle features binary grouping at two levels, that is, two groups of two pulses, grouped again by the metric cycle. The corresponding GNSM code 1010 may be interpreted as an alternation of *strong-weak-strong-weak*. Evidently, the rhythmic articulation of the stronger positions better supports the meter than an articulation of the weaker. However, pairs of rhythmic patterns like [x.xx] and [xxx.], or [x.x] and [xx.] – which both are rotations of a single rhythmic necklace – would make no difference from the current perspective. That is, they provide the same overall support to the meters 1010, or 100 respectively, as the metric weights of the residual pulses of the elementary level are not differentiated. At that point, the principle of *fundamental indispensability* provides an appropriate method to distinguish such cases. This is indeed desirable to account for metric proximity and correspondent effects like meter stabilization by anacruses, exemplified in the following. To calculate the individual indispensabilities for each pulse in a stratified meter, an

> indispensability formula is designed for the successive division of a metre by arbitrary prime number divisors. A prerequisite for each divisor is a *fundamental indispensability* series for the metric level containing the corresponding prime-quantity of pulses, e.g. for 2 the series 1 0, for 3 the series 2 0 1, for 5 the series 4 0 3 1 2 etc.[195]

In other words, if a metric period is divided into shorter periods by introducing a lower level or a faster metric stratum, a gradation of the indispensabilities or metric weights of the resulting periods is assumed.[196] A division by two makes the first period stronger than the second (1 0) and a division by three results in a series with a strong, a weak and a mid-weight period (2 0 1).[197] These patterns suggest certain types of rhythmic movement. The principle of *arsis* and *thesis* in Greek rhythm theory of poetry, dance and music – the *up* and *down* of an organic, breath-like movement – corresponds with the division by two, an alternation of strong and weak.[198] The third period resulting from a division by three follows strong and weak as an anacrusis or upbeat to the next expected strong beat.[199] Note that the formula for metric indispensability implies these differentiations for any level of metric division into two or three parts. It is thus applied recursively for meters with more than two levels.

[195] ibid.

[196] Thus, for a series of n pulses, the gradation of indispensability values corresponds to the set $\{0, ..., n-1\}$, that is, each pulse is associated with a unique integer value modulo n.

[197] For a discussion of the relation of metric multiplication and division to metric accent see Hasty, 1997, p. 116: "accents *produce* multiplication. Conversely [...] division will produce the accents (by the rule: first half accented – second half unaccented)."

[198] cf. for instance Fraisse, 1982, p. 174

[199] Hasty, 1997, pp. 133 ff., refers to the projective effect of an upbeat as a *deferral* of the next strong beat. Therefore, an unexpected upbeat may be inferred retrospectively from a realized downbeat.

Whereas it is commonly agreed that a metric division by two results in an accentual pattern of *strong-weak* (1 0), the gradation 2 0 1 for metrical accent strengths implied by a division by three entails further discussion, involving issues of perceptual grouping and musical predisposition. LERDAHL and JACKENDOFF theoretically separate grouping and meter, and analyze the interaction of these aspects as induced by rhythmic patterns.[200] They assume

> that a beat as such does not somehow belong more to the previous beat or more to the following beat [...]. But once metrical structure interacts with grouping structure, beats do group one way or the other. If a weak beat groups with the following stronger beat, it is an *upbeat*. If a weak beat groups with the previous stronger beat, it is an *afterbeat*.[201]

In other words, metric grouping generally depends on rhythmic grouping structure, which for its part depends on the temporal interval structure of rhythmic patterns. However, metric ambiguity and malleability prevent from definite relationships. As discussed from the outset, already simple rhythmic cycles may be metrically grouped in several ways (see section 1.1 and figure 1.1).

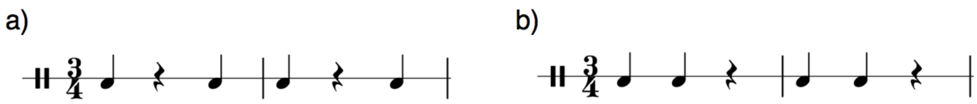


FIGURE 4.11: Two metric interpretations of the rhythmic necklace [2-1]

Figure 4.11 shows two metric interpretations of the rhythmic necklace [x.x], already considered before. In terms of "additive" hearing, we perceive an alternation between a longer and a shorter interval.[202] In theoretical discourses, there has been disagreement about the most natural representation of such patterns, advocating either L-S grouping as in figure 4.11 a), or S-L grouping, like in figure 4.11 b).[203] This relates to metric phase ambiguity (section 3.2.5), as it involves shifts of the regular grouping period, and thus, of metrical accent. Regarding the metric 3 cycle, I claim that – at appropriate pulse rates – articulations on pulses 1 and 3 (upbeat) lead to a clearer perception of the meter than articulations on pulses 1 and 2 (afterbeat). First, this is in accord with the mentioned metrical stabilization effect of anacruses. Secondly, if the shorter interval duration is less than about 300 ms, there is a considerable effect of durational accent (see section 2.3.2). In this case, the rhythmic onset before the longer interval will be perceptually more salient, and therefore it is most likely to coincide with metrical stress. "A well-known principle of Western rhythm is that notes on strong beats tend to be

[200] Lerdahl and Jackendoff, 1981, pp. 494 ff. They claim that "*groups do not receive metrical accent;* and *beats do not possess any inherent grouping.*" (p. 494, italics in source)
[201] Lerdahl and Jackendoff, 1981, p. 496
[202] cf. section 2.3.4 and section 3.2.1.
[203] cf. Benjamin, 1984, pp. 356 f.

longer than notes on weak beats".[204] Thus, the preference for interpretation a) may also a result of (culturally dependent) musical exposure and rhythmic internalization.[205] Summing the fundamental indispensability values of the two articulated pulses in both rhythms as well represents the same preference. According to the gradation 2 0 1 for ternary metric subdivisions, interpretation a) reaches a higher score (3, versus 2 for interpretation b)). This can be interpreted as a better support or embodiment of the ternary metric structure.[206] Nevertheless, it remains an issue of metric ambiguity which at last is subjectively resolved.

The fundamental indispensabilities can be derived from all indispensability series calculated for meters which are constituted by simple (multiplicative) stratification. For instance, the series for the meter 3 x 2 x 2[207] is, according to the formula outlined below, 11 0 6 3 9 1 7 4 10 2 8 5.[208] The following array relates this series to the correspondent GNSM code, making obvious that the indispensability values are distributed to continuous ranges for each level.

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

The higher the level an elementary pulse can be associated to, the higher the range to which its indispensability value belongs. Now, to receive the indispensabilities of a certain level separated from lower levels, a subtraction can be made by

> the difference between the quantity of pulses at the [elementary] level [...] and those at the desired level from the indispensability, keeping only non-negative numbers, e.g. for $\frac{3}{4}$ [...]:
>
> 1st Level (pulse quantity 3: subtract $12 - 3$, i.e. 9): 2 _ _ _ 0 _ _ _ 1 _ _ _
> 2nd Level (pulse quantity 6: subtract $12 - 6$, i.e. 6): 5 _ 0 _ 3 _ 1 _ 4 _ 2 _[209]

[204] Temperley, 2010, p. 362, referring to Lerdahl and Jackendoff, 1983. Temperley, 2010 discusses modeling variants that account for this principle of "common-practice rhythms" by a differentiation of metric weights similar to metric indispensability. Huron, 2006, p. 245, accordingly states that in "Western music, the prototypical meter-defining rhythm is a long-duration note followed by one or two short-duration notes. Two examples are the quarter note followed by two eighth notes, and the dotted eighth note followed by a sixteenth. These types of temporal structures make it easy for the listener to infer beat and downbeat placement."

[205] Also, Parncutt's data reproduced in Appendix A support this view. In particular, the tactus choices for the cyclic patterns [2-1] and [2-1-1], listed in figures A.1 (b) and A.2 (c) prove a strong tendency to favor long-short groupings and correspondent tactus phases.

[206] The theory of indispensability basically claims that the rhythmic support of a meter, or the "*metric field strength*" (Barlow, 2012, Part I, p. 44) is generated and represented by the sum of indispensability values relative to a fixed amount of articulations (cf. Barlow, 2012, Part II, p. 33, Diagram Γ31 d)

[207] The order of multiplication is from higher/slower levels to lower/faster levels, cf. section 4.2.3.

[208] Barlow, 2012, Part I, p. 45

[209] ibid.

After this operation, a fundamental indispensability series for the first (highest) level is returned (2 0 1). On the second level the periods of the first level are divided by two. Hence, the resulting pairs of indispensability values (5 0), (3 1) and (4 2) all show a stronger period followed by a weaker period as in the fundamental indispensability series of two periods (1 0). This principle is continued in the resulting pairs of the next lower binary division level (11 0), (6 3), (9 1), (7 4), (10 2) and (8 5).

The algorithm to calculate the fundamental indispensabilities for a metric level containing any prime number of periods is recursively linked to the main indispensability formula.[210] However, to apply the theory to all possible meters including mixed and hybrid structures (see section 4.2.2), it has to be extended or modified. For the division of a metric period by five, the fundamental indispensability algorithm suggests the pulse weight series 4 0 3 1 2. In contrast, as discussed in section 4.2.1, a metric unit including fife pulses can be regarded as a mixed metric structure which contains a layer of non-isochronous groupings (2 + 3 or 3 + 2). This inherent stratification can be indicated by the possible layer structures 10100 and 10010. To calculate indispensability series for both variants, or for other possible and more complex mixed meters, the system is either not flexible enough or it would rely on personal estimation.

> The prime-number fundamental indispensability Ψ can be personally estimated or worked out by the formula in F32b;[211] the basis here is the indispensability series for a metric level with one pulse less, in which the highest-level divisors are the largest (e.g. to find Ψ for a cycle of 23 pulses, the values are first calculated according to the ψ-formula for 22 pulses, whereby 22 is represented as 11x2; the values for 11 pulses are based on those for 10, understood as 5x2, etc.). The dropped pulse is reinstated between the last two pulses of the 'reduced' metre (e.g. $\frac{4}{8}$ standing in for $\frac{5}{8}$).[212]

This illustrates the necessity to adapt the approach of metric indispensability, to cover all variants of mixed meters and non-isochronous groupings at all metric levels, that is, the set of possible metric structures developed in section 4.2.2.

4.3.3 An extended indispensability algorithm

My approach to an alternative or modified indispensability algorithm which applies to the set of meters explored in section 4.2.2 is based on the analysis of the principles by BARLOW, outlined in the previous section (4.3.2). They are resynthesized to calculate the indispensabilities for any stratified meter in GNSM notation (see section 4.2.3). In the following, a plain description of this algorithm by means of practical examples is given.

210 Barlow, 2012, Part II, p. 34
211 see ibid.
212 Barlow, 2012, Part I, p. 45

1. Input: any well-formed meter represented in GNSM notation (the number of elementary pulses n is therefore given by the length of the GNSM structure).
2. get the sum of pulses (2 or 3) which constitute the top level or layer (l_{max}) and assign a fundamental indispensability sequence (1 0 or 2 0 1) to this level or layer, for instance

   2 0 1 0 1 0 2 0 1 0 1 0 (GSM code)
   1 – – – – – 0 – – – – – (fundamental indispensabilities for top layer)
   or
   2 0 1 0 2 0 1 0 2 0 1 0
   2 – – – 0 – – – 1 – – –

3. Iteration through all levels and layers l, from the second highest to the lowest (starting at $l_{max} - 1$), in four steps (a - d):

   a) copy the values – which were set before to the pulses of all higher levels and layers[213] – to the preceding pulses (in the cycle modulo n) of level or layer l[214] and count the number of copies.

   The following example shows the steps until here for a $\frac{3}{4}$ meter with $\frac{1}{8}$ elementary pulse. The number of copies (c) in an iterative step for a level or layer is always equal to the number of pulses corresponding to all higher strata (see also the next example illustrating step 3b):

   $\frac{3}{4}$ meter

	1 0 1 0 1 0	
step	resultant indispensability values	
2)	2 – 0 – 1 –	
3a)	2 0 0 1 1 2	3 copies ($c = 3$)

   b) add c to the indispensability values of all pulses corresponding to higher strata (see next example below), meaning to perform an equivalent number of additions as the value to be added.

   An iteration of steps 3a) and 3b) through all $l < l_{max}$ would already terminate the operation for all simple (multiplicative) meters of – as expressed in BARLOW's terminology – *order* 0 or 1:

   > a $\frac{12}{16}$ bar [...] exhibits a stratification of 2 x 2 x 3. If this stratification is examined in further depth with 32nd [...], 64th-notes etc., the geometric series given could proceed by constant bisection ..x 2 x 2 x 2 x 2..; the position of the series' last divisor which is not 2 determines the order

[213] Recall that *metric layers*, defined in section 4.2.1, involve non-isochronous groupings, as in the following example.
[214] Note that in a linear sequence this is realized by copying the value of the first pulse to the last.

of the metre – according to this reasoning, $\frac{12}{16}$ is a metre of 3rd order, $\frac{6}{8}$ (=2 x 3) of 2nd, $\frac{3}{4}$ (=3 (x 2 x 2...)) of 1st and $\frac{4}{4}$ of 0th order.[215]

The iterative steps 3c) and 3d) described below are only required for simple meters of order > 1 and for mixed meters. Thus, the algorithm terminates, demonstrated as follows for a $\frac{3}{4}$ meter with $\frac{1}{16}$ elementary pulse (a simple meter of order 1) and a $\frac{2}{4}$ meter with $\frac{1}{16}$ elementary pulse (a simple meter of order 0). First the meters are represented in GNSM notation, and then the iterative steps are shown with the resultant intermediate indispensability series. The last row exhibits the completed indispensability sequences after two iterations:

$\frac{3}{4}$ meter

	2 0 1 0 2 0 1 0 2 0 1 0	
step	resultant indispensability values	
2)	2 – – – 0 – – – 1 – – –	
3a)	2 – 0 – 0 – 1 – 1 – 2 –	$(c = 3)$
3b)	5 – 0 – 3 – 1 – 4 – 2 –	
3a)	5 0 0 3 3 1 1 4 4 2 2 5	$(c = 6)$
3b)	11 0 6 3 9 1 7 4 10 2 8 5	

$\frac{2}{4}$ meter

	2 0 1 0 2 0 1 0	
step	resultant indisp. values	
2)	1 – – – 0 – – –	
3a)	1 – 0 – 0 – 1 –	$(c = 2)$
3b)	3 – 0 – 2 – 1 –	
3a)	3 0 0 2 2 1 1 3	$(c = 4)$
3b)	7 0 4 2 6 1 5 3	

c) if a meter contains any metric group of three periods on any level or layer below l_{max}, like in simple meters of order > 1 or in any mixed or hybrid structure, the situation after step 3b) would look like in the following example involving a mixed meter. In this case, first copy the value from the third period in the group to the second period for all groups of three periods:

	1 0 0 1 0 0 1 0	
step	resultant indispensability values	
2)	2 – – 0 – – 1 –	
3a)	2 – 0 0 – 1 1 2	$(c = 3)$
3b)	5 – 0 3 – 1 4 2	
3c)	5 0 0 3 1 1 4 2	

Second, count the number of three-period groups (d) on the current level or layer. If there is more than one such group ($d > 1$), meaning that there was more than one value copied, look for "gaps" in the numeric series of the copied values. If so, all higher values causing gaps have to be lowered in order to get a continuing series of values, for instance a series like $1, 3, 4, 6$ would be transformed to $1, 2, 3, 4$. Here are two examples to illustrate this:

	1 0 0 1 0 1 0 0
step	resultant indisp. values
2)	2 – – 0 – 1 – –

	2 0 0 1 0 2 0 0 1 0 0
step	resultant indispensability values
2)	1 – – – – 0 – – – – –

[215] Barlow, 2012, Part I, p. 44

3a)	2 – 0 0 1 1 – 2	($c = 3$)
3b)	5 – 0 3 1 4 – 2	
3c)	5 0 0 3 1 4 2 2	(series: $0, 2$)
	5 0 0 3 1 4 1 2	(–> $0, 1$)

($d = 2$)

3a)	1 – – 0 – 0 – – 1 – –	($c = 2$)
3b)	3 – – 0 – 2 – – 1 – –	
3a)	3 – 0 0 2 2 – 1 1 – 3	($c = 4$)
3b)	7 – 0 4 2 6 – 1 5 – 3	
3c)	7 0 0 4 2 6 1 1 5 3 3	(series: $0, 1, 3$)
	7 0 0 4 2 6 1 1 5 2 3	(–> $0, 1, 2$)

($d = 3$)

d) finally, add d to the values of all pulses corresponding to higher strata and to the value of every period on the current level or layer which is not the second period of a three-period group. Accordingly, the two examples from above continue as follows:

	1 0 0 1 0 1 0 0	
step	resultant indisp. values	
...	...	
3c)	5 0 0 3 1 4 2 2	(series: $0, 2$)
	5 0 0 3 1 4 1 2	(–> $0, 1$)
3d)	7 0 2 5 3 6 1 4	

	2 0 0 1 0 2 0 0 1 0 0	
step	resultant indispensability values	
...	...	
3c)	7 0 0 4 2 6 1 1 5 3 3	(series: $0, 1, 3$)
	7 0 0 4 2 6 1 1 5 2 3	(–> $0, 1, 2$)
3d)	10 0 3 7 5 9 1 4 8 2 6	

As demonstrated, the algorithm is actually applicable by hand. For practical tests, an implementation in form of a C++ function is provided in appendix C. A comparison of the output of this algorithm with the output of the method by Barlow for metric indispensability proved that there is complete accordance for all meters, which are possible to construct by the principles suggested by Barlow (see section 4.3.2). Furthermore, the classification developed in section 4.2.2 involves a larger set of possible meters. The extended algorithm was developed to provide plausible indispensability series for all meters in this set.

4.3.4 Summary

The concept of metric indispensability suggests a weighting system for metric positions, involving a level of differentiation that makes it interesting for applications of rhythm generation.[216] It may also be appropriately integrated in probabilistic models of meter induction. For instance, Temperley's Bayesian approach[217] requires a

[216] cf. for instance Bernardes, Guedes, and Pennycook, 2010, Sioros and Guedes, 2011a, and Sioros and Guedes, 2011b

[217] cf. Temperley, 2007

method which estimates the probabilities of the occurrence of particular rhythmic patterns in the context of a given meter.[218] These probabilities have to be calculated as variables for BAYES' theorem, which is employed by TEMPERLEY to infer in turn a meter from a rhythmic pattern (see section 5.1.2).

Metric indispensability may reflect first-oder relations between rhythmic onsets in a metric context, discussed in section 4.3.2 in terms of metric proximity. Though, the weight of metric positions representing "tendency moments" cannot be regarded as independent from concrete rhythm. A rhythmic onset which stabilizes a meter when it is perceived as an anacrusis of a following onset on a stronger metric position can turn into a syncope when the following onset is dropped. In the latter case, its weight or influence is rather "anti-metric", as it turns to destabilize the meter. Another substantial contextual problem of this approach is its ignorance of absolute durations and the influence of pulse rates and durational accents on pulse saliences (see sections 2.2.2, 2.2.3, and 2.3.2), and therefore, on the weights of metric positions. The discussion of figure 4.11 in section 4.3.2 and the examinations in chapter 2 illustrate that particular temporal structures interact with perception in a more complex way, than it can be coded within a system of metric accent. Thus, it is important to recall the difference between schematic or metric accentuation, and phenomenal accents which are raised by a rhythmic pattern.[219]

However, the main reason for an examination of metric weight and corresponding measures like metric indispensability lies on a more abstract level. Such measurement can also be the basis for a quantitative estimation of the degree of relationship between different meters. This issue is discussed in the following sections in terms of *metric coherence* (section 4.4.1) and *metric contrast* (section 4.4.2), and finally incorporated in the heuristics of metric malleability (see sections 5.2.5 and 5.3).

4.4 Relations between meters

When musical meter is developed and metamorphosed in the course of a piece, individual levels of awareness may vary according to different perspectives. For the performer, meter as a framework for musical communication and synchronization has to be well defined (see section 1.3). Listeners' metric construals can diverge (termed metric ambiguity throughout this thesis), but they can also deviate from the framework established for musical coordination among performers. BENADON's distinction of a

[218] This may be done by summing the indispensability values of the m rhythmic onsets of a pattern including n elementary pulses, and then subtracting this sum from the maximum sum for m onsets, which equals $\{n - 1 + ... + n - m\}$. The (normalized) result may be interpreted as an inverse of the requested probability. This idea is similar to "metrical expectedness (or simplicity)", as suggested by Toussaint, 2013, p. 283, for his measure of metrical complexity (cf. section 4.3.1).

[219] See also Povel, 1984, pp. 327 ff. for a precise and instructive discussion on the interplay of accent and "tempo" in regard to the choice of "metrical grids".

predefined *conceptual* tempo from several possible *emergent* tempos determines these differences from the perspective of tempo, as discussed in section 2.2.2. Moreover, the perceptual impact of metric development and change depends on the way how musically confronted meters are embodied in the rhythmic texture. A metric modulation can be so subtle that it is hardly recognized. In other cases, contrasting and disrupting effects can be the result. Regarding for instance the avant-garde of the 20th century, Elliott CARTER is known for the former,[220] whereas Igor STRAVINSKY is a representative of the latter.[221]

However, there may be abstract and formally definable distances between different meters, represented by aspects which can be detached from these contexts. To that effect, metric structures, as described in section 4.2, can be related to each other by a distance matrix. Among the relevant approaches, two types are distinguished in the following. First, relational and combinatorial spaces of musical meter have been constructed "analogous to tonal pitch space; and as with models of tonal pitch space, models of metrical spaces have traced a pattern of increasing complexity and sophistication."[222] For instance, COHN developed a two-dimensional space to represent relations between simple meters which reflect their distances in terms of hierarchic structure.[223] More recently, GOTHAM elaborated a comprehensive apparatus to assess formal relations among a wider set of meters, focusing on mixed meters (see section 4.2). Figure 4.12 exemplifies some of these relations which are inscribed in a correspondent three-dimensional metric space.

Meters are indicated by the numbers of elementary pulses each beat includes (see section 4.2.1). The number of beats per metric cycle (beat cardinality B) rises from left to right, and the space is conceptualized to be further extended to the right to connect meters of higher cardinalities. Accordingly, the origin at the left end recursively bifurcates to the right to set out distance paths along which different meter types are related to each other. The space is circumscribed by the two lines starting from the origin; the upper (left from the perspective of the origin) represents a progression of simple meters with beats of duple subdivision, and the lower (right) displays the parallel progression of simple meters comprising beats of triple subdivision. In between, the space unfolds to combine the possible mixed meters by relating their beat structures (1) in regard to *unordered subset relations* (U) across the horizontal grid, and (2) as permutations of *vector-identical* (V) meters in the vertical.[224] The meter 2233 is picked to demonstrate

[220] Carter's well-known procedure has been described as *metric modulation* or *tempo modulation*, as for instance in Berry, 1976, pp. 307 ff., who analyzes a passage of *String Quartett No. 1* where pulse- and activity-tempo are temporarily disconnected to sublimate a progressive metric modulation (cf. section 2.3.1).

[221] cf. footnote 9 in section 3.1

[222] Gotham, 2015b, section 1.3

[223] Cohn, 2001. These distances are also expressed by the different paths which can be taken in a *ski-hill graph*, discussed in section 3.2.4. This graph was introduced in the same article to examine the formal framework of hemiola.

[224] The exponentially growing number of mixed meters with rising cardinality, already discussed in section 4.2.2, here appears to be derivative of the unfolding abundance of relations.

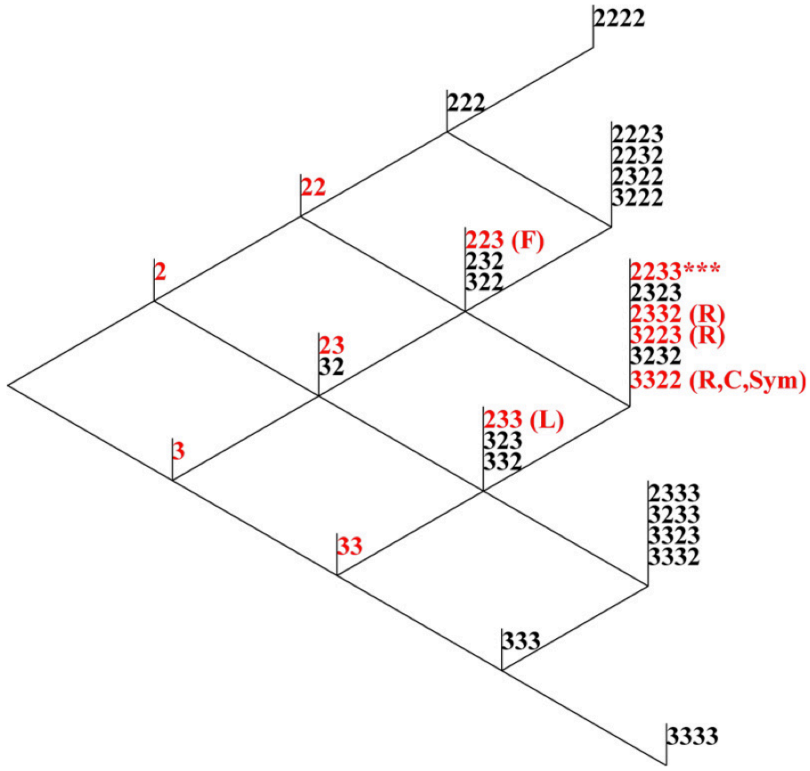


FIGURE 4.12: Gotham, 2015b, figure 26: a space for meters and metrical relations, highlighting meter 2233 and its relations to other meters which comprise up to four beats.

some particular relations to the closely linked meters marked in red. Among the meters along the same vertical, which hold vector-identity in relation to 2233, *rotational equivalence* (R, see section 4.2.2), *complementation* (C, "opposite beat types at every corresponding position"), and *symmetry* (Sym) occur.[225] The labeled meters of smaller beat cardinality are *ordered subsets* of the meter, that is, their beat structures are contained in the meter 2233 (O-relation).[226]

As in the present example, metric spaces specify networks of formal relations which can serve as an orientation for the exploration of possible perceptual and musical relations between meters. The latter may also be more directly accessed by another approach which is based on a quantitative model of metric accentuation (see section 4.3). When synchronized on the elementary pulse level, quantified accentual patterns of different meters can be compared pulse by pulse. In this way, measures of coherence or

[225]cf. Gotham, 2015b, sections 6 to 8. "Symmetry holds among pairs of meters for which the order of beats in one may be reversed to form the other [...]. All meters are symmetrical to either one other meter (pairs such as 23 and 32) or to none (internally symmetrical meters such as 2332)." (section 8.1)

[226]cf. Gotham, 2015b, section 8.1: "Three types of O relation are distinguished according to the position of the shorter meter within the longer—at the start (F), end (L), or within (W). For instance, 23 is L-related to 2223."

contrast can be developed, which can similarly establish a basis for distance networks. This concept is further amplified in the following sections 4.4.1 and 4.4.2.

Finally, meter relations and meter distances can be regarded as an important aspect of metric malleability. When a rhythm is reinterpreted according to an alternative metric framework, the relation between the involved meters may have impact on the way the rhythm seems to be metamorphosed. Thus, a rhythm appears to be more malleable, that is, it is transformed in a more contrasting way, if the distance of the metric alternatives is far. This assumption is an essential part of the heuristics of metric malleability, put forward in chapter 5 (see section 5.2.5 in particular).

4.4.1 Metric coherence

BARLOW developed the measure of *metric coherence* to quantify the similarity, or the accentual correspondence, of two meters.[227] It is based on a serial comparison of the associated *metric indispensabilities* (see section 4.3.2). The series are aligned so that temporally coinciding metric categories of the two meters match. That is, the method of calculation requires a common elementary pulse. Depending on their relative cycle lengths or pulse cardinalities, the superposition of the two accentual patterns can involve increasing levels of relational complexity. Therefore, a simple example of the calculation, involving the metric cycles 2 x 3 (33) and 3 x 2 (222) which are of equal length, may first illustrate the basic method. Initially, their correspondent indispensability series are computed and aligned as follows.

2 x 3: 5 0 2 4 1 3
3 x 2: 5 0 3 1 4 2

The next step is to "multiply the *relative indispensability* (nominal value divided by the maximum value, i.e. by the number of pulses minus one) of each elemental pulse of one of the metres by the relative indispensability of the concurrent pulse of the other metre,"[228] and then squaring the products. This is exemplified in table 4.7.

The final measure of metric coherence turns out to be basically a calibration of the average product-square.[229] Higher values of relative indispensability contribute most effectively to the magnitude of the average product-square when they coincide in both

[227]Barlow, 2012, Part I, p. 46. As already noted in section 4.3.1, an alternative notion of metric coherence was brought forward by Volk (cf. Fleischer, 2002, Volk, 2008). First, the *inner metric structure* of a musical piece (as opposed to the *outer*, notated meter) is derived by a metric weighting system based on actually occurring rhythmic onsets. Then, "metric coherence describes the correspondences of varying degrees between the outer and inner metric structure" (Fleischer, 2002, abstract). This concept is closely related to metric ambiguity, as it "can be described as divergence between inner and outer metric structure." (ibid.)
[228]ibid.
[229]ibid. See also Barlow, 2012, Part II, pp. 35 f. The average product-square is calculated by the sum of all product-squares, divided by the number of product-squares or elementary pulses.

TABLE 4.7: basic calculations for *metric coherence* between meters 2 x 3 and 3 x 2

pulse	relative indisp. (2 x 3)	relative indisp. (3 x 2)	product	squared product
0	1	1	1	1
1	0	0	0	0
2	0.4	0.6	0.24	0.0576
3	0.8	0.2	0.16	0,0256
4	0.2	0.8	0.16	0,0256
5	0.6	0.4	0,24	0.0576

aligned series. Thus, metric coherence corresponds to temporal coincidences of metrically accented categories, or metric positions which are assumed to have a high metrical weight.

When the metric cycles to be related have different pulse cardinalities m and n, that is, when they contain different numbers of elementary pulses, the method suggests to first construct a hypermetric cycle of pulse cardinality $lcm(m, n)$[230] in order to ensure the embedment of full cycles of each meter in the combined cycle. Accordingly, a comparison of, for instance, a 3 x 2 meter with a 2 x 2 x 2 meter would take into account a cycle of $lcm(6, 8) = 24$ elementary pulses. Thus, to compute the product squares of the relative indispensability values for these 24 pulses, the combined cycle embeds four cycles of the 3 x 2 meter and three cycles of the 2 x 2 x 2 meter.

Another type of temporal relation occurs when metric cycles of equal duration include a different number of elementary pulses. In this case, the computation of metric coherence can be achieved after the stratification of the involved meters is extended below the elementary level by additional divisors. If the 3 x 2 meter and the 2 x 2 x 2 meter taken from the example above would have the same cycle duration, a lower/faster elementary pulse level is shared by their extensions 3 x 2 x 2 x 2 and 2 x 2 x 2 x 3. Again, the combined cycle contains 24 elementary pulses. This type of temporal relation may be musically described as *polymeter*[231] or *polytempo*, that is, the original elementary pulses of the meters have different durations. The increase of temporal resolution in search for a common minimal pulse may be only theoretic, as this resulting pulse may be too short to serve as a metric category. In other words, it may fall beneath the *metric floor* of perceptually discriminable units (see section 2.2.2).

As mentioned before, the non-consideration of temporal constraints on pulse sensation is a common flaw of many models of metrical accent, including the measures of metric indispensability and metric coherence. Another shortcoming of the latter is that the mentioned calibration does not properly work when a meter is related to a temporal shift of itself or to shifted versions of other meters. There may be possible adjustments to allow for the calculation of such cases, as they correspond to type B dissonance

[230] The least common multiple of m and n.
[231] See sections 3.1.1 and 3.1.2.

(section 2.3.3), metric phase ambiguity, or the reinterpretation of rhythmic cycles in terms of the metric starting point (section 3.2.5). At any rate, the perceptual contrast of particular phase relations between changing meters should be considered an important aspect of metric malleability. It may for instance be less contrasting when a metric cycle is perceptually shifted by half of its length than by a more complex ratio. That is, the coherence of the 2 x 2 x 2 meter to itself, shifted by four elementary pulses, would be high, as the two strongest accent categories, the cycle downbeat and the half-cycle downbeat would coincide. In contrast, the coherence would be much lower if the meter is displaced by only one elementary pulse. There would be no coincidences of metric accents, as they fall in between each other. From the perspective of pulse perception, the latter change would accordingly be more disruptive when a tactus had been established which coincided with a metric level above the elementary pulse.[232] If the tactus would have grouped two elementary pulses, it would be preserved by a metric shift by two or four pulses but not by a shift by one pulse. More concrete musical examples of the perceptual impact of meter changes are for instance described by OSBORN.[233] He concluded from an analysis of meter changes in rock songs that the degree of disruption depends on the relation of the primary pulse and a *pivot pulse*, that is, "the slowest pulse stream preserved in a given meter change."[234]

In sum, these issues emphasize the requirement for a coherent quantification of metric relations, more precise, for an estimate of perceptually relevant distances between meters, including their relative phases and reflecting the temporal constraints on pulse sensation. In the following section 4.4.2, an alternative measure of *metric contrast* is therefore suggested which accounts for the mentioned aspects in order to support a heuristics of metric malleability.

4.4.2 Metric contrast

Before suggesting a definition of metric contrast, it is instructive to register and distinguish "outer" and "inner" factors for the perceptual evaluation of meter relations. In section 2.2.1, meter has been characterized by the cognitive processes of entrainment and projection, which are activated by rhythmic input. Meter can therefore be regarded as an abstract mental template, embodied by rhythm, though, implying formal constraints (see section 4.2). Accordingly, it was mentioned before (section 4.4) that perceptual judgements about similarity or distance between meters depend on the way they are conveyed by rhythm. This specific interaction is finally explored in sections 5.2.5 and 5.3 in the context of metric malleability. In order to prepare the necessary and appropriate tools for this study, metric contrast is introduced as a formal

232 For a correspondent example see sequence e in figure 2.13 (section 2.3.3).
233 Osborn, 2010
234 Osborn, 2010, p. 48

distance measure for hierarchic metric frameworks. It is distinguished from metric coherence by two features. First, it is based on the assumption that two specific aspects of metric contrast can be estimated independently: metric *period contrast* and metric *phase contrast*, derived from the correspondent types of metric ambiguity (sections 3.2.4 and 3.2.5) and metric dissonance (section 2.3.3). Second, while metric coherence is a measure of similarity, metric contrast is inversely oriented to dissimilarity. This facilitates the formulation of the above-mentioned assumption that the metric malleability of a rhythmic cycle increases with the amenability to be heard in contrasting metric frameworks. More precisely, when a rhythm likewise fits to different frameworks, the perceptual contrast of a metric modulation or metric gestalt flip of the rhythm from one to the other depends on the metric contrast between the involved meters. This heuristic aspect of metric malleability is further discussed in section 5.2.5.

According to the following definition, metric contrast corresponds to the amount of conflict or asynchronism between higher-level pulses of different meters, which are simultaneously or successively aligned by means of a common elementary pulse. These types of metric conflict have also been distinguished as simultaneous or successive metric dissonance by KREBS (section 2.3.3). Informally, the algorithm may be summarized as follows. An overall value of metric contrast is estimated as the sum of individual contrast values between the different pulses or metric strata of two hierarchic frameworks. To account for different numbers of summands, an average contrast is calculated. The procedure is tempo sensitive (see above) but the tactus or main pulse is not predefined.

As outlined above, metric contrast is proposed as a combined measure of metric period contrast and metric phase contrast. These components can be separately estimated to differentiate these aspects for analytical clarity. They may be included with individual weights in an overall measure of metric contrast. The calculation method is parallel in respect to both components, and may be first overviewed as follows:

- let A and B be stratified meters according to the definitions in section 4.2
- if A and B have different pulse cardinalities m and n, a common cycle of pulse cardinality $LCM(m, n)$ is examined (equal to the algorithm for metric coherence, see the previous section 4.4.1)
- every stratum a of A is compared to every stratum b of B
- the calculation of metric contrast between two strata (pulses) a and b is based on the number of temporal intersections of coinciding stratum elements or temporal categories. Thus, to compare two meters, a common elementary-pulse rate is specified

- the weighting of metric contrast between a single pair $\{a, b\}$ relative to the overall metric contrast between A and B depends on the product of the pulse-period saliences of a and b (see equation 2.1 in section 2.2.3)

Let us now continue to specify the procedure in regard to both types of metric contrast. For two metric strata a and b, synchronized at elementary pulse 0 (initial intersection), their period contrast $C'_\delta(a, b)$ determines a value for type A dissonance (section 2.3.3) – the deviation from metric consonance or mutual embedment – between a and b. It can be related to

1. the sum of intersections or temporal coincidences $\sigma_i(a, b)$ of pulse categories in a and pulse categories in b within one common cycle, and
2. the lower sum of pulses σ_p within one common cycle, either of a or b, determined by $min(\sigma_p(a), \sigma_p(b))$,

as defined in equation 4.1.

$$C'_\delta(a, b) = log(\frac{min(\sigma_p(a), \sigma_p(b))}{\sigma_i(a, b)}) \tag{4.1}$$

Then, the overall period contrast $C_\delta(A, B)$ between two stratified meters A and B can be estimated by a summation of $C'_\delta(a, b)$ of all strata combinations $\{A_a, B_b\}$, weighted by the product of pulse-period saliences S_p of A_a and B_b (equation 4.2).

$$C_\delta(A, B) = \sum_{a=1}^{\sigma_s(A)} \sum_{b=1}^{\sigma_s(B)} C'_\delta(A_a, B_b) S_p(A_a) S_p(B_b), \tag{4.2}$$

where $\sigma_s(M)$ is the sum of strata in a meter M. The weighting reflects the prototypical perceptual saliences of particular metric strata, dependent on their pulse rates (see section 2.2.3). As it is an abstracted estimate, no phenomenal accents are involved in this heuristics,[235] and therefore, perceptual relevance can only be related to pulse-period salience. It is based on the assumption that *metric contrast is the more salient, the more salient the involved pulse streams* are. If only one of two contrasting pulses is perceptually salient, their contrast may hardly be recognized.

Metric phase contrast between two metric strata a and b can simply be defined as a binary decision: $C'_\varphi(a, b) = 1$ if $\sigma_i(a, b) = 0$ (otherwise, $C'_\varphi(a, b) = 0$ if $\sigma_i(a, b) > 0$). That is, metric strata can be either in or out of phase.[236] Therefore, an interval of an

[235] As in the notion of pulse-match salience (cf. sections 2.2.3 and 5.2.1).

[236] The number of phase constellations between two isochronous pulse streams, the periods of which can be expressed as integral multiples of an elementary unit, is equal to the greatest common divisor of the two periods (cf. Parncutt, 1994, p. 440). Accordingly, two isochronous metric strata with relatively prime period ratios, as for instance 3:2, 5:3, or 7:4, can never be out of phase, and thus, their metric phase contrast is always 0. (This may also hold for some combinations of metric strata which involve mixed pulse durations.)

integral number of elementary pulses has to be defined in advance which expresses the temporal shift between the two meters in terms of their initial pulses. Then, metric phase contrast $C_\varphi(A,B)$ between two stratified meters A and B can be defined parallel to metric period contrast (equation 4.3).

$$C_\varphi(A,B) = \sum_{a=1}^{\sigma_s(A)} \sum_{b=1}^{\sigma_s(B)} C'_\varphi(A_a, B_b) S_p(A_a) S_p(B_b) \tag{4.3}$$

The differentiation of metric contrast into period and phase aspects also reflects the differentiation of musical concepts. The examples of hemiola, discussed in section 3.2.4, are related to metric period contrast. Shifts of metric pulses on different levels, for instance of the cycle downbeat in *Gahu* drum music, or of a hypermetric pulse in Western classical music, correspond to metric phase contrast (see section 3.2.5). These aspects of metric contrast can be incorporated as part of a heuristics of metric malleability, as proposed in section 5.2.5. The level of differentiation is further enhanced by the reflection of pulse rates, though, further studies are needed to qualify the amount of influence on metric contrast by rhythmic presentation rates.

Metric contrast will be employed in the model of metric malleability, as discussed in section 5.2.5. Furthermore, a selective comparison of the measures of metric period contrast and metric coherence is found in appendix D. It involves fife simple, hierarchic meters (2 x 2 x 3, 2 x 3 x 2, 3 x 2 x 2, 3 x 3, and 2 x 2 x 2), yielding ten meter relations for which calculations are applied. The correspondent data are displayed to provide additional orientation regarding these measures.

Chapter 5

A quantitative heuristics of metric malleability

In this chapter, a quantitative heuristics of metric malleability is tackled, based on the information which is collected, developed, and evaluated so far. As it was discussed from several perspectives, the metric malleability of a rhythmic pattern results from the cognitive flexibility to interpret this pattern according to different metric frameworks. Also, it was set out that metric malleability is a feature that may inhere in a rhythm but usually is not recognized. Rather, it is quickly resolved by spontaneous metric interpretation, implying a perceptual framework which finalizes the rhythmic gestalt. To that effect, many studies have concentrated on the identification of "criteria that might lead a listener to favor a particular rhythmic interpretation".[1] Computational models of metric interpretation however follow divergent paradigms to simulate those cognitive decision processes. From the perspective of this study, a quantitative approach to metric malleability needs to employ a computational model of metric interpretation, which is able to suggest a preference ranking of plausible metric frameworks for a given rhythm. In the following section 5.1, such models are therefore evaluated, concerning their flexibility and openness for alternative interpretations. This is important, to reflect the space of metric ambiguity involving the potential for metric malleability. Finally, in section 5.2, a model of metric malleability is advanced by the development of heuristic tools which are designed to analyze specific features of cyclic rhythmic patterns with respect to metric malleability.

5.1 Modeling metric interpretation of rhythm

According to FLANAGAN, metric ambiguity can be modeled "as the degree of variety of listener responses to a given rhythm."[2] This notion is derived from experimental

[1] Longuet-Higgins and Lee, 1984, p. 425. The phrase "rhythmic interpretation" as used by Longuet-Higgins and Lee, 1984 is synonymous to what is termed "metric interpretation" in this study.

[2] Flanagan, 2008, p. 635

investigations in tactus responses,[3] as described in section 2.2.3, and emphasizes proportionality between the variety of such responses and the amount of metric ambiguity. The approach put forward in section 5.2 correspondingly includes an account of metric ambiguity as the "dispersion of probabilities over candidate meters",[4] which is then treated as an important aspect of the quantitative estimate of metric malleability. To realize this, I follow FLANAGAN's step to employ "a model of meter induction that, for a given rhythm, produces numerical values for an array of potential meters".[5]

In view of this methodological design, a categorization of existing quantitative models of metric interpretation will show in advance, which features of such models will be advantageous for an implementation in our context. Chapter 2 provides a comprehensive survey of findings about the perceptual and cognitive bases of metric entrainment. Advanced theoretical and computational models of *beat induction* or *beat tracking* apply that knowledge in order "to predict the period and phase of the psychological pulse series (the beat) of a complex acoustic stimulus"[6] or a symbolically categorized musical sequence. The cognitive process of meter identification "has been the subject of a huge amount of research, both experimental and computational".[7] A comprehensive survey of computational models of beat or meter induction and their experimental groundings would thus go beyond the scope of this thesis. Given the array of models involving various assumptions, conditions, and types of calculus, confirmed by different empiric studies, it is reasonable to restrict the discussion to paradigms which yield viable approaches to the ambiguity implied in emergent processes of metric interpretation. Moreover, our interest is limited to models dealing with rhythmic input which conforms to the abstraction level discussed in chapter 4 and section 2.3.4. Thus, both audio and performance data,[8] a meanwhile common input for contemporary models,[9] are not taken into account.

5.1.1 The problem of evaluating induction models

The majority of the models in question are designed in order to compute a single solution, that is, to find the most plausible metric interpretation of a musical sequence. Studies undertaking systematic evaluations of these models similarly tend to predefine an ideal output to establish a principal criterion for a model's performance quality.[10] From the perspective of metric ambiguity, the problem arises that several meters could be plausible solutions for the predictive task simulated by the models. In other words,

[3] For instance, of Parncutt, 1994 and Povel and Essens, 1985
[4] Flanagan, 2008, p. 640
[5] Flanagan, 2008, p. 635
[6] Large, Fink, and Kelso, 2002, p. 15
[7] Temperley, 2010, p. 373
[8] For instance from MIDI recordings.
[9] See for instance Sethares, 2007 for an illustration of many of the techniques developed for this purpose.
[10] For instance Desain and Honing, 1999, andTemperley, 2004.

it may be an inappropriate simplification of an evaluation method to identify a reference solution in advance of an assessment in the sense above. TEMPERLEY discusses this problem in order to define a measure for a model's efficiency.

> Whatever the goals and assumptions of a metrical model, an important and obvious question to ask is, "How good is it?" That is, what percentage of the time does it actually produce the correct result? (The "correct result" can be defined as the metrical structure inferred by competent listeners. There might, of course, be some subjective differences among listeners; one might also take the music notation for the piece to represent its meter. But in most cases, I would argue, there will be agreement among these sources.)[11]

I would instead object that the basic occurrence of metric ambiguity, demonstrated by empirical findings like those mentioned above, prohibits premises like that. Two issues are pushed aside here. First, the dilemma of meter induction models is to simulate a subjective process which makes an objective and systematic evaluation difficult. Second, the general presence of ambiguity as a consequence of subjectivity is thus ignored by the assumption that there exists a "correct" result.

The evaluation study by DESAIN and HONING on rule-based beat induction models[12] differentiates the reference problem according to different sets of test corpora. Three types of temporal patterns are used as input data for the models to compare. For a corpus of rhythmic patterns derived from national anthems,[13] DESAIN and HONING compare the output of the models to the notated meters in the source, a common approach, as already mentioned above by TEMPERLEY. Two other types of input rhythms are derived from a combinatorially complete set of patterns: a sample of patterns, collected by Monte Carlo simulation and therefore evenly distributed within that combinatorial space, and a "strictly metrical" subset.

> These patterns have a simple metrical interpretation in which each durational interval fits one level of a metrical hierarchy directly. The patterns are strictly metrical in the sense that there are no syncopations or tied notes. Note that they can still be ambiguous – some patterns can be generated from different meters.[14]

The last sentence quoted has special significance from our perspective. The mentioned fact is indeed symptomatic for the problem we just examine. Consequently, DESAIN and HONING suggest that for "the set of strictly metrical patterns, correct beats can be defined to be those that fit one of the metrical levels of one of the generating meters."[15]

[11] Temperley, 2004, p. 28
[12] Desain and Honing, 1999
[13] Desain and Honing, 1999 refer to Shaw, M. and H. Coleman (1960). *National Anthems of the World*. London: Pitman.
[14] Desain and Honing, 1999, pp. 32 f.
[15] Desain and Honing, 1999, p. 36

The tolerance inherent in this definition of correctness allows for two types of metric ambiguity. First, several metrical levels are accepted as correct beat levels or tactuses (corresponding to reference-level ambiguity, see section 3.2.3), and second, several meters (or types of meter) are accepted as a correct choice (corresponding to metric period- and phase ambiguity, see sections 3.2.4 and 3.2.5).

In sum, it is generally questionable under which conditions a reference can be determined. For the purposes of this study, the idea of a reference is meaningless and any *a priori* judgement or metric preference will therefore be avoided. An ecologically and perceptually valid measure of correctness or performance quality for beat-induction models may only be achieved by comparing the results of a model's output to distributions of human judgements. The study by PARNCUTT[16] follows this approach which is one of several reasons to choose this model as a basis for the model, presented in section 5.2, where further arguments for this choice will be discussed as well.

> Some metrical analysis systems are clearly intended to model human cognition; others are simply designed to solve the practical problem of meter-finding in whatever way seems most effective.[17]

In other words, it is for practical reasons that many models are restricted to produce only one suggestion. In contrast, models of human cognition, dealing with subjective, emergent processes and their inter-individual differences, have to account for the ambiguity, inherent in metric interpretation. Chapters 2 and 3 thus anticipate the problems, cognitive models have to cope with. Due to the examined properties and ambiguities of rhythm perception and cognition, the models follow divergent paradigms and thus highlight different aspects of the cognitive process.

> Beat induction is a fast process. Only after a few notes (5-10) a strong sense of beat can be induced (a "bottom-up" process). Once a beat is induced by the incoming material it sets up a persistent mental framework that guides the perception of new incoming material (a "top-down" process). [...] However, this top-down processing is not rigidly adhering to a once established beat-percept, because, when in a change of meter the evidence for the old percept becomes too meager, a new beat interpretation is induced. This duality, where a model needs to be able to infer a beat from scratch, but also to let an already induced beat percept guide the organization of more incoming material, is hard to model.[18]

[16]Parncutt, 1994, see section 2.2.3
[17]Temperley, 2004, p. 28
[18]Desain and Honing, 1999, pp. 29 f. See also section 2.3.3 about metric priming, and Honing, Bouwer, and Háden, 2014, p. 307: "The perception of a beat is a bi-directional process: not only can a varying musical rhythm induce a regular beat, a regular beat can also influence the perception of the very same rhythm that induces it."

DESAIN and HONING here summarize many of the aspects discussed before. They assume that these considerable challenges led to a "wide variety of computational formalisms that have been used to capture the process."[19] Rule-based approaches,[20] probabilistic algorithms,[21] neural-networks,[22] and systems of coupled, adaptive oscillators[23] belong to the frequently used types. The following section 5.1.2 provides a selective discussion of some of these approaches and related aspects which are relevant in the context of this study.

5.1.2 Model types

As found in the previous section 5.1.1, it is a different task to model the cognitive process of metric interpretation instead of developing a meter-finding algorithm which works for practical purposes. Correspondingly, the goals of cognitive models include further aspects than just technical solutions for real-time beat tracking or the transcription of performance data to a musical score. The latter demands a decisive attitude: the problem has to be solved by a definite solution, that is, by the definition of an unequivocal meter or beat. In contrast, cognitive models have to account for multiple cognitive aspects of metric interpretation, as discussed in chapter 2, that is, expectation, entrainment, short term memory, internalized metric schemes in the long-term memory, and so forth. In this context, metric ambiguity plays an essential role, though, the area of research on inter-individual differences is generally understudied.[24] In the same context, it is also important to recall that a perceived rhythm does not always evoke a metric framework of regular beats. Instead, as discussed in section 2.3, it may only afford figural grouping based on temporal proximity between rhythmic events.[25]

> With respect to rhythm, there are numerous anecdotal reports that suggest that there are large individual differences in the ability to perceive a beat; that is, some people appear to have much more difficulty perceiving a beat than others.[26]

In section 2.2.1, entrainment theories were compared to interval theories of temporal perception. The latter are grounded in "an information-processing framework. Formal models developed within this framework typically posit distinct clock, memory, and decision stages of temporal processing."[27] For the purpose of simulating beat induction, the concept of an internal clock (as the cognitive representation of a beat) is

[19] ibid.
[20] for instance Longuet-Higgins and Lee, 1984, Eck, 2001
[21] for instance Temperley, 2007
[22] cf. for instance Böck, Krebs, and Widmer, 2016 and Krebs et al., 2016. For a combined approach including (Bayesian) probability and neural networks see Holzapfel and Grill, 2016.
[23] See for instance Sethares, 2007, pp. 147 ff. and pp. 199 ff., for an instructive overview.
[24] cf. McAuley, 2010, p. 193
[25] cf. for instance Essens and Povel, 1985, Parncutt, 1994
[26] McAuley, 2010, p. 193
[27] McAuley, 2010, p. 169

proposed as the basis of rhythm processing, as in the model of Povel and Essens.[28] More precise, possible clocks are ranked by summing up negative evidence. The less complex, a rhythm can be coded within a clock framework, the more likely this clock may be chosen as the metric framework for this rhythm. Other related approaches are in contrast based on positive evidence, reducing some flaws of the original method.[29] These rule-based models work well only on a symbolic level. Meter tracking models based on self-sustaining oscillators, and grounded in the mentioned entrainment theories, provide a better approach to (subsymbolic) beat tracking from audio and performance data as they are "able to cope with a certain amount of expressive deviation in the time domain."[30] They also claim to correlate with biological reality, that is, "the peaks in oscillator amplitude in models of musical entrainment are assumed to represent periodic changes in gross neural activity".[31] The success of oscillator models in subsymbolic contexts and real-time beat tracking is based on flexible entrainment. Periods and phases of the oscillators can adapt to a certain extend to temporal variations in the input signal.[32]

Lerdahl and Jackendoff's meter theory, established in the *generative theory of tonal music* (GTTM), and including the collection of *metric preference rules*[33] mentioned before (section 2.3), has had considerable impact on subsequent rule-based approaches to model meter induction.[34] However, it also received criticism for that it "is based on intuition and music theory principles that were not supported by psychological experimental data".[35] Nevertheless, Palmer and Krumhansl[36] experimentally proved that for Western common-practice music, "the strength of a pulse location" in a metric hierarchy according to Lerdahl and Jackendoff "correlates well with the degree of the expectancy of occurrence of an onset at that particular location".[37] This is in line with the assumption that the establishment of a hierarchical metric framework is influenced by expectations generated by accumulated musical knowledge. Models which are designed to simulate meter induction, specifying several metric levels, may therefore behave too specific. Indeed, "a second criticism of GTTM has been that it is a theory applicable only to Western tonal music, and that its claims of universality have not been supported by intercultural research".[38] Thus, regarding the universality

[28] Povel and Essens, 1985

[29] Eck, 2001, and, to a certain extend Parncutt, 1994.

[30] Noorden and Moelants, 1999, p. 44

[31] McAuley, 2010, P. 170

[32] cf. footnote 23 above

[33] Lerdahl and Jackendoff, 1983, pp. 74 ff. The metric preference rules complete the whole set of rules concerning grouping structure and metric structure. Both aspects are treated with a collection of well-formedness rules and preference rules (pp. 37 ff.)

[34] For instance Lee, 1991

[35] Toussaint, 2015, p. 3, referring to Hansen, Niels C. (2011). "The Legacy of Lerdahl and Jackendoff's *A Generative Theory of Tonal Music*: Bridging a Significant Event in the History of Music Theory and Recent Developments in Cognitive Music Research"." In: *Danish Yearbook of Musicology* 38, pp. 33–55.

[36] Palmer and Krumhansl, 1990

[37] Toussaint, 2015, p. 3

[38] ibid., again referring to Hansen (2011, see footnote 35). See also section 3.2.1.

of cognitive models, deviating scientific perspectives on the perceptual and cognitive bases of meter have to be considered. In sum, sensory cues may guide initial beat induction, but the establishment of a hierarchic metric framework largely depends on internalized metric schemes which develop individually with musical education and exposure.

To cope with this influence of long-term memory and schematic expectations about meter, some tracking mechanisms implement probabilistic methods. "Models of statistical periodicity do not presume that the signal itself is periodic; rather, they assume that there is a periodicity in the underlying statistical distributions."[39] In this context, BAYES' rule has become an influential computational theorem, employed for instance by TEMPERLEY, to infer meter "from the probabilities of different patterns of regularity generated by a given rhythmic input."[40] According to this approach, first the probabilities of rhythms (onset patterns) for a given meter, and a prior probability of that meter are calculated. Then, the probabilities of all possible meters of a given rhythm are ranked according to the product of both values.[41] Statistical data for the "memory" of the system (the prior probabilities) may be derived from music corpora or likewise a set of metric preference rules. The calculation of meter probabilities can then be interpreted as an optimization process according to these rules or the statistical regularities, derived from existing music.

Generally, rule-based models operating on temporal intervals derived from a rhythm and dynamical-systems accounts based on the entrainment of oscillators, "are similarly limited in their reliance primarily on temporal accents to infer beat and meter."[42] Rule-based models often rely on temporal accentuation principles derived from empirical studies (see section 2.3.2). For instance, POVEL and ESSENS[43] employ three assumptions about the occurrence of temporal accents.

> First, when an onset is isolated relative to other onsets, it sounds like an accent. Second, when two onsets are grouped together, the second onset sounds accented. Finally, for groups of three or more onsets, the first and/or last tone of the group will be perceived as an accent.[44]

[39] Sethares, 2007, p. 175

[40] Vuust and Witek, 2014, p. 3. "Bayesian approaches can also be seen as the basis of perceptual processing more generally, from the level of individual neurons, to subjective affective experience." (ibid.), see also Sethares, 2007, pp. 175 ff. and Temperley, 2007

[41] Temperley, 2010, referring to Temperley, 2007, summarizes: "In probabilistic terms, the meter-finding problem can be defined as the problem of determining the most probable metrical structure given a pattern of notes - Bayes' Rule then tells us that
(meter | note pattern) $\propto P$(note pattern | meter) $\times P$(meter)
The meter that maximizes the right side of this expression will be the most probable meter given the note pattern" (p. 373).

[42] McAuley, 2010, p. 191

[43] Povel and Essens, 1985

[44] Honing, Bouwer, and Háden, 2014, p. 308

Those subjective durational accents are thought of as cues for beat and meter, like the often discussed Western norm that longer notes appear on strong metrical positions (see section 4.3.2, footnote 204). If this rule is broken, for that a long note is initiated on a weak position of a previously established meter, one would probably speak about syncopation in the context of Western common practice. Though, whereas such norms are conditioned by experience and musical culture, as mentioned above and in section 4.3.1, temporal accents may be a more general, low-level condition of rhythm perception (see section 2.3.2).

Beyond this involuntary level of accent perception, induction models are fundamentally challenged by high-level cognition. Once an accent pattern is processed, the projection of a metric framework is open to influences from consciousness,[45] that is, cognitive top-down processes which depend on attention, or which are voluntarily controlled by interest (see sections 2.2.1 and 2.3.1). Therefore, a model should not be considered as representing individual psychological reality, but merely as a formalism or a quantitative heuristic of general tendencies.

5.1.3 Induction models and aspects of metric ambiguity

The previous sections concern different concepts for models of metric interpretation (5.1.2) and the challenge of their evaluation due to metric ambiguity (5.1.1). According to this perspective, the crucial distinctions among the range of approaches are summarized by means of the following six items. The involved aspects are related to the aspects of rhythm perception discussed so far.

1. Some models merely process interval ratios derived from a rhythm,[46] whereas others also take absolute interval durations into account.[47] Thus, models are either sensitive or ignorant about the temporal parameters, which demonstrably influence rhythm perception and cognition (section 2.2.2). LONDON's notion of *tempo-metrical types* (section 3.2) accounts for the fact that a metric hierarchy integrates several metric strata with different perceptual implications: the measure (or cycle level), the tactus (or beat level), the subdivision level, and so forth. Consider for example the frequency or period duration of the tactus, the "counting" level which carries subjective tempo. If it is changed, other levels are affected and may change their function in the metric construal, for instance, another level may then represent the tactus. From the perspective of this study, this is addressed as *reference-level ambiguity* in section 3.2.3. At very fast tempi, for instance,

[45] As further discussed in the context of subjective grouping in section 2.3.1, Honing, Bouwer, and Háden, 2014 note that "the perception of a beat can also be guided by conscious effort. By consciously adjusting the phase or period of the regularity we perceive, we can influence which tones we hear on the beat." (p. 309)

[46] Povel and Essens, 1985

[47] Lee, 1991, Parncutt, 1994

the measure level can become the tactus level, as in the second movement of BEETHOVEN's Ninth Symphony.[48] In other words, the interplay of temporal accents gets modified and may favor another metric interpretation. This is also an important factor of metric ambiguity as it influences the scope of possible metric interpretations. Reinterpreting the statement by HANDEL and OSHINSKY cited at the beginning of section 2.2.2, that "rhythm emerges at a specific tempo"[49], it may be distinguished that meter emerges at a specific elementary-pulse rate. The concept of *preferred tempo* has been incorporated in the model of PARNCUTT[50] as one determinant of pulse-period salience. Van NOORDEN and MOELANTS refined this approach by the formulation of a resonance curve for beat perception in regard to tempo. They compared different empirical data sets of tapping responses to find appropriate parameter settings for their curve. In other words, these models define parametric filter functions for preferred tempo regions which are adjusted to empirical data (section 2.2.3).[51] GOTHAM studied "tempo attractors" for meter and shows that these filters can be successfully employed to model optimal tempi for specific metric types.[52] Relating to LONDON's tempo-metrical types, the interaction of (subjective) tempo with the temporal thresholds of rhythmic perception[53] leads specific meters to "gravitate" towards attractor tempi, which optimize the sum of pulse-period saliences across metric levels.[54]

2. Beat-induction models are generally designed to derive categorically isochronous beats from a rhythmic input. In contrast, some rhythmic patterns particularly afford non-isochronous or mixed metric beats, and some listeners are trained to generate corresponding metric expectations (see section 3.2.1). To my knowledge, these possibilities have not yet been consistently represented by induction models. The variety of metric frameworks, augmented by this means, naturally increases the space of metric ambiguity and makes predictive modeling even more difficult. The present study explicitly reflects this issue and integrates mixed-meter perspectives for meter induction (section 5.2.3).

3. Metric ambiguity and malleability also get apparent by the fact that different

[48] cf. London, 2012, p. 72. By means of this symphony, London, 2012 also demonstrates that a musical passage played in different tempi in fact evokes different perceived meters (cf. example 4.3 on p. 70). Thus, metric interpretation is guided by tempo, in that we "are looking for periodicities in particular temporal ranges, and we require that they relate to each other in certain ways." (p. 69)

[49] Handel and Oshinsky, 1981, p. 9

[50] Parncutt, 1994

[51] See also Moelants and McKinney, 2004

[52] Gotham, 2015a

[53] For tempo-metrical types, see section 3.2, and for temporal thresholds of rhythmic perception see section 2.2.2. See also London, 2012, pp. 95 f.

[54] See the notion of *aggregate pulse salience*, discussed in sections 2.3.1 and 5.2.3.

beat-induction models can yield different suggestions respective to the same input.[55] This suggests to specify meter induction as an emergent cognitive resolution of latent metric ambiguity and malleability, modeled as the potential of alternative perceptual scenarios and their relative probabilities. A rhythm may also raise a bistable perception, which is described in the context of polyrhythm in section 3.1.1 and, in terms of figure-ground ambiguity, in section 3.2.2. In such situations, metric malleability becomes sensible. To this effect, cognitive models are challenged to simulate this process, given that "rhythms are inherently ambiguous".[56] Another important aspect of metric interpretation is the temporal expectancy about immediate rhythmic development which arises in the course of rhythmic processing. In section 2.2.1, the formalization of basic expectancy by DESAIN is discussed as a consequence of a single perceived IOI. This model moreover defines *complex expectancy*, emerging from interval relations during rhythm perception, as the summation of basic expectancies within the perceptual present (see figure 5.1).

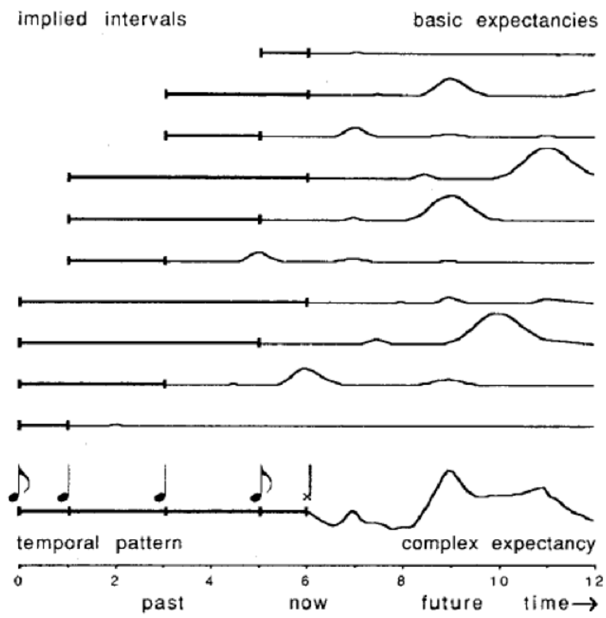


FIGURE 5.1: Complex expectancy emerging as a composition of basic expectancies induced by rhythmic IOIs (Desain, 1992, figure 2, p. 443).

The function of complex expectancy could be considered both as a beat-induction model and an indicator of metric ambiguity. Regular peaks may be interpreted as a clear perceived beat whereas a more complex structure obviously points to

[55] See section 5.1.1. Desain and Honing, 1999, p. 31, add that, accordingly, "different assertions about the state maintained during processing can be made for the different models."

[56] Rosenthal, 1989, p. 325, paraphrasing Longuet-Higgins and Lee, 1984, p. 424 (cf. section 3.1, footnote 11).

metric ambiguity, malleability or another form of complexity.[57]

4. Some models are restricted to find the main beat or tactus of a rhythmic sequence, whereas others suggest several levels of the assumed underlying meter. They may therefore be distinguished as beat-induction and meter-induction models. A metrical framework can simply consist of a tactus or main beat. More complex metric hierarchies are usually built up in relation to this central metric reference. Many induction models restrict themselves to predict the pulse which is most likely to be perceived as the tactus. Some models furthermore account for possible metric hierarchies emerging from the main pulse derived from a rhythmic input. As discussed in section 2.2.1, rhythmic textures dynamically embody a varying number of metric strata, correlating with a variable hierarchic depth of metric entrainment.[58] It may be assumed that the cognitive depth of a metric framework, that is, the number of metric strata which are recognized by metric interpretation, runs parallel to the hierarchic depth of those embodiments. Nevertheless, the extend of cognitive processing of metric strata above or below the tactus is difficult to prove empirically.[59] Put together with the argument that possible hierarchic depths of metric frameworks are culturally facilitated or constrained,[60] meter-induction models may be less reliable than beat-induction models. The taxonomy of metric ambiguity proposed in section 3.2, predominantly distinguishes types of tactus ambiguity, though, reference-level ambiguity (section 3.2.3) implies a hierarchical construal of meter. Thus, from our perspective, an appropriate model of metric interpretation should reflect the possible metric depths, which may be embodied by a rhythmic pattern.

5. Some approaches specifically model initial beat induction as a bottom-up process,[61] whereas others focus on metric top-down interpretation. The latter type is particularly appropriate for the holistic perception of cyclic rhythms.[62] In section 3.2.5, metric phase ambiguity is exemplified by the observation that the meter of a cyclic rhythm can be variably conceptualized by means of different *starting points*, also referred to as *regulative time points*[63] or metric *clasps*.[64] This type

[57] A comparable approach is the concept of *inner meter* (Fleischer, 2002, Volk, 2008), discussed in section 4.3.1 (see also section 4.4.1, footnote 227). Inner meter is similarly estimated by the sum of metric projections of single IOI in a rhythmic texture.

[58] London, 2012 distinguishes between static or abstract notions of meter, and meter as a dynamic system which runs through various modes of excitation or resonance with different hierarchical depths. This dynamics is similarly deduced from the concept of metric projection by Hasty, 1997. Section 4.3.1 discusses the correlation between "hierarchical persistence and depth" of metric entrainment, and metric weight or "metrical strength" (see footnote 146 in section 4.3.1).

[59] cf. for instance Vuust and Witek, 2014, p. 2, and Povel, 1984, p. 333.

[60] See section 3.2.1 (footnote 92) and section 3.2.3 (footnote 152).

[61] For instance Lee, 1991 and Longuet-Higgins and Lee, 1984.

[62] For instance Eck, 2001, Parncutt, 1994 and Povel and Essens, 1985.

[63] cf. Anku, 2000, Cuthbert, 2006

[64] cf. Yu, Getz, and Kubovy, 2015

of ambiguity may not be restricted to cyclic patterns. A metric framework, established for an arbitrary rhythm, is a cyclic conceptual structure for which a phase is defined by the temporal location of the strongest beat in the metric hierarchy (the cycle downbeat). For cyclic patterns, the empirically derived run and gap principles give a basic idea, how particular structures may correlate with metric phase ambiguity. Another hint may be derived from studies about the *robustness* of rhythmic cycles (section 4.1.3). Cyclic rhythms which occur in several metric rotations within different musical contexts may be more malleable in regard to phase, that is, they may be perceived according to different metric phases. Induction models which linearly process a rhythmic input – thus allowing real-time cognitive modeling – define metric phases by a bottom-up process.[65] Typically, the first attack is primarily assumed to represent a strong metrical position. If counter-evidence against this initial hypothesis arises with further rhythmic progression, the strong metrical accent or downbeat position is shifted to a later attack and the beginning is regarded as an anacrusis.[66] However, this method is basically inappropriate to identify loud rests on initial downbeat positions, as in figure 1.1.

6. A model either processes time-discrete or time-continuous rhythmic information, that is, it either requires an elementary pulse grid or is capable to simulate a temporal categorization process of continuous temporal intervals.[67] In the previous section 5.1.2, it is already mentioned that models based on the entrainment of oscillators, are able to adapt to time-continuous information. However, exploring the relation between categorial rhythm perception, metric priming, and metric entrainment is a complex matter, as surveyed in section 2.3.4. Thus, models of metric interpretation have to account for the fact, that the "question of the relation between the rhythmic-categorisation process and the metre-induction processes is still open."[68]

In sum, the list exemplifies the relevant features of induction models related to the previously discussed aspects of metric ambiguity. The approach developed in the following section 5.2, accounts for most of them by implementing the involved types of metric ambiguity (section 3.2) as aspects of metric malleability. Items 1 and 4 refer to reference-level ambiguity (section 3.2.3), item 2 alludes the problem of simple versus mixed metric frameworks (section 3.2.1), and item 5 corresponds to metric phase ambiguity (section 3.2.5). Finally, items 3 and 6 take a back seat, though, processes of metric

[65] cf for instance Lee, 1991, Longuet-Higgins and Lee, 1984

[66] Povel, 1984 assumes that the corresponding "process of perception may be seen as a hypothesis-testing activity in which the successive intervals in the sequence are used to construct temporal grids that are subsequently tested as a potential frame to specify the sequence under consideration. Only if either the location or the duration of a latter interval suggests a grid that appears to yield a more economical description will the latter grid be adopted". (pp. 326 f., referring to Steedman J. (1977). "The perception of musical rhythms and meter". In: *Perception* 6, pp. 555–569.)

[67] cf. for instance London, 2012, pp. 19 ff.

[68] Desain and Honing, 2003, p. 363, as already more extensively quoted in section 2.3.4 (footnote 344).

priming and categorical rhythm perception are implied in the abstraction level of the following examination.

5.2 Modeling the metric malleability of rhythmic cycles

In this section, some basic devices for the development of a heuristics of metric malleability are suggested and reviewed in a confined context. So far, the wider context of metric malleability has been examined, and the level of abstraction has been defined at which a quantitative model can be formulated.

A main insight of the psychological research reported during this study, is the correlation of metric entrainment and pulse sensation (see section 2.2). A perceived rhythm may evoke several pulse responses. Against this background, stratified meter can be conceptualized as an aggregation of embedded pulses, and metric ambiguity and malleability can be characterized as the listeners' opportunity to entrain to different pulses, featuring diverging periods and/or phases. The types of ambiguity which are distinguished in section 3.2, are correspondingly related to these two categorical properties.

To quantify the metric malleability of rhythmic cycles, a computational method to estimate the plausibility of possible metric frameworks for a given cycle is needed as a starting point. Due to its principal focus on metric ambiguity of cyclic patterns, PARNCUTT's quantitative model of pulse salience and metrical accent constitutes an excellent basis for this module. A central assumption of this paradigm

> is the inherent ambiguity of the underlying pulse (tactus) and meter of a rhythm. The model does not output a single solution, but instead considers many possible pulse and meter sensations, estimating the relative importance or salience of each.[69]

In particular, this approach accounts for two aspects derived from the experimental data discussed, among others, in sections 2.2.3 and 3.2. It combines a quantitative account to model the dispersion of different pulse responses – that is, the metric ambiguity represented by the data – with a concept of pulse salience, reflecting the preference for moderate tactus rates.[70] Furthermore, the model is restricted to cyclic patterns, implementing a top-down approach to estimate metrical accent strengths for pulse categories in rhythmic necklaces.

[69]Parncutt, 1994, p. 423, also mentioning the similarity of this model to Lee, 1991, Povel and Essens, 1985, and Rosenthal, 1992 in this respect.

[70]Due to this wide range, the model received plenty of attention and is reflected in numerous subsequent models (for instance in Eck, 2001 and Flanagan, 2008, which is also influential for this study).

In his model of metric ambiguity and syncopation, FLANAGAN employed the model of PARNCUTT with little modifications.[71] In the following section (5.2.1), this approach is analyzed and further developed, resulting in a revised account for pulse salience. Thereafter, I follow PARNCUTT and FLANAGAN in characterizing the metric ambiguity of a rhythmic cycle by the distribution shape of possible metric interpretations.

5.2.1 A revised model of pulse salience

The short rhythmic necklace [1-2-3] from the introductory example in section 1.1 (see figure 1.1) will now be reused to illustrate the adaptation of the existing models of pulse salience for the present approach. To explore the metric malleability of this cycle, first, the tactus ambiguity may be estimated by a comparison of quantified saliences of candidate pulses. Two relevant aspects of pulse salience are already discussed in section 2.2.3. First, a relative salience could also be interpreted as a probability, for that a particular pulse is likely to be chosen as tactus. Second, according to PARNCUTT's model, pulse saliences are highly dependent on the presentation rate of the sequence, as pulse salience is defined as the product of *pulse-match salience* and *pulse-period salience*.

Pulse-match salience represents "the goodness of fit [...] of a pulse of period P and phase Q"[72] in relation to the onsets of the rhythmic pattern. As the model operates on the basis of a grid of elementary pulses, the period is defined by the number of elementary pulses and the phase is denoted by an index of an elementary pulse on which the pulse falls.[73] In the words of PARNCUTT, the "perception of pulse trains in musical rhythms may be conceptualized as a process by which sound onsets are matched against the elements of an *isochronous template*."[74] Thus, pulse-match salience reflects the aspect of pattern recognition or pattern matching (see section 2.2.3). This implies that a pulse sensation may be the more salient, the more onsets coincide with the pulse. In other words, pulse-match salience measures the *positive evidence* for a pulse sensation found in the onset pattern of a rhythm.[75] However, this approach is obviously limited to categorically isochronous pulse sensations. The present approach instead accounts for possible non-isochronous pulses, as amplified later in section 5.2.3.

Pulse-match salience is differently formalized in the models to be employed. The basis of PARNCUTT's implementation is formulated in his assumption as follows.

[71] Flanagan, 2008. In the current context, the syncopation aspect of this model, which is shorty surveyed in section 4.3.1, is of little relevance. Metric malleability concerns the "state" of rhythm before its meter is chosen, a feature which is resolved by metric interpretation. In contrast, syncopation can only arise in relation to a metric framework, that is, as a consequence of metric interpretation.

[72] Parncutt, 1994, p. 434

[73] As already mentioned in section 4.4.2 (footnote 236), the number of different phases in which a pulse with a certain period can occur, is equal to the greatest common divisor between the pulse period and the cycle period.

[74] Parncutt, 1994, p. 433 (emphasis in source)

[75] cf. Eck, 2001

> *Each pair of events in a rhythmic sequence initially contributes to the salience of a single pulse sensation, in proportion to the product of their phenomenal accents; and contributions to the salience of a pulse sensation from different pairs of events add linearly.*[76]

In this model, numeric values for phenomenal accent strengths are derived by the formalism for *durational accents*, outlined in section 2.3.2 and more precisely described in the following (see equation 5.4). Thus, only temporal aspects of accentuation phenomena are considered, and other parameters like loudness, timbre, and so forth, are excluded. The assumption above also implies that "as few as two events can produce a sensation of pulse."[77] This was already discussed in section 2.2.1. There are however cases in musical practice, where the main beat coincides with less than two rhythmic onsets, that is, "some rhythms contain no inter-onset intervals (IOIs) that map onto candidate meters that we know to be viable meters."[78] As PARNCUTT's algorithm for pulse-match salience sums the products "of the phenomenal accents that occur on neighboring pulses",[79] those cases score a zero-value for pulse-match salience. FLANAGAN exemplifies this with the *Bossa Nova* clave [3-3-4-3-3] which is conventionally conceptualized in 4/4 meter (see figure 5.2).

FIGURE 5.2: The *Bossa Nova* clave, interpreted in 4/4 meter (Flanagan, 2008, p. 637, figure 2).

Here, only the first (quarter) beat is phenomenally accented, while the others are only cognitively present. Although the sequence [3-3-4-3-3] hardly invokes a pulse sensation with period 4 and phase 0 when presented repetitively as a pure rhythmic necklace without musical context, FLANAGAN suggests a modification to the pulse-match algorithm to account for those highly syncopated musical idioms. Instead of adding phenomenal-accent products of matching rhythmic and metric IOIs, he suggests to

[76]Parncutt, 1994, p. 434 (italics in source)

[77]ibid. Parncutt, 1994 further notes that this assumption "conflicts with Krebs' claim" (Krebs, 1987, footnote 16) "that at least three regularly recurring events are needed to evoke a level of motion. Krebs pointed out that two equal IOIs are needed to imply isochrony, and that three events are needed to define two IOIs. According to this line of reasoning, a sequence of alternating 1/2 and 1/4 notes (the *waltz* sequence in [figure A.1]) would never imply a pulse with a period of 1/4 note. Given that many listeners [...] tapped at 1/4 note intervals in response to the *waltz* sequence, Krebs' hypothesis may be rejected in favor of the principle quoted above. [...] if the above definition of pulse salience is valid, then pulse salience is *defined* by the results of Experiment 1", the data of which are reproduced in appendix A.

[78]Flanagan, 2008, p. 635, using the term "meter" for "an isochronous, tactus-level pulse, characterized by a period and phase expressed in discrete integer multiples of some smallest unit of musical time. The more common term for this isochronous pulse is, of course, 'the beat.' 'Meter,' however, is the preferable term because this paper uses the conventionally understood downbeat of musical examples as a reference point for the tactus-level pulse." (ibid.)

[79]ibid.

simply sum "the phenomenal accents that coincide with a candidate meter regardless of their adjacency."[80] Figure 5.3 shows the different results for pulse-match saliences of pattern [1-2-3], calculated upon durational-accent values according to an elementary pulse IOI of 0.2 s.

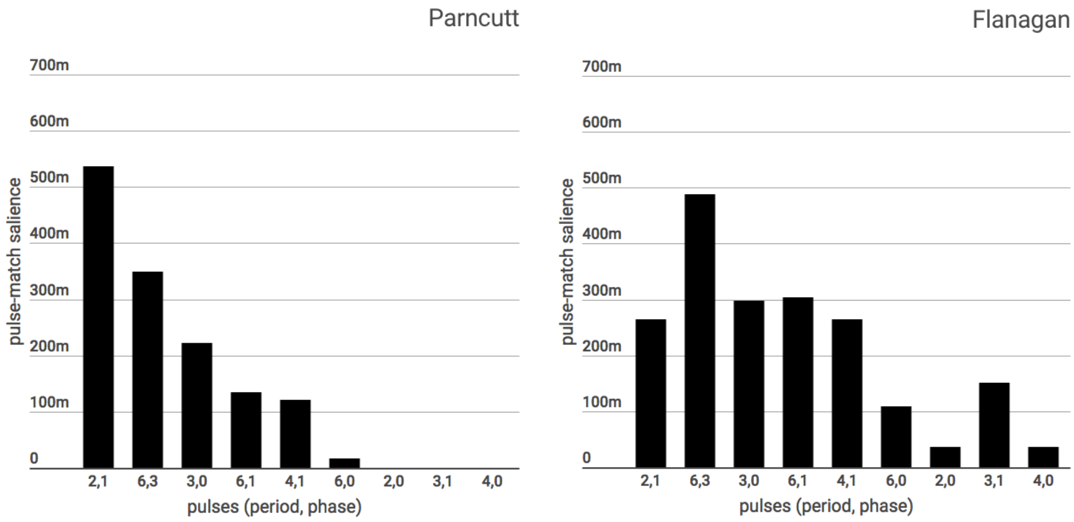


FIGURE 5.3: Pulse-match saliences for [1-2-3] according to Parncutt, 1994 and Flanagan, 2008 (elementary pulse IOI = 0.2 s).

Only pulses with periods that stand in relatively simple ratios to the length of the rhythmic pattern are listed in figure 5.3, and extremely high or low (fast or slow) pulse rates are excluded. In both approaches, pulse-match saliences are normalized to equalize effects of different pulse-period durations. PARNCUTT and FLANAGAN use diverging methods to accomplish this. The common purpose is "to introduce explicit tempo dependence",[81] that is, to separate the accounts for pattern matching (pulse-match salience) and for pulse rates (pulse-period salience, derived from the phenomena of preferred tempo and the existence region of pulse sensation, discussed in section 2.2.3). As PARNCUTT assumes essential independence between the "two contributions to pulse salience – goodness of fit between the rhythm and the corresponding isochronous template, and the tempo of the pulse"[82] – pulse salience S is estimated by multiplying pulse-match salience S_m and pulse-period salience S_p (see equation 5.1).[83]

$$S = S_m S_p \tag{5.1}$$

[80] Flanagan, 2008, p. 636
[81] Parncutt, 1994, p. 435
[82] Parncutt, 1994, p. 438 (written in italics in the source)
[83] cf. Parncutt, 1994, p. 439, where the original proposition states $S = (S_m S_p)^j$. The *pulse-salience index* j (which is generally assumed to exceed 1) accounts for the amount of metric ambiguity, regarding different pulse candidates: the higher the index value, the more the saliences of different pulses differ. This index is not implemented by Flanagan, 2008.

To assimilate and combine advantages and to eliminate inconsistencies of both measures, I suggest to define pulse-match salience in a sequence of rhythmic onsets for a pulse with period P and phase Q, expressed in elementary-pulse units, as follows:

$$S_m = \frac{\sum_{n=0}^{N-1} \min(A_d(IOI(mod_C(nP+Q))), A_d(eP))}{NA_d(eP)} \tag{5.2}$$

where $N = lcm(P, C)/P$ and mod_C is the *modulo* to the base of C (the cycle length, expressed in elementary-pulse units). $IOI(T)$ is the categorical interval duration between a rhythmic onset coinciding with an elementary pulse or temporal category T and the next temporal category which contains an onset. The durational accent $A_d(\Delta t)$ was originally defined by PARNCUTT as $A_d(T)$, that is, the durational accent of an onset falling in T.[84] It is adopted here in terms of an interval Δt in the sense above, according to equation 5.4.

$$A_d(\Delta t) = (1 - \exp(\frac{-\Delta t}{\tau}))^{i} \tag{5.4}$$

[85]

If we insert eP – the elementary-pulse interval e times P – we receive the durational-accent value for the pulse itself, denoted as $A_d(eP)$. This value is basically needed for the mentioned normalization of pulse-match salience. For this purpose, I prefer to assimilate PARNCUTT's precise calibration for pulse-match salience, "so that it equals one in the simple case where the stimulus is an isochronous sequence and the pulse sensation corresponds exactly to the stimulus."[86] Therefore, the function $min(a, b)$ outputs the minimal value found in its argument list. This operation also limits the accentual influence of longer IOI in the context of faster pulse sensations. In such situations, loud rests or silent pulses emerge which may interfere with durational accent.

To compare these adaptations to the original measures, figure 5.4 displays the pulse-match saliences of pattern [1-2-3] for the same pulses as in figure 5.3. Durational-accent values are again calculated according to an elementary pulse IOI of 0.2 s. In general, the relations between saliences in figure 5.4 combine features of both distributions shown in figure 5.3. Though, it is striking that some pulses are estimated to have a relatively greater salience compared to that according to the other measures, as the pulses (3,0),

84

$$A_d(T) = (1 - \exp(\frac{-IOI(T)}{\tau}))^{i}, \tag{5.3}$$

where $IOI(T)$ is the interval duration between T and the next elementary pulse or temporal category which contains a rhythmic onset; cf. Parncutt, 1994, p. 430 (equation 3). This definition is as well employed by Flanagan, 2008, p. 636 (equation 1).

85 The free parameters τ and i account for perceptual constraints and can be varied to adapt the model to experimental data: "τ is the saturation duration (assumed proportional to the duration of the echoic store), and i is the so-called *accent index*, accounting for the minimum discriminable IOI." (Parncutt, 1994, pp. 430 f.) For further discussion about durational accent see section 2.3.2.

86 Parncutt, 1994, p. 435

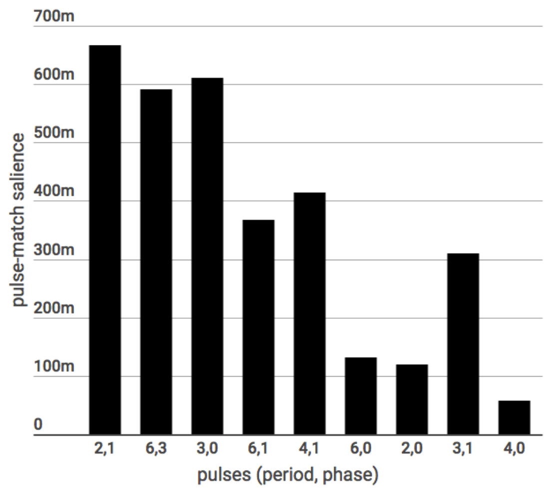


FIGURE 5.4: Pulse-match saliences for [1-2-3] according to equation 5.2 (elementary pulse IOI = 0.2 s).

(4,1), and (3,1). To a certain extend, their prominent role may be interpreted as a combination of the relatively high saliences, suggested both by PARNCUTT's and FLANAGAN's method. However, in particular the relative strength of pulse (4,1) highlights a type of pulse which may often be underestimated because its period ($4e$) is not a divisor of the cycle length ($6e$). According to *metrical preference rule 1* by LERDAHL and JACKENDOFF (section 2.1.3), this pulse may not be supported but rather obscured by the repetition of the rhythmic pattern due to the period of the cycle. Nevertheless, it may readily emerge under convenient temporal conditions, as two rhythmic cycles and three beats form a structural periodicity. This and other similar metric interpretations were not yet considered in the introductory example, illustrated in figure 1.1. If we assume the pulse (4,1) to emerge as the tactus level, the three metric interpretations shown in figure 5.5 can be construed in extension of figure 1.1.

FIGURE 5.5: Two cycles of [1-2-3] with tactus (4,1), interpreted according to different measure beats in 3/2 time signature.

Obviously, this leads to a considerable extension of the space of metric ambiguity, and thus, of the metric malleability of this simple rhythmic cycle.[87] Such possible interpretations are ignored by FLANAGAN, as he only studies pulses having periods that are factors of the period of the rhythmic cycle. Hence, the salience value for pulse (4,1) in figure 5.3 according to his method is estimated by considering two rhythmic cycles as

[87]See also the parallel discussion of pattern [x.xxx.] in section 3.2.4.

one. PARNCUTT explicitly integrates those cases, not at least because his experimental data indicate the emergence of corresponding pulse sensations.[88]

Equipped with the proposed method to estimate pulse-match salience, I adopt equation 5.1 to account for the tempo dependence of pulse salience. The discussion of possible function curves for pulse-period salience in section 2.2.3 suggests that the optimal function may vary according to various properties of the rhythmic or musical input.[89] However, for the sake of consistency, I follow PARNCUTT[90] and FLANAGAN[91] to model pulse-period salience according to the Gaussian function over the logarithm of the temporal interval corresponding to the pulse period, reproduced in section 2.2.3 (equation 2.1). As mentioned, this aspect is regarded as independent from pulse-match salience and thus independent from the structure of the rhythmic input. Figure 5.6 provides the resulting pulse saliences of rhythm necklace [1-2-3] according to the three measures of pulse-match salience, combined with pulse-period salience (according to equation 5.1). It shows that the effect of pulse-period salience enhances pulse periods near the *preferred tempo* (see sections 2.2.2 and 2.2.3), a parameter, fixed at $0.6s$ for the calculations. With an elementary-pulse IOI of $0.2s$, period 3, and to a lesser extend 4 and 2, gain more preference compared to the ranking merely according to pulse-match salience.

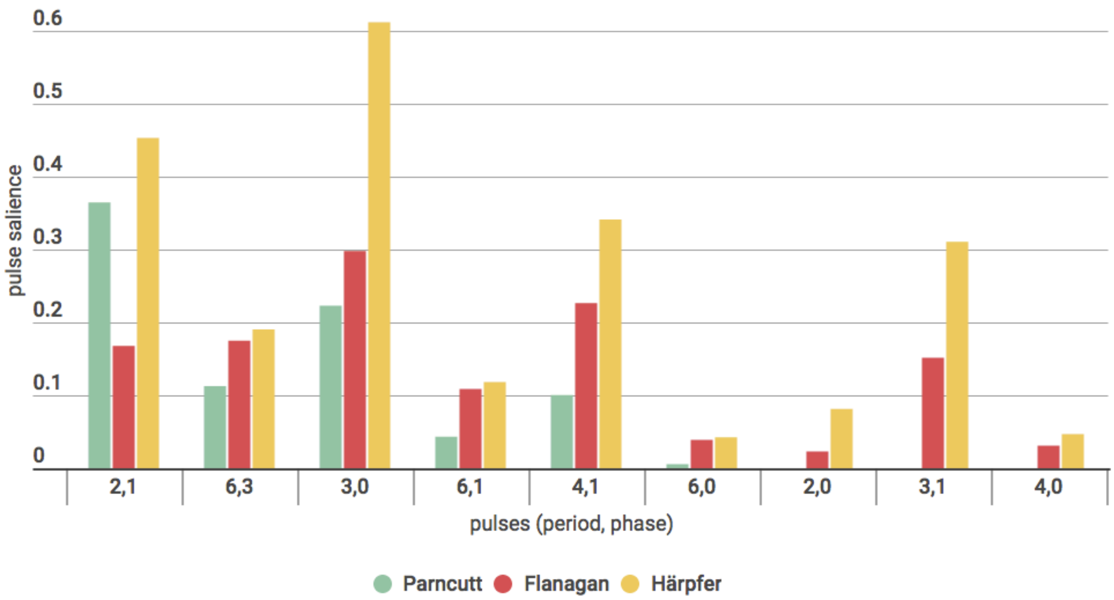


FIGURE 5.6: Pulse saliences for [1-2-3] according to equation 5.1 and the three measures of pulse-match salience (elementary-pulse IOI = 0.2 s).

[88] cf. figure A.2 (e) and (f). Although there are no responses to rhythm (e) (exactly representing the necklace [1-2-3]) which correspond to the pulse (4,1) in figure 5.3 and 5.4, the pulse period $4e$ was tapped to the similar rhythm (f) with period $6e$.

[89] cf. Noorden and Moelants, 1999, Moelants and McKinney, 2004, and Gotham, 2015a.

[90] Parncutt, 1994, p. 438 (equation 6)

[91] Flanagan, 2008, p. 636 (equation 3)

As mentioned before, the illustrated methods to estimate pulse salience are restricted to isochronous pulse streams. In regard to mixed metrical beats, implied by numerous musical styles (as amplified in section 3.2.1), a more flexible approach to pulse salience would be desirable. It should integrate both isochronous and possible non-isochronous pulse sensations. The present approach suggests to realize this by an alignment of the rhythmic cycle to all simple and mixed metric structures, conforming to the metrical well-formedness criteria which are mainly discussed in section 4.2.1. To this effect, pulse salience may be generalized and specified at the same time, by developing a measure of *meter salience* (see section 5.2.3).

5.2.2 Salience distributions, metric ambiguity and malleability

According to the present model of pulse salience, it may be preliminarily inferred that a rhythm is the more metrically ambiguous the more similar the perceptual saliences, calculated for different pulses.[92] More differentiated, the type of metric ambiguity involves the formal relation of pulses with similar saliences. If those pulses are consonant, reference-level ambiguity may arise (section 3.2.3). If they are dissonant, period ambiguity (section 3.2.4) or phase ambiguity (section 3.2.5) may be present. Regarding metric malleability, the model of pulse salience may be employed for the following hypothesis: *a rhythm is notably malleable, that is, amenable for metric modulation or reframing, if another pulse has a similar or a higher salience than the currently established tactus*. In other words, a metrically malleable rhythm can be similarly well integrated into different metric contexts.

Most of the diagrams in the previous section illustrate salience distributions for different pulses implied by a rhythmic cycle. Note that, according to the definition of pulse salience as a probability for choosing that pulse as the tactus of a rhythmic sequence (section 2.2.3), a dispersion of pulse saliences could also be employed as a probabilistic weighting scheme for random choices. The evenness or flatness of such distributions initially tells us something about the global amount of metric ambiguity of a cycle, without specifying the particular interpretations competing with each other. Statistic measures of the evenness of a distribution are manifold and employed in multiple disciplines.[93] In the context of rhythm complexity, for instance, TOUSSAINT demonstrates the benefit of applying *information entropy*[94] as a measure for the flatness of (temporally categorized) IOI histograms derived from a rhythm. Thereby he acknowledges "entropy's ability to measure the flatness of any histogram, no matter what its origin."[95]

[92] This corresponds to the notion of Flanagan, 2008, as amplified later.

[93] cf. Toussaint, 2013, p. 111: "There are an uncountable number of ways to measure the flatness of a probability distribution".

[94] The notion of entropy in information theory was developed by Shannon, 1948. Toussaint, 2013 relates to several applications of information entropy and other information measures to aesthetics and music perception (see p. 111).

[95] ibid.

For a probability distribution $P^n = (p_1, p_2, ..., p_n)$ with an arbitrary n, entropy can be defined by

$$H(P) = -\sum_{i=1,n} p_i log_n(p_i).^{96} \quad (5.5)$$

Specifically related to the current topic, FLANAGAN suggests to calculate "the inequality of the dispersion of induction strengths across a population of candidate meters"[97] by means of the GINI coefficient. Developed "in economics as a measure of income inequality",[98] it is advantageous as "its value is independent of the scale of the data (i.e., unit of measurement) and the sample size."[99] Comparable to entropy, where we can regard P as a collection of pulse saliences $p_1, p_2, ..., p_n$, normalized to a sum of 1, a discrete version of the GINI coefficient can be defined as

$$G(P) = \frac{\sum_{i=1,n} \sum_{j=1,n} |p_i - p_j|}{2n^2 \mu}, \quad (5.6)$$

where μ is the average value of P, in our case, the mean pulse salience (or mean "induction strength" in FLANAGAN's terminology).[100]

Thus, by means of the entropy, the GINI coefficient, or another comparable method, shapes of statistical dispersion can be translated into single quantities. To continue the analysis of the rhythmic necklace [1-2-3], entropies and GINI coefficients of the pulse-salience distributions displayed in figure 5.6 are compared in figure 5.7.

The left part of the diagram illustrates the reciprocal relation of the two measures: the higher the entropy the lower the GINI coefficient. As a higher entropy value indicates a flatter distribution (of pulse saliences in this case), the same applies for a lower GINI coefficient. "In music theoretical terms, lower values denote greater metrical ambiguity, and vice versa."[101] If we flip the scale for the GINI coefficients like in the right part of figure 5.7, we see that both measures provide similar analytic tools. It turns out that in the case of pattern [1-2-3] PARNCUTT's distribution of pulse saliences implies considerable lower metric ambiguity compared to the others. The main reason are the four pulses listed rightmost in figure 5.6. According to PARNCUTT, their saliences are

[96] Toussaint, 2013 (ibid.) basically employs this formula, but statically uses log_2 instead of log_n, specifying the *information content* of a variable related to a binary decision. In this case, the maximal entropy $H_{max} = log_2(n)$ is > 1 with $n > 2$. For broader comparability, equation 5.5 implies a normalized $H_{max} = log_n(n) = 1$, as it will be useful for our purposes.

[97] Flanagan, 2008, p. 636

[98] ibid.

[99] ibid.

[100] Flanagan, 2008 (ibid.) notes that the Gini coefficient "produces values with a range of 0 to 1, where 1 signifies perfect inequality and 0 signifies perfect equality." Though, the present version has a well-known flaw: the smaller the list of values, the lesser a perfect inequality ϕ approaches a result of 1. A list P^n_ϕ of n values with $p_{1,..,n-1} = 0$ and $p_n > 0$ (an arbitrary positive value) yields $1 - 1/n$. To gain a normalized result for different n, I suggest to calculate $G_{norm} = G(P^n)/G(P^n_\phi)$.

[101] Flanagan, 2008, p. 637

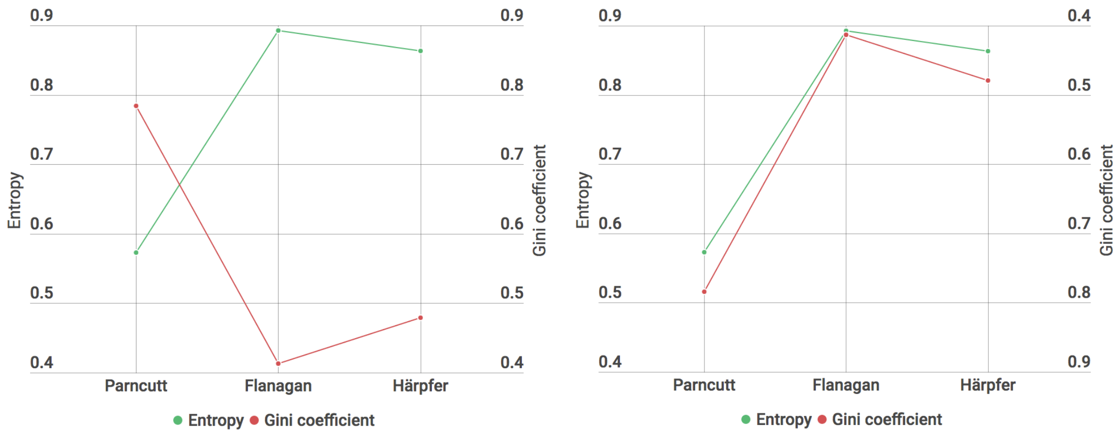


FIGURE 5.7: Two comparisons between entropies and GINI coefficients of the pulse-salience distributions for [1-2-3] according to figure 5.6.

zero or almost zero which makes the global shape of distribution less flat. The measure derived from the suggestions of PARNCUTT and FLANAGAN and introduced in the previous section is more similar to FLANAGAN's, as it produces as well relatively high values of entropy (corresponding to low GINI coefficients). Accordingly, it follows the idea that rhythmic cycles basically tend to invite a broader range of metric interpretation and provide notable metric malleability.

To summarize, measures of distribution evenness can be employed to globally estimate metric ambiguity and malleability from pulse salience distributions, such as those resulting from the procedure proposed in the previous section. Though, the loss of information about specific pulses involved in the reduction to a value of overall flatness indicates limited insight by the use of these tools.

5.2.3 Meter salience

Basically, any metrically well-formed pulse – that is, any metric stratum – can be aligned to a rhythm to estimate its pulse salience. Thus, meter salience may be conceptualized as a combination of the pulse saliences of all perceptually relevant metric strata. This implies that all meters, including simple and mixed structures, could be integrated in a more generic approach to metric malleability. Extrapolating PARNCUTT's procedure, the saliences of consonant pulses, or of the strata in any well-formed metric hierarchy, may be added to measure meter salience.

> The perception of meter may be assumed to involve the simultaneous perception of concurrent pulses, and the salience of a perceived meter may be modeled by adding calculated pulse saliences.[102]

[102] Parncutt, 1994, p. 424

This paradigm is based on theoretic frameworks of hierarchic meter discussed in section 1.3 and 4.2, and on the notion of metrical consonance, developed by KREBS and others (section 2.3.3). These concepts led PARNCUTT to the assumption that a "prerequisite for the perception of meter is the concurrent perception of consonant pulses",[103] and that a "consonant set of pulses may be conceived as a hierarchical structure."[104]

Conceptualizing meter salience as an aggregate salience of consonant pulses also underlies GOTHAM's proposition of "attractor tempos" for specific hierarchical meters.[105] Given that metric contexts incline musical performance towards particular tempos,[106] he basically assumes that such behavior is attracted by optimizing the saliences of consonant pulses that make up the metric context. This is particularly interesting, because precisely the deviations from such "optimal" or "attractive" tempos may be the crucial information to be derived for tempo interpretation, that is, to judge a passage of music as fast or slow. Attractor tempos are thus related to the topics of preferred tempo, "moderate" tempo, and maximal pulse salience (sections 2.2.2 and 2.2.3). GOTHAM developed his method to calculate attractor tempos according to the idea of meter salience, but his aggregations of pulse-period saliences merely correspond to abstract metric templates, as they were examined in section 4.2. To establish meter salience as an aspect of a concrete rhythm, it is necessary to employ the particular pulse saliences – comprising the pulse-match aspect – of a particular rhythm, as outlined above. However, PARNCUTT's hypothesis is not yet empirically verified in a systematic way.[107] In the following, I therefore propose a theoretic procedure to approach a heuristics which may be further developed and adapted according to more specific knowledge in the future.

In spite of its abstract level, GOTHAM's approach turns out to be a suitable basis for a model that calculates meter saliences for individual rhythms. He hypothesizes meter salience for "full-represented" meters, that is, meters which are embodied in full metric depth by the texture of rhythmic components.[108] Thus, his measure of *metrical salience*, "defined by the combined saliences of all the pulse streams present in the given structure" is equivalent to "the metrical weight for the highest level present."[109] Both are defined for simple meters by

$$M = \sum_{n=1,p,q,r,\ldots} \exp(-\frac{(\log(nx/\mu))^2}{2\sigma^2}), \tag{5.7}$$

[103] Parncutt, 1994, p. 443 (written in italics in the source)
[104] ibid.
[105] See sections 2.2.3, 2.3.1, and 5.1.3 (item 1).
[106] Gotham, 2015a, p. 23
[107] cf. item 1 in section 5.1.3
[108] cf. sections 2.1 and 2.2.1
[109] Gotham, 2015a, p. 39

where x represents the period of the elementary pulse. The periods of higher metric levels are then denoted "by the relevant multipliers $(px, qx, rx, ...)$"[110] and n is iterated according to these integer multipliers representing consecutive metric levels. This implies that higher levels get more weight by adding up their own and all lower-level saliences.[111] In this way, GOTHAM's notion of metrical salience as a measure of metric weight considers that

> musical intuition and theoretical accounts of the metrical hierarchy tell us that metrical positions corresponding to longer durations have greater metrical weight. These principles are perfectly compatible as long as metre is modelled as the sum of several periodicities [...]. Modelling metre in this way necessarily preserves the positive correlation between duration and metrical weight while also leaving room to include a weighting of the constituent pulses' importance.[112]

The measure represented by equation 5.7, can thus be interpreted as another approach to a model of metric weight in the line of the propositions discussed in section 4.3. Metric weights, which are associated with metrical positions or temporal categories, are estimated by summing up pulse-period saliences of the pulses converging at these positions.

As argued before, distribution features of pulse saliences, calculated for a rhythm, can tell us something about its metric malleability. Moreover, the approach to estimate saliences of candidate pulses which could emerge as the tactus of a given rhythmic cycle, may be further differentiated by the involvement of meter salience. In the following, metric malleability is thus appraised on the basis of the saliences of possible hierarchic metric interpretations, rather than on the basis of pulse saliences. To employ metric hierarchies according to the definitions in section 4.2, accounting for both simple and mixed meters, they are encoded as GNSM strings. Modeling meter salience as the sum of constituent pulse saliences then requires to redefine pulse-match salience for any well-formed metric stratum. To formalize this in analogy to equation 5.2, it is practical to represent single metric strata M_s by series of indices enumerating the temporal categories T_i of an ongoing elementary pulse that belong to that stratum. Two instances may illustrate this. First, the strata of the simple meter described by the GNSM string (3 0 1 0 2 0 1 0) can be defined as index series

$$M_0 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, ...\};$$
$$M_1 = \{0, 2, 4, 6, 8, 10, ...\};$$

[110] Gotham, 2015a, p. 31

[111] Gotham, 2015a, p. 29: "the weight of a metrical position is given by the sum of the saliences for the pulses which coincide there (as defined by the metrical hierarchy)." For an explicit formalization of this principle see Gotham, 2015a, p. 30 ("Equation 3" and the illustration in "Table 1").

[112] Gotham, 2015a, p. 29

$M_2 = \{0, 4, 8, 12, ...\};$
$M_3 = \{0, 8, 16, ...\};$

or secondly, the strata of the mixed meter (2 0 0 1 0 0 1 0) are represented by

$M_0 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, ...\};$
$M_1 = \{0, 3, 6, 8, 11, 14, ...\};$
$M_2 = \{0, 8, 16, ...\};$

both following the rule: $T_i \in M_s$ if $gnsm(mod_M(i)) \geq s$.[113] A generalized pulse-match salience of a metric stratum M_s can then be defined by

$$S_m(M_s) = \frac{\sum_{n=0}^{N-1} \min(A_d(IOI(mod_C(M_s(n) + Q))), A_d(e(M_s(n+1) - M_s(n))))}{NA_d((M, C)e/N)}, \tag{5.8}$$

where $N = lcm(M, C)B(M_s)/M$ and $B(M_s)$ is the sum of beats belonging to M_s over the length M of a metric cycle in elementary pulses. For instance, in the examples above, $B(M_1)$ equals 4 in the first case and equals 3 in the second case. As in equation 5.2, mod_C is the *modulo* to the base of C (the length of the rhythmic cycle in elementary pulses) and Q represents the phase relation between the rhythmic and the metric cycle (see section 5.2.1).

Reconsidering pulse salience S, it is evident to define the salience of a metric stratum $S(M_s)$ in parallel to equation 5.1 as the product of $S_m(M_s)$ and $S_p(M_s)$. To estimate $S_p(M_s)$, it is necessary to determine a pulse period (section 2.2.3). As we want to include non-isochronous pulses like M_1 in the second example above, a practical approach would be to calculate $S_p(M_s)$ on the basis of the average pulse rate. This does not change anything for isochronous pulses but fairly represents the range of periodicity embodied by non-isochronous, metrically well-formed pulses. Equation 5.9 thus generalizes pulse-period salience for any metric stratum. The average pulse period $\bar{x}$ (1/ average pulse rate) can be determined by $eM/B(M_s)$.

$$S_p(M_s) = \exp(-\frac{(log(\bar{x}/\mu))^2}{2\sigma^2}) \tag{5.9}$$

The concept of meter salience presented here is supposed to express the fitness between a rhythmic and a metric cycle. Both are assumed to be synchronized according to an elementary pulse (a series of temporal categories) and to share a period ratio and phase relation defined in terms of that pulse. Meter salience may then be preliminarily formalized as the sum of pulse saliences of all relevant metric strata $S(M_s)$ (determined by $S_p(M_s)S_m(M_s)$, see equation 5.1), divided by the number of metric strata L:

[113] The function $gnsm(i)$ is supposed to return the GNSM value, that is, the number of interpretive metric strata (above the elementary-pulse level) converging at elementary pulse i (see section 4.2.3).

$$S_M = \frac{1}{L} \sum_{m=0}^{L-1} S(M_s) \tag{5.10}$$

To give a first illustration, consider the different metric interpretations in figures 1.1 and 5.5. For almost each interpretation, a different meter salience can be assessed according to the proposed algorithm. Table 5.1 lists the interpretations in figure 1.1 in descending order of their correspondent saliences, that is, from highest to lowest meter saliences, calculated for two different periods of elementary pulses ($e = 0.20s$ and $e = 0.25s$). The listings may be interpreted as rankings of plausibilities for particular metric frameworks which may emerge with the presentation of the cyclic pattern [1-2-3].

TABLE 5.1: Metric interpretations in figure1.1 and corresponding meter saliences for $e = 0.20s$ and $e = 0.25s$

$e = 0.20s$		$e = 0.25s$	
bar number	meter salience	bar number	meter salience
10	0.276	4	0.263
7	0.227	6	0.253
4	0.223	10	0.244
6	0.199	2	0.232
2	0.160	7	0.221
12	0.152	12	0.145
9	0.113	9	0.125
1	0.050	1	0.075
5	0.036	5	0.067
3	0.036	3	0.067
11	0.009	11	0.027
8	0.009	8	0.027

Within one tempo condition, the meter saliences of very different frameworks are estimated quite similar, in particular for period $e = 0.25s$. As both rates are only slightly different (they stand in ratio 4/5), the changes of the order among the five interpretations which gain the highest meter saliences are remarkable. At the slower rate, the quarter pulse gets more salient compared to the dotted-quarter pulse which is estimated stronger at the faster rate. Thus, the slower rate prefers a 3/4 beat whereas the faster rate favors a 6/8 beat. Note, that the tempo change is not as relevant for interpretations of low meter salience, respectively, of low plausibility. These aspects – extensive similarities of meter saliences and different rankings of preferable metric frameworks at similar rates – strongly indicate the considerable metric ambiguity and malleability of pattern [1-2-3].

The interaction of meter saliences with pulse rates can be further illustrated by the study of salience distributions, as introduced in the previous section 5.2.2. Figure 5.8 plots meter saliences for the mentioned five interpretations, which change order under the two tempo conditions compared in table 5.1, against elementary-pulse period.

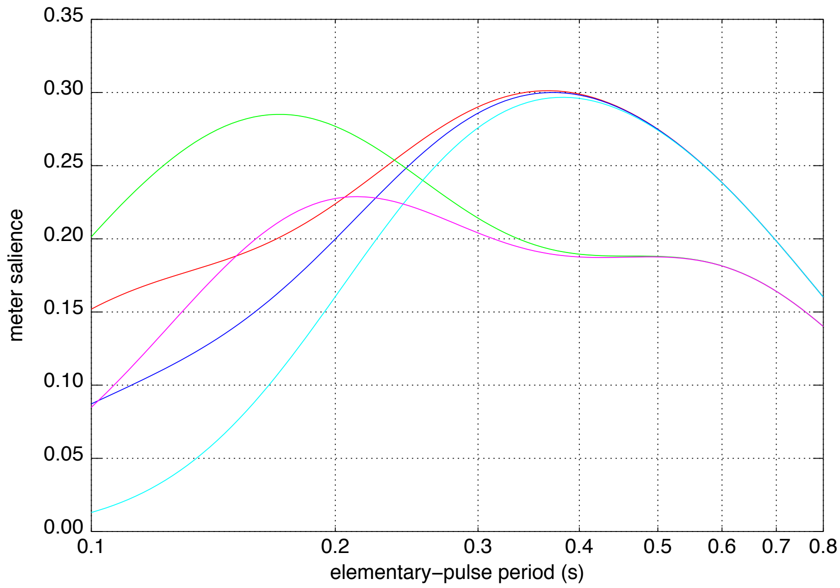


FIGURE 5.8: Meter saliences of five interpretations in figure 1.1 plotted against pulse rate (red: bar 4, green: bar 10, blue: bar 6, magenta: bar 7, cyan: bar 2).

A clearer picture is revealed, as the plot substantiates the assumed relation between the variants of 3/4 and 6/8 frameworks. All shown meter saliences bundle around similar values at pulse periods about $0.25s$. At slower rates, 3/4 variants converge on a higher salience level, whereas the 6/8 variants run together at lower saliences. All frameworks downgrade and converge with further slowing, as the saliences of most decisive component pulses become insignificant. At the fastest rates, the situation is more complex. Individual metric interpretations may be differently affected by slight tempo changes at these rates, if the slopes of the salience curves and the similarities of their values are regarded as aspects of metric malleability. This may be further corroborated by interpreting functions of the introduced measures of distribution, dependent on pulse rate. Figures 5.9 and 5.10 display the correspondent normalized GINI coefficients and normalized entropy values of the meter-salience set shown in figure 5.8.

As it could be already supposed by interpreting figure 5.8, the distribution of meter saliences is most even at pulse periods around $0.25s$. Both measures reach maximum – respectively minimum – values under that temporal condition. Irrespective of similar values at the slowest rates due to a general leveling of meter saliences, the region next to an elementary-pulse period of $0.25s$ can be considered as an attractor tempo for the metric malleability of pattern [1-2-3]. Transferring GOTHAM's notion of "attractor tempos for metrical structures"[114] to metric malleability, a rhythmic pattern may be particularly malleable at specific tempos. Thus, for the pattern [1-2-3], an attractor tempo about 40 cycles per minute can be assumed for its metric malleability in respect

[114]Gotham, 2015a

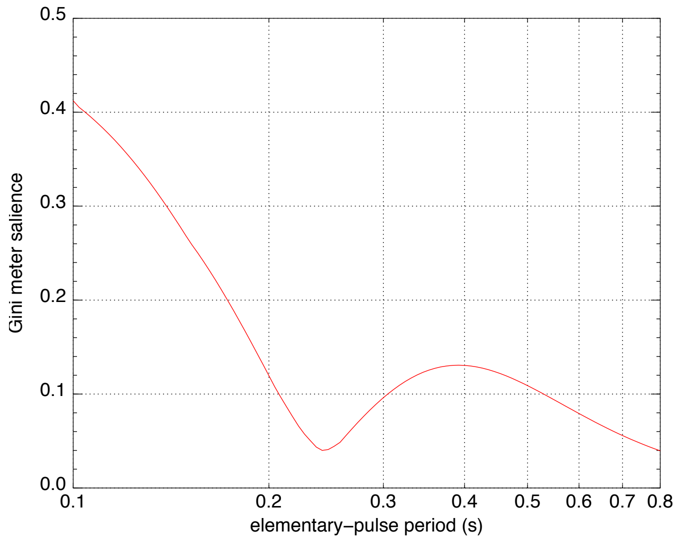


FIGURE 5.9: Normalized GINI coefficients derived from the data in figure 5.8

to the five interpretations taken into account for the present calculations. If we would include all twelve metric interpretations listed in figure 1.1 and in table 5.1, the distribution measures, compared for different pulse rates, would yield a less clear image: there would be no clear temporal attractor for metric malleability. In this respect it seems to be more meaningful to compare just a few candidate frameworks for which considerable peaks of meter salience can be found. It will be demonstrated later that it is particular interesting to focus on two or more specific candidate frameworks for a certain pattern, and collect those data which may characterize metric malleability pertaining to these candidates. Certainly, the present considerations prove the complexity of metric ambiguity and malleability even of relatively simple rhythmic patterns, especially when regarded as a function of elementary-pulse rate.

A comparison of the examples in figure 5.5, provided in table 5.2, reveals as well the marginal impact of higher (or "slow") metric strata on meter salience. They only minimally differ in that for each case, a different pulse-match salience is suggested merely for the pulse of the measure level. Though, as the period of this level which corresponds to the data in table 5.2 is $12e = 2.4s$, its pulse-period salience is very low. Thus, we can almost neglect its influence on meter salience.

TABLE 5.2: Metric interpretations in figure 5.5 and corresponding meter saliences for $e = 0.20s$

bar number	meter salience
2	0.206915
1	0.206408
3	0.205578

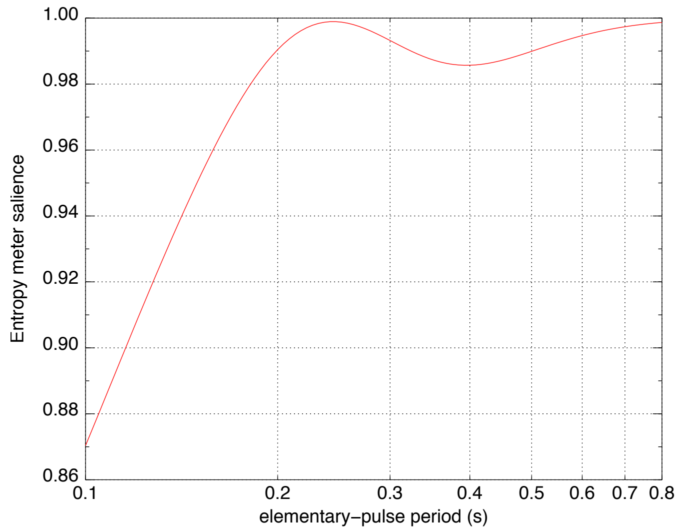


FIGURE 5.10: Normalized entropy values derived from the data in figure 5.8

On the one hand this may be counterintuitive in regard to what was said earlier about the correlation between metric level and metric weight. As said, GOTHAM assumes equivalence of the weight of a metrical position – estimated by summing the weights of all levels converging at the temporal category of the position – and meter salience up to the level of that position. This is basically in line with the models of metric weight, discussed in section 4.3, which generally propose higher metric weights for positions belonging to higher metric strata.

On the other hand, a low influence of long metric periods on meter salience, due to low pulse-period saliences, properly represents typical equivocalities arising with higher-level metrical grouping. Those were discussed by a number of authors, in particular in terms of the ambiguity of hypermetric levels.[115] Corresponding to higher-level metric ambiguity caused by the weak perceptual salience of longer metrical periods, metrical grouping at those levels may also be less consistent. Indeed, phrase structures of Western tonal music often suggest some irregularity regarding metrical grouping on levels above the notated measure.[116] *Metrical Preference Rule 10*[117] by LERDAHL and JACKENDOFF therefore

[115] cf. Gotham, 2015a, p. 24 (footnote 3): "The clarity of the hierarchy at the tactus level as compared with the relative ambiguity of more remote, hypermetrical levels is a key consideration which the present article addresses." See also Hasty, 1997, pp. 174 ff., and Parncutt, 1994, p. 423: "Metric ambiguity is particularly important in the case of hypermeter – meter that is perceived but not notated and whose period exceeds that of a notated measure." (referring to Rothstein, W. (1989). *Phrase rhythm in tonal music*. New York: Schirmer.)

[116] cf. for instance Benjamin, 1984, and Temperley, 2008. See also section 3.2.5.

[117] Lerdahl and Jackendoff, 1983, p. 101: "MPR 10 (Binary Regularity) Prefer metrical structures in which at each level every other beat is strong."

> "allows metrical irregularity, but, in the absence of other information, imposes duple meter. This seems to reflect musical intuition about hypermetrical structure."[118]

Compared to the meter saliences listed in table 5.1 for $e = 0.20s$, those of the 3/2 frameworks (cf. figure 5.5) given in table 5.2, are relatively strong at that rate. A different rate of elementary pulses would change those relations as well. An elementary-pulse period of $0.15s$ would for instance shift the period of the half-note beat in figure 5.5 to $0.6s$, corresponding to estimates of maximal pulse-period salience. Thus, a higher pulse rate would make the interpretations in figure 5.5 even more suggestive, reflected as well by rising meter saliences in relation to the frameworks in figure 1.1.

5.2.4 Studying meter salience for complex patterns

PARNCUTT developed another method to model the perception of a metric hierarchy on the basis of pulse saliences. Accordingly, the strength of a *metrical accent* can be quantified for each temporal category T by "linear addition of the saliences of all the pulse sensations converging on time T".[119] This approach corresponds as well with the principles of schematic models for metric weight, discussed in section 4.3, but suggests to calculate weights of metrical positions according to a particular rhythmic cycle. However, this method is also restricted to simple metric structures, as the estimate is based only on the saliences of isochronous pulses. The notion of meter salience developed in the previous section is in contrast designed to include mixed metric frameworks. The following examination evaluates this aspect, accompanied by a discussion of how to balance the saliences of isochronous and non-isochronous pulses.

I have argued that the metric malleability of simple repetitive patterns like [1-2-3] also allows for more complex metric frameworks, that is, metric cycles of higher cardinalities, like in figure 5.5. Nevertheless, repetitions of such short rhythms are usually strong metric cues, as basically discussed for parallel musical structures (section 2.4). They reduce metric ambiguity, that is, the variety of plausible metric interpretations. It is thus instructive to explore longer and more complex rhythmic cycles to gain a more profound view on meter salience and its relation to metric malleability. In particular, longer patterns more likely suggest metric cycles of higher cardinalities which offer a rich repertoire of both simple and mixed metric hierarchies.

As illustrated in section 4.2, different cycle lengths (cardinalities of elementary pulses) allow for either simple or mixed, or both categories of metric frameworks. Table 5.3

[118] ibid.

[119] Parncutt, 1994, p. 440. This approach fairly models experimental data relating metrical schemes in long-term memory and metrical position usage by Palmer and Krumhansl, 1990 (cf. sections 4.3.1 and 4.3.2).

summarizes the numbers of possible metric types, divided in the two categories for cycle cardinalities from 4 to 16 elementary pulses.

TABLE 5.3: Possible types of metric interpretation for cycles from 4 to 16 elementary pulses (cf. section 4.2.2)

elementary pulses	metric types: simple	mixed
4	1	-
5	-	2
6	2	-
7	-	3
8	1	3
9	1	4
10	-	6
11	-	14
12	3	14
13	-	30
14	-	40
15	-	61
16	1	81

In this range, cycles which comprise 8, 9, 12, and 16 elementary pulses enable both simple and mixed metric frameworks. Cardinality 12 permits the most variants of simple metric hierarchies, as one ternary level can be combined with two binary levels in three different ways: 2 x 2 x 3, 2 x 3 x 2, and 3 x 2 x 2. Assuming the elementary pulse at the eighth-note level, these structures correspond to the classical meters 12/8, 6/4, and 3/2 (a sixteenth-note pulse level would yield 12/16, 6/8, and 3/4). At the same time, this cardinality provides a relatively large, but still manageable number of mixed metric structures. It is therefore promising to conduct a series of calculations, estimating meter saliences over a range of elementary-pulse rates for a group of rhythmic patterns of that pulse cardinality. Though, from the perspective of metric malleability, the exploration of meter saliences and their dispersion properties (section 5.2.2) for meters of higher cardinalities is a combinatorially vast enterprise. Our instance, the metric 12 cycle enables three simple and 14 mixed meters, thus 17 possible metric frameworks in all. The number of possible phase relations between a rhythmic and a metric pattern is equal to the greatest common divisor of the cardinalities of both.[120] Thus, for every rhythmic necklace of cardinality twelve, 17 metric 12 cycles in twelve possible phase relations can be conceived, that is, 204 potential interpretive cycles.[121] Figure 5.11 provides a global impression of such meter-salience constellations for some rhythmic cycles of cardinality twelve. The selection covers four rhythmic necklaces with five onsets, one with six onsets (d), and one with seven onsets (f).

[120]See footnote 236 in section 4.4.2.

[121]Additionally, metric cycles consisting of 6, 8, and 24 elementary pulses may be included, as their cardinalities stand in simple ratios to the rhythmic cycle.

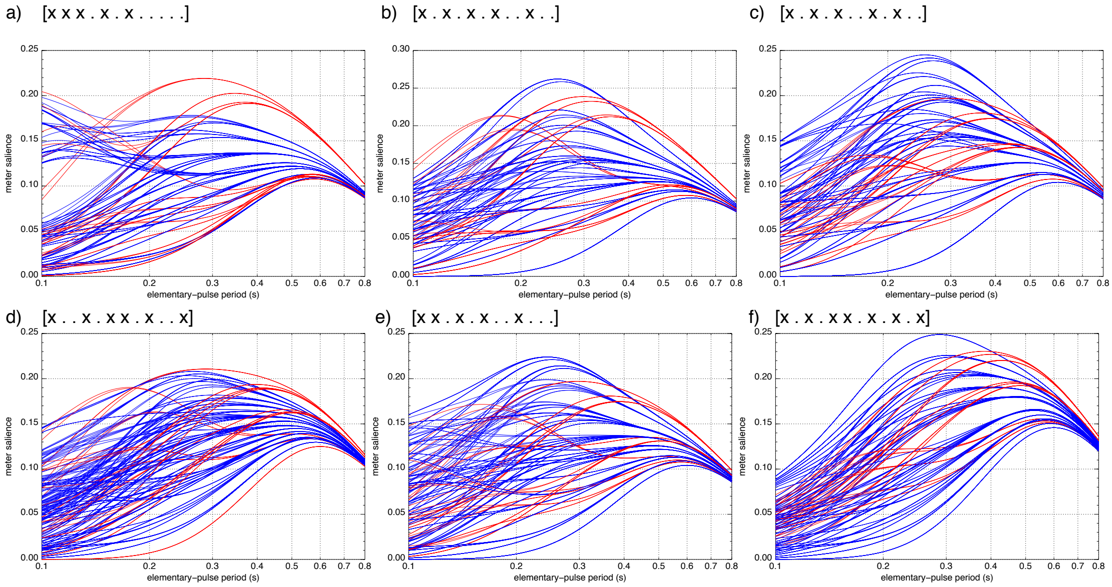


FIGURE 5.11: Saliences of 204 potential metric interpretations as functions of elementary-pulse period for some rhythmic necklaces of cardinality 12 with fife onsets (respectively six (d) and seven (f) onsets).

These overviews merely distinguish the salience curves of simple (red colored) and mixed metric frameworks (blue colored). According to the measure of meter salience developed so far, the distributions generally suggest to assume a relatively high prominence of mixed metric frameworks. This has to be put into perspective, as many factors interact, concerning their actual impact upon metric interpretation. Throughout this study, it is emphasized that long term exposure to specific metric structures has a crucial influence on metric schemes, that is, the top-down processing in metric interpretation. As amplified in section 3.2.1, such cultural backgrounds influence listener's preferences, to either choose simple or mixed metric frameworks when malleable rhythms would suit both.[122]

Before we further discuss how to balance these aspects, it may be informative to compare the examples in figure 5.11. Patterns *a* and *d* favor interpretations according to simple meters, as the highest meter saliences indicate. The contrary seems to hold for the others, at least for patterns *c* and *e*. Rhythm *f*, the "standard" pattern [2-2-1-2-2-2-1],[123] is also employed in section 2.3 to illustrate grouping variants (see page 56). It is argued there, that it can either be "metrically" grouped, suggesting a simple (isochronous) metric framework, or by contrast, "figural" or "serial" grouping takes

[122] cf. for instance Hannon and Trehub, 2005, Huron, 2006, Magill and Pressing, 1997, and Toussaint, 2015. Among simple metric frameworks, the preference for binary or duple meters discussed in sections 2.3.1 and 2.3.4, may also be fortified by the prevalence of these meters in many musical cultures.

[123] This pattern is further discussed, for instance in sections 3.2.2 and 4.1.1.

place. The latter may correspond to a mixed metric framework whose accents all coincide with rhythmic attacks. This would be impossible with an isochronous metric template below the half-cycle level. These variants may compete, depending on the pulse rate, as it can be interpreted from the salience curves of pattern *f*. Rhythm *c* – the other "standard" pattern examined before – is the complementary pattern of *f*. Both complements are maximally even, though, pattern c has less attacks. Therefore, it rather acts like a metric scheme which strongly suggests a mixed metric framework.[124] Finally, we have already referred to necklace b in different contexts, for instance, as the *guajira* pattern in flamenco,[125] or as a common pattern in African rhythm called the "hemiola" pattern by TEMPERLEY.[126] It can similarly act like a meter, although it is not maximally even. Indeed, particular metric patterns like this illustrate why LONDON's *metrical well-formedness constraint* 4.2.2 does not presume maximal evenness for any metrical level, but rather links an overall, hierarchic optimization of evenness to metric well-formedness (cf. section 3.2.2, footnote 113). In contrast to pattern *e* which has the property of rhythmic oddity (sections 3.2.2 and 4.1.3), pattern *b* establishes an attack pair on the half-cycle level. It is thus comprehensible that the highest saliences are estimated for simple meters at the fastest rates, where the isochronous cycle and half-cycle pulses are salient. At an elementary-pulse-period range around $0.2s$ to $0.3s$, the rhythm itself coincides with the most salient pulse, which is non-isochronous in this case. This is exactly what is meant by a rhythm that "acts as a meter".[127]

To return to the issue of balancing the saliences of possible simple and mixed metric interpretations, it is instructive to recall the assumption that "consonant" pulse sensations mutually enhance each other. It was stated in section 2.3.1, that this type of enhancement can explain the bias for the emergence of binary meters induced by isochronous rhythms, and the related tendency to prefer binary over ternary metric frameworks considered in sections 2.3.1 and 2.3.4. The measure of meter salience implicitly reflects this on a general level, due to the following. Duple meters naturally possess a higher temporal density of metric strata than all other meters. More precisely, the ratio between the periods (or rates) of adjacent metric strata can be 1:2 (duple), 1:3 (triple), or both (mixed meters). The temporal range where pulse-period salience is of relevant magnitude (the existence region of pulse sensation, see section 2.2.3), spans a ratio of four-odd octaves (1:16). Thus, emergent metric hierarchies with solely duple relations may afford pulse sensations on up to five levels (1:2:4:8:16), which considerably add to meter salience. The remaining possible hierarchies will all accommodate less metric strata within the span of four octaves. The perception of pulse in pure triple meters is definitely restricted to maximal three levels (1:3:9). Mixed meters and simple

[124]See section 3.2.2 on maximal evenness and "meter rhythms".
[125]cf. sections 3.2.2 (footnote 98) and 4.2.2
[126]cf. 3.2.4
[127]See the corresponding quotes of Huron, 2006, p. 202, in sections 1.3 and 3.2.2. See also the notion of *meter rhythms* in section 3.2.2.

meters including both duple and triple relations between different metric strata perform somewhere in between. The latter may afford pulse sensation on four levels if maximally one triple relation occurs in the metric hierarchy. This corresponds to the meters of cardinality 12 which were examined above (1:2:4:12, 1:2:6:12, 1:3:6:12).

Hence, pure duple meters have better temporal preconditions to gain higher saliences than other meters. Though, to take effect, all metric levels should first be well embodied by the rhythmic pattern to reach high scores of pulse-match salience. To that effect, meter salience is influenced by the metric depth (section 2.2.1) conveyed by a rhythm. Secondly, their periods need to be advantageously distributed within the temporal window of pulse sensation to optimize pulse-period salience.[128] Thereby, the above mentioned and widely acknowledged preference for binary metric frameworks may be adequately reflected. It is nevertheless questionable if this sufficiently accounts for culturally trained biases for certain types of meter. The assumption of mutual enhancement of consonant pulse sensations could additionally be integrated in the measurement of meter salience to introduce a flexible tool for the adaptation to specific listener constitutions. Due to the lack of evidence, it is reasonable to postpone the suggestion of a definite implementation in the computational method outlined so far. It could however be generally assumed that the perceptual salience of isochronous pulses differs from that of pulses which deviate from nominal isochrony (see section 2.2.1). Then, the deviation can be calculated as a factor of pulse salience, essentially independent from a specific weighting. If we claim a preference for isochronous pulses, the factor could be defined as 1 for isochrony, and deviations could have a downscaling effect on this factor to attenuate the salience of non-isochronous pulses. The amount of damping may then be estimated in some way proportional to the amount of deviation.

5.2.5 Metric contrast and metric malleability

The previous investigations show that the metric malleability of a single rhythmic necklace can be related to an extensive potential of metric interpretation. However, the global look at exhaustive spaces of possible metric frameworks obstructs the view on more specific properties of a rhythmic cycle, concerning its malleability related to particular meters.[129] In the following, the focus is therefore narrowed to some characteristic aspects of malleability. More specific, the relations among metric frameworks which score the highest meter saliences according to a rhythmic pattern will be regarded as a key feature of the malleability of that pattern.

[128] This also corresponds to the concept of *attractor tempos* proposed by Gotham, 2015a, which depend on the amount of aggregate pulse salience (cf. sections 2.3.1 and 5.2.3).

[129] For instance, including a large number of different metric frameworks in the calculation of meter-salience distributions, their profiles get less significant in regard to pulse rate (cf. section 5.2.3).

In sections 4.2 and 4.4, formal relations between meters and methods to quantify their distance or similarity are examined. The measure of metric contrast (section 4.4.2) is developed in this context to be employed in a heuristics of metric malleability according to the following general hypothesis: *strong contrasts between salient meters of a rhythmic cycle – that is, metric frameworks highly suggested by that cycle – indicate pronounced metric malleability*. Thus, the measure of metric contrast may contribute a distinctive aspect for the characterization of a rhythm's metric malleability. This aspect may underlie musically striking effects, such as radical changes of the perceptual gestalt when the same rhythmic pattern is cast in contrasting frameworks. That is, correspondent metric gestalt flips (section 3.1) may be the more exciting and surprising, the more contrasting those frameworks are.

The general notion of metric contrast is differentiated in section 4.4.2 into two components: *period contrast* and *phase contrast*. This scheme relates to the perceptual effects which are distinguished in the context of metric-ambiguity types and metric-dissonance types (sections 2.3.3 and 3.2). In particular, metric period contrast reflects KREBS' type A dissonance and metric phase contrast corresponds to his type B dissonance. The two measures provide an elaborated tool to quantify these aspects separately, which also supports the analysis of metric malleability. To that effect, they are applied in the following survey to evaluate and differentiate the above-mentioned hypothesis.

In section 5.2.3, the relation between meter-salience distributions and metric malleability has been examined by means of the fife most salient meters, which are calculated for the rhythmic necklace [1-2-3] for the pulse rates listed in table 5.1.[130] To gain more specific insight into the metric malleability of pattern [1-2-3] pertaining to these meters, estimates of metric period and phase contrast between the ten individual pairs in this group of meters are provided in table 5.4. The bilateral meter relations are indicated by the correspondent metric interpretations listed in figure 1.1, referred to by the bar numbers in the left column.

Three blocks are separated in the table which correspond to three relational categories. The upper part shows three relations including different phases of a 3 x 2 meter (3/4 in figure 1.1), the middle section lists six constellations between 3 x 2 and 2 x 3 meters (3/4 vs. 6/8), and the lowest line represents the relation between the two 2 x 3 (6/8) meters in this group.

As only two metric types are involved, the figures for period contrast are not much differentiated. There is either a certain amount of period contrast between two different metric types (middle section) or no period contrast between two instances of the same

[130] See also figures 5.8, 5.9, and 5.10.

TABLE 5.4: Bilateral period and phase relations between the five metric interpretations in figure 1.1, yielding the highest meter saliences for $e = 0.25s$ (see table 5.1 and figure 5.8)

bar numbers	period contrast	geodesic phase distance (in e)	phase contrast
4 vs. 6	–	2	0.02
4 vs. 2	–	2	0.02
6 vs. 2	–	2	0.02
4 vs. 10	0.57	0	–
4 vs. 7	0.57	3	0.15
6 vs. 10	0.57	2	0.14
6 vs. 7	0.57	1	0.27
2 vs. 10	0.57	2	0.14
2 vs. 7	0.57	1	0.27
7 vs. 10	–	3	0.02

metric type. According to the definitions in section 4.4.2, period contrast is partly determined by the absolute durations of contrasting metric periods.[131] Table 5.4 presents values for $e = 0.25s$, a duration in the region of pulse periods where the distribution of meter saliences, estimated for the fife meters under consideration, is maximally flat (see section 5.2.3). This naturally coincides with the fact that metric period contrast between 3 x 2 and 2 x 3 meters reaches its maximum of about 0.57 around $e = 0.25s$. At this rate, the contrasting metric periods – presumably the tactus levels of both meters – are around $0.5s$ and $0.75s$, thereby optimizing their aggregate pulse-period salience.[132]

Concerning metric phase contrast, the example bears a more subtle gradation, as six different phases can be distinguished according to figure 1.1 in section 1.1. Due to the cyclicality of the patterns, their relations can however be expressed in only four different geodesic phase distances (see third column in table 5.4), for which different figures of phase contrast can be distinguished according to the definitions in section 4.4.2. Geodesic phase distance between two metric frameworks corresponds to the shortest temporal shift[133] of one of the two meters, which is necessary to get them congruent. All phase distances between two frameworks of equal metric type, included in table 5.4, are equal to the period of the tactus level of that metric type ($2e$ or $3e$). Therefore, phase contrast is estimated very low (0.02) between these pairs, as a phase difference only occurs on the highest level, that is, the bars are shifted to each other whereas the phase of the tactus level is preserved, or shared among all mentioned cases.[134] A look at the estimates of phase contrast between two different metric types (see the middle section of table 5.4) reveals a very fine gradation between the phase distances $2e$ and

[131] More precisely, the pulse-period saliences of contrasting periods are used as weighting factors for their period – and also phase – contrast (see equations 4.2 and 4.3). Basically, perceptually less salient pulse periods are assumed to convey less impact on metric contrast, both for the aspects of period and of phase.

[132] cf. sections 2.3.1 and 5.2.3

[133] The temporal interval of this shift is measured in elementary pulses.

[134] This can be related to the marginal differences of meter salience, calculated for the instances in figure 5.5 (see table 5.2), as the pulse-period salience of higher (or "slower") metric strata is assumed to be relatively low.

$3e$ (0.14 versus 0.15).[135] Phase distance $1e$ yields the highest score of phase contrast among the listed relations. This is easily reproducible because it results in crosswise phase shifts between the tactus level of one meter and the measure level of the other (and vice versa).

Based on this analysis, the overall picture of the metric malleability of pattern [1-2-3] can be enriched by aspects of metric contrast. Among the most salient metric frameworks, the major contrasts are obviously found between two meters of different type (3 x 2 vs. 2 x 3). Beside period contrast on the tactus level, those pairs feature several intensities of phase contrast. Among the interpretations exemplified in figure 1.1, the most conflictive meter relations are thus found between bars 6 versus 7, and bars 2 versus 7. Consequently, the metric malleability of the [1-2-3] pattern is reflected most peculiarly by these relations, as the respective metric interpretations create very different perceptual shapes for the same pattern. The involved frameworks are salient meters of this pattern, and thus, the intersections of their meter-salience functions, related to pulse rate in figure 5.8, specify attractor tempos for metric malleability (section 5.2.3), itemized for these peculiar relations: the meter-salience curves of the frameworks shown in bars 6 and 7 of figure 1.1 cross at $e \approx 0.22s$, and the interpretations of bars 2 and 7 intersect at $e \approx 0.24s$.

In sum, the figures for metric contrast in table 5.4 indicate that, among the relations of salient meters of pattern [1-2-3], a double contrast (of metric period and phase) correlates with increased metric phase contrast. However, estimates of phase contrast are much higher for many meter relations among the twelve frameworks displayed in figure 1.1. Yet, they apply either to relations between one of the fife salient meters and a less salient meter of the same metric type, or to relations between two (as well equal-typed) meters of little salience. In particular, among the 3 x 2 frameworks, the figure of metric phase contrast is 1.13 for all meter relations holding phase distances of $1e$ or $3e$. Regarding 2 x 3 meters, phase contrast is estimated at 1.06 for all relations featuring phase distances of $1e$ or $2e$.[136]

This again exemplifies a more general aspect, already described in section 4.4.1. A metric framework, related to a phase-shifted version of itself, produces high or low phase contrast, depending on the relation of the phase distance to the subcycle periods of the meter. The more metric strata are shifted against each other, the higher the metric phase contrast. As mentioned in section 4.4.1, the concept of *pivot pulse*, suggested by OSBORN,[137] can indicate the amount of perceptual disruption caused by a metric shift. The pivot pulse corresponds to the highest metric stratum which is preserved

[135] Part of the reason is that phase distance $2e$ keeps metric periods in relation $2e$:$6e$ in phase, whereas phase distance $3e$ keeps metric periods in relation $3e$:$6e$ in phase. This leads to very similar overall measures. For details see the definition of metric phase contrast in section 4.4.2.

[136] These figures are as well calculated for $e = 0.25s$, like those in table 5.4.

[137] Osborn, 2010, see also Benadon, 2004 about the role of pivot pulses for the organization of metric modulation or “tempo modulation” (cf. section 4.4).

in a change of the metric framework. Correspondingly, the highest phase contrasts are estimated when only the elementary pulse level is preserved, as in the relations above, featuring phase distance $1e$. As reported in section 2.2.1, it has been experimentally established that in the context of metric entrainment, the adaption to changes of a metric period is slower, and thus more intricate than adapting to a phase shift of that period. This may have impact on the musical effects of metric shift, or metrical gestalt flip. For instance, a metric shift between the frameworks shown in bars 4 and 10 of figure 1.1 may be harder to follow than a respective shift between bars 4 and 6. The former involves period contrast (or type A dissonance), whereas the latter involves phase contrast (or type B dissonance) between the respective two metric frameworks (sections 2.3.3 and 4.4.2). However, a change between bars 4 and 3 may be highly challenging as well, as the phase distance is $1e$ and the pivot pulse equals the elementary pulse. This is properly reflected in the huge differences between the mentioned phase contrast estimates (1.13 for bars 4 versus 3, and 0.02 for bars 4 versus 6).

More complex patterns which allow for simple and mixed metric frameworks, like the metric types of the 12 cycle involved in the examination in section 5.2.4, require a more extensive analysis of relations between metric contrast and metric malleability. In the present example, the aspect of phase contrast needs more discussion than period contrast, as there are only two metric types involved. A look at table 5.3 recalls the rising number of metric framework types for longer metric cycles, and therefore, the rising complexity of meter relations involving both aspects of metric contrast. Approving the preliminary state of the heuristics put forward in this thesis, comprehensive analyses may be postponed to subsequent studies. However, from the present discussion we can get an idea of the analytic paths to go about relating the metric malleability of more complex patterns to aspects of metric contrast. In the following section 5.3, an exemplary matrix of parameter settings is used to explore such relations for the class of cyclic rhythms defined by the class of necklaces $n = 12$, $k = 2$ with fixed content $n_0 = 5$, $n_1 = 7$ (see section 4.1.3). This examination is designed to exemplify one of many possible heuristics to approach metric malleability.

5.3 An exemplary application for two related classes of cyclic rhythms

The analytic devices developed in the previous section 5.2 can serve as a basic equipment for a more extensive exploration of the metric ambiguity and malleability of complex rhythmic patterns. The concept of a flexible quantitative heuristics seems to be particularly adequate to tackle the combinatoric universe of possible subjective relations between rhythmic and metric cycles of higher cardinalities. Metric malleability may also be a driving motive for processes of rhythm selection or generation, for instance in the context of compositional strategies, or the evolution of rhythms in musical traditions.[138] It is therefore of interest to explore and characterize metric malleability in the context of classes of cyclic rhythms, to establish differentiated criteria for rhythmic composition and selection according to specific properties. The subsequent examination thus heuristically generates rankings of individual patterns according to precisely defined features, as the main parameters of metric malleability – meter salience, metric contrast, and pulse rate – are exemplarily set into relation by means of the following question.

> For which rhythmic cycle of pulse cardinality n and k onsets at an elementary-pulse duration e, maximal metric contrast is found among the m most salient meters of the cycle?

According to the model defined so far, this problem is explored for the class of cyclic rhythms defined by the class of binary necklaces $n = 12$ (alphabet size $= 2$) with fixed content $n_0 = 5$, $n_1 = 7$, where n_0 represents onsets and n_1 represents rests, or vice versa (see section 4.1.3). Thus, two complementary classes of rhythmic 12-pulse cycles are examined, either including five or seven onsets. Most of the rhythmic patterns in figure 5.11 (section 5.2.4) belong to these two classes, as well as other popular patterns like the variants of the 12-pulse flamenco pattern *guajira* [x..x..x.x.x.] (figure 5.11 b),[139] its metric rotations *seguiriya* [x.x.x..x..x.] and *soleá* [..x..x.x.x.x], as well as *bulería* [..x...xx.x.x].[140] As already discussed in section 3.2.2, the above-defined constraints also include the two "standard" patterns, widely-used as "ternary rhythm timelines in sub-Saharan Africa and the Caribbean".[141] These ubiquitous patterns, isomorphic to the pentatonic and the diatonic scales (see also figure 5.11 c and f), play as well a predominant role in the present examination. The present approach verifies that they afford contrasting metric frameworks of similarly high salience, which is demonstrated in the following.

The study comprises 36 calculations according to different settings of the variables defined in the context of the problem stated above. Each results in a ranking of necklaces in order to estimate the pattern that solves the problem, or represents the answer to that question. The number of settings arises from the differentiation of the two mentioned classes of rhythmic cycles, subsequently denoted as $12/5$ (five onsets) and $12/7$ (seven onsets), and three different elementary-pulse durations ($e = 0.15s$, $e = 0.22s$, and $e = 0.3s$). For either of the six resultant temporal patterns, a set of possible frameworks, including all well-formed metric 12 cycles, is ranked according to the salience

[138] cf. for instance Colannino, Gómez, and Toussaint, 2009, Toussaint, 2010, and Toussaint, 2013, pp. 265 ff.

[139] cf. the other references in sections 3.2.2 (footnote 98), 3.2.4 ("hemiola pattern"), and 4.2.2.

[140] cf. Toussaint, 2013, pp. 266 ff.

[141] Toussaint, 2013, p. 45

of each meter in the set in relation to the rhythm.[142] As mentioned in section 4.2.2 and section 5.2.4, the 12 cycle allows for 17 meter types (three simple and 14 mixed, see table 5.3), and thus maximally 204 metric frameworks for each necklace.[143] Further estimations are based on three different subsets of the most salient meters, derived from the meter-salience ranking ($m = 3$, $m = 24$, and $m = 192$). For each pair of meters in a subset, their metric contrast is calculated in order to find the most contrasting constellation.[144] The motivation to vary the size of the subsets is to focus either on maximal meter salience (by means of a small subset like the three most salient meters), or on maximal meter contrast (using a big subset like $m = 192$, that is, most of the set of the 204 frameworks). A middle-sized subset ($m = 24$) represents a compromise between both, that is, the aspects of meter salience and metric contrast are balanced. Finally, for each of the 18 constellations combined so far,[145] two solutions are differentiated by performing the described calculation either as a function of metric period contrast, or according to metric phase contrast.

5.3.1 Prevalent instances of malleable rhythmic necklaces

The following description of the results illustrates the metrically ambiguous and malleable character of the necklaces which are favored by the procedure. From each class of 66 necklaces, eight (of 18 possible) different necklaces reach the top of the rankings. Three, respectively two necklaces are favored more than once. As indicated before, the predominant necklaces in this respect are the two "standard" patterns, that is, [2-2-3-2-3] among the $12/5$ class, and [1-2-2-1-2-2-2] among the $12/7$ class. As discussed before, these two patterns are strongly interrelated.[146] They are (1) complementary,[147] (2) the latter includes the former,[148] and (3) both are maximally even (section 3.2.2). The [2-2-3-2-3] pattern also has a well-formed metric structure (see section 4.2.2). The results confirm the prominent position of these well-known and often discussed patterns in

[142] No other meters of cardinalities which stand in simple ratio to twelve are taken into account for economic reasons, in spite of the suggestion in section 5.2.4.

[143] Cyclic rhythms which are represented as prime necklaces or Lyndon words (see section 4.1.3) afford 204 frameworks, as they can be metrically rotated in twelve variants in relation to each of the 17 meter types (metric phases). Non-prime necklaces result in less rotational variants, respectively. Due to the relative primeness of 12 and 5 (respectively 12 and 7), all necklaces of the class, chosen for the current study, are prime.

[144] Thus, for the subsets of 3, 24 and 192 meters, 3, 276 and 18336 contrast values are respectively estimated for each necklace in the context of each of the six constellations.

[145] 18 constellations result from the combination of two necklace classes with three pulse rates and three subset sizes.

[146] cf. sections 3.2.2, 4.1.1, and 5.2.4.

[147] Note that every necklace of the $12/5$ class has a complement in the $12/7$ class. Both classes therefore have the same number of elements, that is, 66 necklaces.

[148] More precisely, the [2-2-3-2-3] pattern is a *shelling rhythm* of the [1-2-2-1-2-2-2] pattern in the sense declared in section 4.1.1.

music. They feature pronounced metric malleability which will be specified by the following details of the study. This corresponds to the musicological findings about the *robustness* of these necklaces,[149] discussed in section 4.1.3.

To specify the overall outcome of the estimations relating to the $12/5$ class, we first take a look at the [2-2-3-2-3] necklace which is favored seven of 18 times. It is top-ranked according to both maximal period contrast and maximal phase contrast with $m = 192$ at any of the three pulse rates (six times), and additionally for maximal period contrast, with $m = 24$ and $e = 0.3s$. In the six cases with $m = 192$, less salient meters drive the choice in favor of higher metric contrast. Figure 5.12 illustrates these cases for maximal period contrast at pulse durations $e = 0.15s$, $e = 0.22s$, and $e = 0.3s$ (from left to right, where the right diagram also holds for $m = 24$). In this figure, as well as in the following ones, rhythmic necklaces are graphed as introduced in sections 1.1 and 4.1.3. Black beads stand for rhythmic onsets and white beads represent rests or continuations. They are flanked by an inner and an outer circle, representing two alternative metric frameworks for the rhythm. The diagrams are designed for an easy comparison, and to reproduce mental switching between alternative metric frameworks. In this way, metric malleability and metric contrast can be evaluated by perception. Metric strata are represented by the size of the beads, and additionally, the digits show indices identifying the metric strata according to the GNSM code (see section 4.2.3). The elementary-pulse level is "zero" (not marked) and all higher interpretive strata (subcycles or metric groupings) are indexed with rising cardinalities. In the case of the 12 cycle, a 3 denotes the highest level, that is, the cycle pulse. In figure 5.12, always two meters of a different type class are combined, one simple and one mixed metric type. Such constellations often cause strong metric period contrasts. Further explanation will be given later in comparison to figure 5.16.

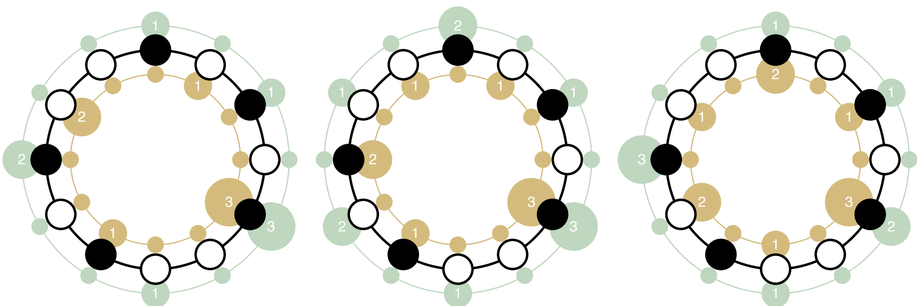


FIGURE 5.12: Rhythmic necklace [2-2-3-2-3] features maximal period contrast with $m = 192$ at elementary-pulse periods of $0.15s$ (left diagram), $0.22s$ (middle diagram), as well as with $m = 24$ and $m = 192$ at $e = 0.3s$ (right diagram).

[149] Toussaint, 2013, pp. 77 ff., lists occurrences of all onset rotations of the two standard patterns in the context of several musical traditions.

In contrast, maximal phase contrast is found with $m = 192$ for the two constellations shown in figure 5.13. Both feature a simple metric framework which is rotated in two alternative ways related to the rhythmic cycle. At a pulse duration of $0.15s$ (left diagram in figure 5.13), a 2 x 2 x 3 meter (300100200100) competes with a rotation of itself. For $e = 0.3s$, the same constellation as for $e = 0.22s$ was chosen (right diagram in figure 5.13). Here, a 3 x 2 x 2 meter (301020102010) competes with its own rotation. In both constellations, the metric phase distance is $5e$, which is relatively prime to the 12 cycle. As discussed in the previous section 5.2.5, this yields maximal phase contrast because all metric strata except the elementary pulse level are shifted against each other. The choice of these two different meters for different pulse rates is driven by the interplay between the pulse-rate dependencies of metric phase contrast, and of their individual meter saliences.

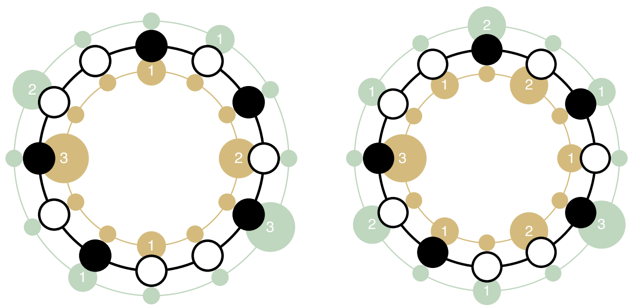


FIGURE 5.13: Rhythmic necklace [2-2-3-2-3] affords maximal phase contrast with $m = 192$ for elementary-pulse periods of $0.15s$ (left diagram), as well as $0.22s$ and $0.3s$ (right diagram).

Two other necklaces play a prominent role according to the results of the examination of the $12/5$ class: the pattern [2-2-2-3-3], extensively discussed before in the context of flamenco and hemiola,[150] and corresponding to a well-formed metric pattern, as well as the pattern [1-3-3-1-4], including the metric pattern [4-4-4]. Each is top-ranked for three different parameter settings. Figure 5.14 illustrates these cases for pattern [2-2-2-3-3]. In both constellations, a simple meter is confronted with a mixed meter. Due to the rather small set sizes m, the meter saliences are high and the contrast values are rather low.

In figure 5.15, the constellations for the pattern [1-3-3-1-4] are shown. The simple 3 x 2 x 2 meter (301020102010) is involved in both constellations, confronted with mixed meters (3+2+2)+(2+3) (300101020100, left diagram) and (3+2)+(2+2+3) (300102010100, right diagram) which have a parallel structure of beats but different half-cycles (7+5 versus 5+7).

[150] cf. sections 3.2.2 (footnote 98), 3.2.4, 4.2.2, 5.2.4, and 5.3.

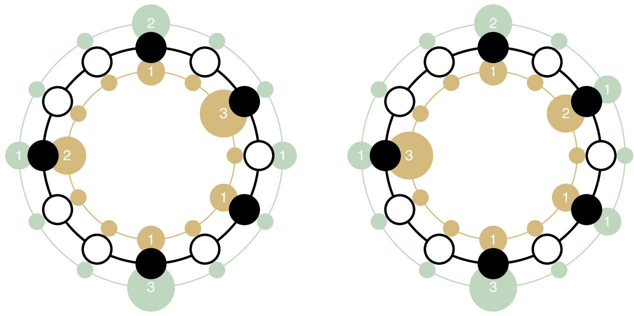


FIGURE 5.14: Rhythmic necklace [2-2-2-3-3] allows for maximal period contrast with $m = 24$ at elementary-pulse period $0.15s$ (left diagram), and for maximal phase contrast with $m = 3$ at elementary-pulse periods $0.22s$ and 0.3 (right diagram).

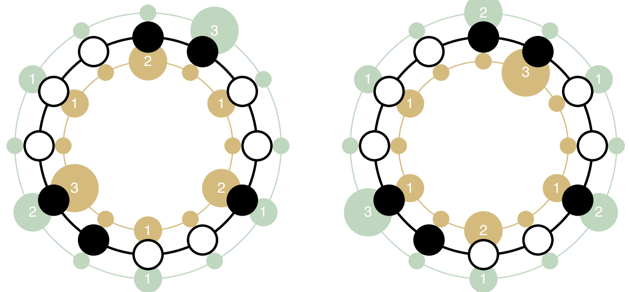


FIGURE 5.15: Rhythmic necklace [1-3-3-1-4] features maximal period contrast with $m = 3$ at elementary-pulse periods $0.22s$ and 0.3 (left diagram), as well as with $m = 24$ at elementary-pulse period $0.22s$ (right diagram).

Five other necklaces are favored once according to one of the remaining configurations for the $12/5$ class. The constellations in which they are involved are described in appendix E and illustrated in figure E.1. In contrast to the prevalent necklaces discussed so far, these five necklaces are all *chiral* (see section 4.1.3), that is, there exists another correspondent necklace representing the mirror image of each, which is equivalent to the rhythmic retrograde.[151] It may be worthwhile to compare the scores of necklaces related in this way to gain insight into the relation of symmetry and metric malleability.

Henceforth, the evaluation is continued with a description of the results related to the $12/7$ class. Similar to the standard pattern of class $12/5$, [2-2-3-2-3], the seven-onset standard pattern [1-2-2-1-2-2-2] is always top-ranked with $m = 192$ at any pulse rate, affording both maximal period and phase contrast. Additionally, it allows for maximal phase contrast with $m = 3$ at elementary-pulse durations $0.22s$ and $0.3s$, and for maximal period contrast with $m = 24$ at pulse duration $0.3s$. Thus, in sum it is favored nine times, that is, for 50 % of all 18 parameter configurations. Figure 5.16 shows the cases

[151] In other words, two correspondent chiral necklaces belong to the same *bracelet* (cf. section 4.1.3).

of maximal period contrast with $m = 192$, from left to right, for $e = 0.15s$, $e = 0.22s$, and $e = 0.3s$.

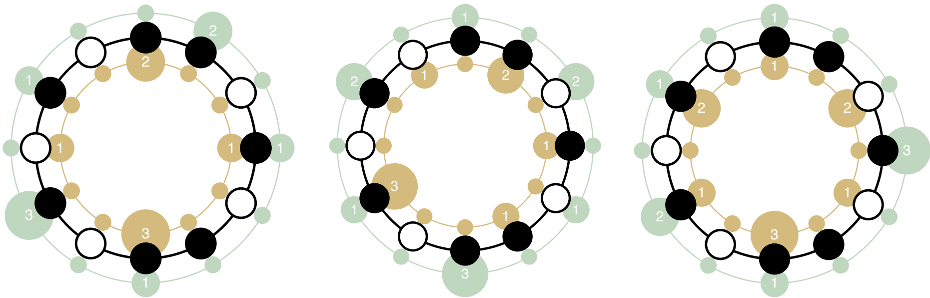


FIGURE 5.16: The standard pattern [1-2-2-1-2-2-2] allows for maximal period contrast with $m = 192$ (left to right: elementary-pulse periods of $0.15s$, $0.22s$, and $0.3s$)

From the present figure, extensive parallels can be drawn to the constellations involving pattern [2-2-3-2-3] in figure 5.12. Like there, one simple and one mixed metric type is combined in all three cases. In the left diagram ($e = 0.15s$), the simple 2 x 2 x 3 meter occurs in parallel to the left diagram in figure 5.12, though the mixed meters are different. In the middle diagram ($e = 0.22s$), both meters are the same as in the middle diagram of figure 5.12 (3 x 2 x 2 and (3+2)+(2+2+3)). Finally, a comparison of the right diagram ($e = 0.3s$) to the right diagram in figure 5.12 reveals again the same simple meter (3 x 2 x 2), as well as two mixed meters, exhibiting the same beat structure (3+2+2+2+3) but different half-cycles (7+5 versus 5+7). These two meters also occur in figure 5.15, as discussed before.

In respect to maximal phase contrast, the same constellations are found for [1-2-2-1-2-2-2] (see figure 5.17) and [2-2-3-2-3], as already described in detail in connection with figure 5.13. A simple meter is confronted with a rotation of itself, yielding metric phase distance $5e$ between the two alternative interpretations. Also, for $3 = 0.3s$ the same constellation as for $e = 0.22s$ gains maximal phase contrast with $m = 192$ (right diagram in figure 5.17).

The [1-2-2-1-2-2-2] pattern is favored for three other parameter settings, yielding two different meter constellations which are illustrated in figure 5.18. Here, the left diagram shows a delicate phase shift between two similar mixed meters, representing metric-phase contrast between the cycle and the half-cycle strata, though exhibiting exactly parallel beat structures in relation to the rhythmic cycle. The other type of relation between a simple and a mixed meter, graphed in the right diagram, is again typical for high period contrast.

To summarize, the prominence of the pattern [1-2-2-1-2-2-2] in manyfold musical contexts which is analytically validated in numerous studies,[152] may also be motivated

[152]Some of them are reported and discussed in sections 3.2.2, 3.2.3, 4.1.1, and 4.1.3.

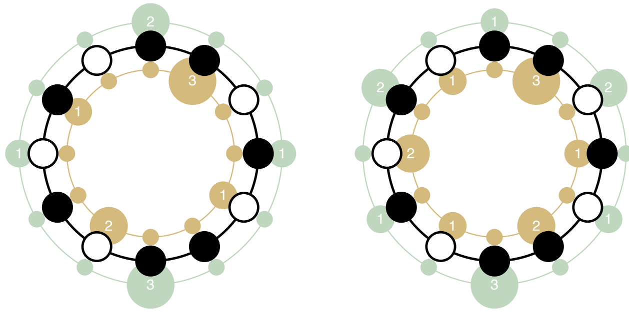


FIGURE 5.17: The pattern [1-2-2-1-2-2-2] features maximal phase contrast with $m = 192$ (left diagram: elementary-pulse period $0.15s$, right diagram: elementary-pulse periods $0.22s$, and $0.3s$)

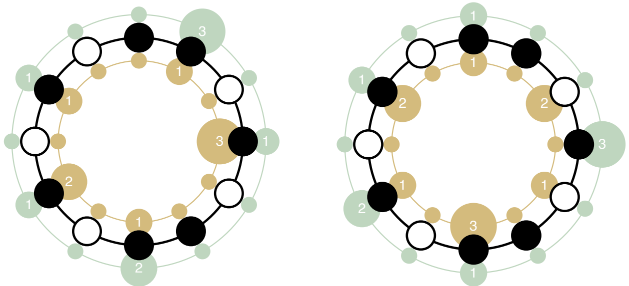


FIGURE 5.18: The pattern [1-2-2-1-2-2-2] affords maximal phase contrast with $m = 3$ at elementary-pulse periods $0.22s$ and 0.3 (left diagram), as well as maximal period contrast with $m = 24$ at elementary-pulse period $0.3s$ (right diagram).

by its outstanding metric ambiguity and malleability. This fact has been indicated by several authors, as already quoted in section 3.2.2: "the standard pattern is almost maximally ambiguous, as it samples several different meters (12/8, 6/4, and 3/2) – and different phases of those meters – almost equally."[153] The present approach substantiates and differentiates this potential by quantitative means. Moreover, the whole dimension of metric malleability can be systematically accessed for any pattern on a more general basis.

Among the 12/7 class, one other necklace apart from the standard pattern is top-ranked in this evaluation for more than one parameter configuration. The pattern [1-2-1-2-2-1-3] yields maximal scores for three settings, resulting in two different configurations which are illustrated in figure 5.19. Again, in the left diagram, a delicate phase shift between two similar mixed meters yields metric phase contrast between the cycle and the half-cycle strata, but parallel beat structures related to the rhythmic cycle. However, in this case, this constellation yields maximal period contrast. The right diagram displays another instance of maximal phase contrast, this time between the simple meter 3 x 2

[153] see footnote 135 in section 3.2.2

x 2 and the mixed meter (2+2+2)+(3+3), where both shifted to each other by phase distance $5e$. Note that the mixed meter corresponds to the pattern [2-2-2-3-3], establishing an *intermediate level*,[154] – that is, an isochronous pulse – at the half-cycle.

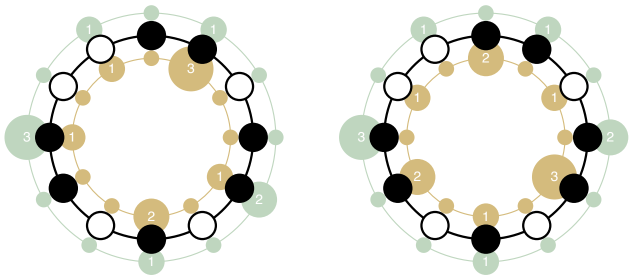


FIGURE 5.19: Pattern [1-2-1-2-2-1-3] features maximal period contrast with $m = 3$ at elementary-pulse periods $0.22s$ and 0.3 (left diagram), and maximal phase contrast with $m = 24$ at elementary-pulse period $0.3s$ (right diagram).

The most contrasting constellations of the six other necklaces, which are favored once in respect to one of the remaining parameter settings for the $12/7$ class, are described in appendix E and displayed in figure E.2.

5.3.2 General observations and concluding remarks

We conclude this examination with some more general observations. At faster rates, more different necklaces gain the leading position, allowing maximal contrast between two possible metric frameworks, than at slower rates. Table 5.5 illustrates this variety. Each element in the table represents six calculations for each class of rhythmic cycles and for each elementary-pulse rate. They result from the differentiation of the two types of metric contrast (period and phase), and the three different set sizes m. The figures indicate the number of different necklaces which are favored across the six configurations.

TABLE 5.5: Numbers of favored necklaces as a function of pulse rate and necklace class (for further explanation see text)

	$e = 0.15$	$e = 0.22$	$e = 0.3$
necklace class $12/5$	5	4	4
necklace class $12/7$	5	4	2

Some of the more prominent necklaces described before are favored in any pulse-rate condition, in particular, the two standard patterns [2-3-2-2-3] (see figures 5.12 and 5.13)

[154] cf. section 4.2.2

and [1-2-2-1-2-2-2] (figures 5.16, 5.17, and 5.18), as well as the pattern [2-2-2-3-3] (figure 5.14). Other prevalent patterns are by trend favored at slower rates, namely, the necklaces [1-3-3-1-4] (see figure 5.15) and [1-2-1-2-2-1-3] (figure 5.19).

It is striking that all discussed patterns feature a relatively high degree of temporal dispersion of onsets. Some of them exhibit clusterings of two, or at most three onsets on adjacent elementary pulses, but not more. Others are actually maximally even, as mentioned before. Most of the favored patterns in this examination are *balanced*, a property which has been discussed by TOUSSAINT in the context of maximal evenness.[155] It can be illustrated by means of the circular graph of a necklace and its diameter, imagined as a line, rotating around the center

> through a full revolution until it returns to its starting orientation. If for every position of the line such that it is not incident on an onset, the number of onsets on one side of the line differs by at most one from the number of onsets on the other side, then the rhythm is called *balanced*.[156]

In other words, every arbitrary division of a rhythmic necklace into half-cycles balances the number of onsets between the two parts. Only four of the 16 different necklaces, which are top-ranked according to the 36 parameter configurations, are not balanced: the patterns [3-2-1-2-4] (figure E.1, top-left), [1-1-3-3-4] (figure E.1, bottom-middle), [2-2-1-1-2-1-3] (figure E.2, top-middle), and [1-2-1-1-2-1-4] (figure E.2, bottom-right). However, the degree of evenness of balanced patterns can be very different. Compare for instance the maximal even necklace [2-3-2-2-3] with the rather uneven necklace [1-3-3-1-4] (see figure 5.15). According to what is said above about the rankings as a function of pulse rate, less even necklaces tend to be favored at slower rates. Though, regarding the patterns in figures E.1 and E.2, this is less consistent. It would need a more comprehensive examination, targeted to this issue, to establish a clearer picture. At least, the tendency for balanced necklaces discards patterns where onsets cluster to an extend that greater inter-onset intervals than $4e$ occur, which is the maximum IOI found among the selected necklaces. Interestingly, patterns with clustered onsets exhibit as well a high degree of metric ambiguity. FLANAGAN calculated GINI coefficients of all 12-pulse necklaces for different pulse rates, in this case, indicating the salience distribution of metric beats with durations that are factors of the cycle duration (see sections 4.3.1, 5.2.1, and 5.2.2). He estimated the lowest values for "those formed by one of the generator cycles of the modulo 12 system, the 5- (or 7-) cycle, the generator of the diatonic scale."[157] Pattern and scale generation is discussed in section 4.1.1, where exactly these generator cycles are exemplarily mentioned. Thus, the generated patterns [2-3-2-2-3] and [1-2-2-1-2-2-2], the predominant results according to the problem defined in the present examination, as well give rise to the most equal

[155]Toussaint, 2013, p. 125
[156]ibid.
[157]Flanagan, 2008, p. 639

salience distribution among all pulses which integrally divide the cycle. FLANAGAN also found that when pulse-match saliences are calculated (according to his model described in section 5.2.1) by applying equal values for the phenomenal accents of all onsets, correspondent patterns of generator cycles 5 and 1 gain also equal values for their GINI coefficients.[158] That is, for instance, the 5-cycle generated, maximally even pattern [2-3-2-2-3] and the correspondent 1-cycle generated, minimally even, unbalanced pattern [1-1-1-1-8] are both metrically ambiguous, though, in a very different way. Whereas the former is metrically complex and malleable, the latter is metrically shallow and indifferent below the cycle level. To that effect, FLANAGAN states that

> maximal ambiguity is hardly a guarantor of musical interest. [...] One possible reason is that while bc sets[159] generated from the 1-cycle are metrically ambiguous, their grouping structure is all too obvious. The maximally even dispersion of onsets in the 5-cycle generated sets creates the potential for ambiguity of grouping as well as metrical structure, allowing grouping and metrical structures to interact in interesting ways. Alternatively, perhaps the maximally even sets are more interesting because they constantly flip the beat around, forcing the listener to retrospectively reevaluate the location in the metrical grid of what has already been heard.[160]

Thus, a quantified heuristics for the differentiation of metric malleability, as demonstrated in the present examination, can provide more eclectic insight into features of musical interest, compared to one-dimensional estimates of metric ambiguity. The examples in figures 5.12 to 5.19, as well as in figures E.1 and E.2 include many constellations implying musically challenging possibilities for metric shifts or metric gestalt flips, specifically, due to high contrasts between two alternative metric frameworks represented in the diagrams. It may nevertheless be instructive to define a global score for the metric malleability of a cyclic rhythm. However, regarding the variety of aspects (meter-salience distribution, metric contrast, influence of pulse rate) which are distinguished in the course of this study, it seems to be more promising to focus on individual features involving only pairs or groups of meters. The general observations described in this section – (1) greater variance of favored necklaces at faster rates, (2) individual conditions on the prevalence of particular necklaces, and (3) the tendency for balanced patterns and even onset-distributions – may then be studied as a function of the interaction between the above-mentioned aspects (meter salience, metric contrast, and pulse rate). However, the heuristics underlying the present examination merely represents one of many possible approaches to further immerse into the complexity of

[158] Flanagan, 2008 relates this finding to the "M operation" in musical set theory, under which "the 5 cycle transforms into the 1-cycle" (ibid.).

[159] Flanagan, 2008 employs the notion of beat-class set types, known from musical set theory, to refer to rotational invariance of a cyclic rhythm, that is, a rhythmic necklace.

[160] ibid.

metric malleability. Other perspectives can be developed by different problem definitions which are only limited by issues of manageability and computational cost.

Appendix A

Experimental data

Figures A.1 and A.2 reproduce the results of PARNCUTT's experiment on pulse salience.[1] Subjects tapped along six rhythmic cycles, presented in six different event rates (rhythmic attacks per minute, see figure 2.10 in section 2.2.3). In each panel, the variants of pulse responses (which are assumed to correspond to the perceived tactuses) are listed leftmost in terms of period and phase related to the elementary pulse and the cycle length of the stimuli. More to the right they are shown in musical notation related to the linear progressions of the stimuli. The rightmost columns show "the number of times each pulse response was selected (plain text). The results are compared with predictions according to the model (italics)."[2]

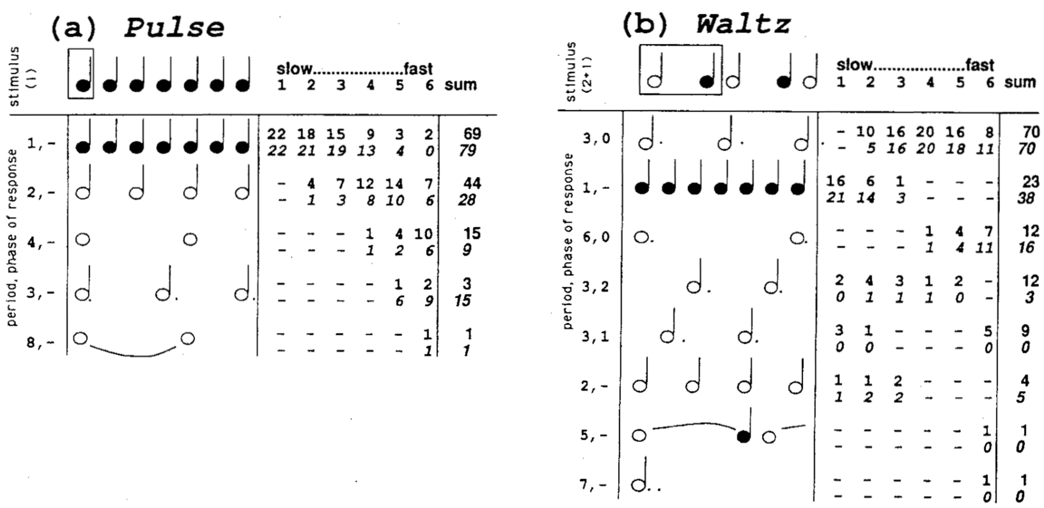
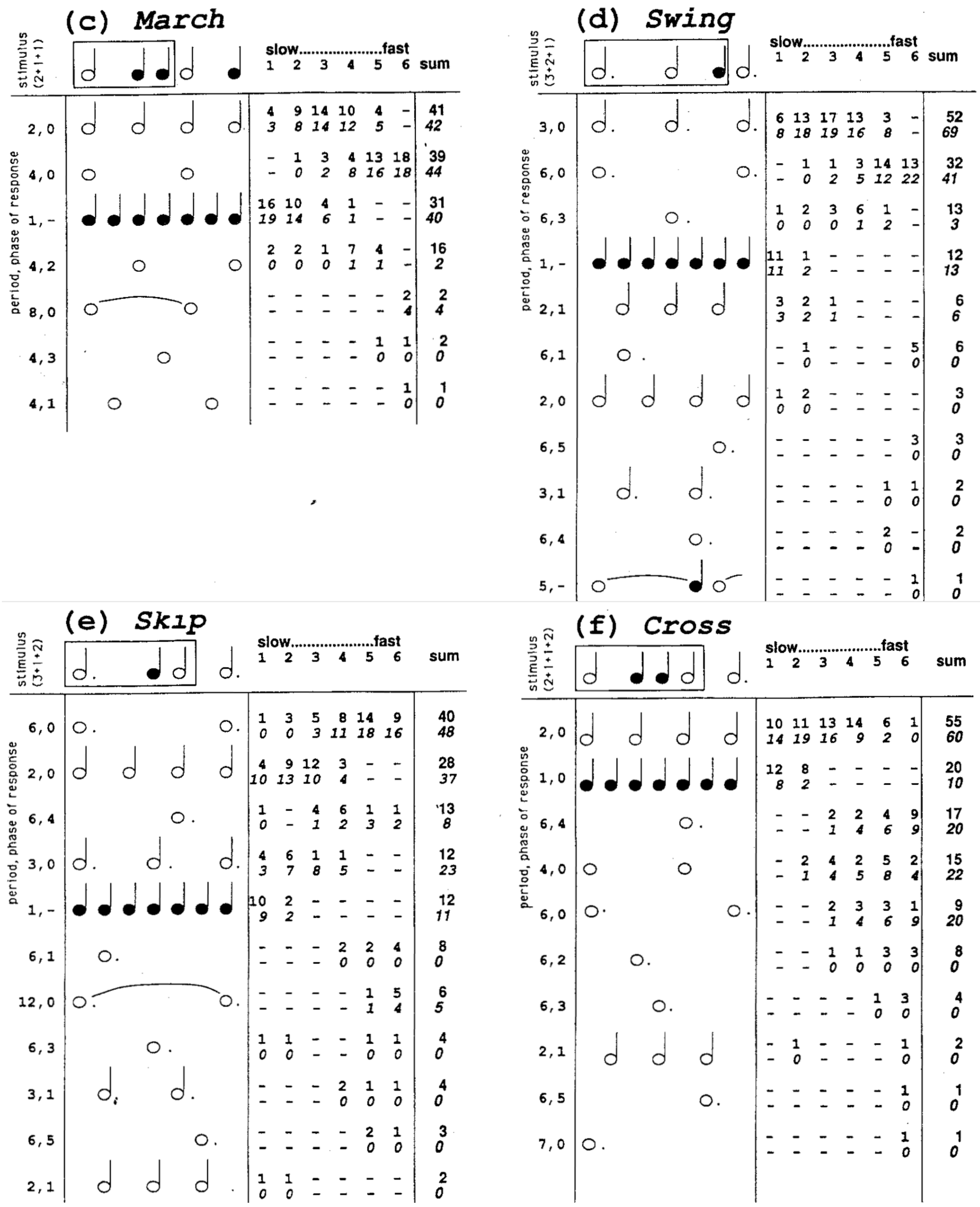


FIGURE A.1: Parncutt, 1994, Figure 2(a) and 2(b) on p. 416 (for Figure 2(c) - 2(f) see figure A.2). Figure 2 displays the results "of Experiment 1 (pulse salience) for all 161 selected pulses. In the table of the right side of each panel [...] Each column corresponds to a particular combination of rhythmic pattern and note rate, or to one of the 36 trials in the experiment." (figure caption, see figure A.2 for continuation)

[1]Parncutt, 1994, pp. 413 ff.

[2]Parncutt, 1994, p. 416. For the explanation of the mentioned model, which is also an important basis for the model of metric malleability proposed in this study (section 5.2), see pp. 423 ff.



(c) March — stimulus (2+1+1)

period, phase of response	1	2	3	4	5	6	sum
2,0	4	9	14	10	4	-	41
	3	8	14	12	5	-	42
4,0	-	1	3	4	13	18	39
	-	0	2	8	16	18	44
1,-	16	10	4	1	-	-	31
	19	14	6	1	-	-	40
4,2	2	2	1	7	4	-	16
	0	0	0	1	1	-	2
8,0	-	-	-	-	-	2	2
	-	-	-	-	-	4	4
4,3	-	-	-	-	1	1	2
	-	-	-	-	0	0	0
4,1	-	-	-	-	-	1	1
	-	-	-	-	-	0	0

(d) Swing — stimulus (3+2+1)

period, phase of response	1	2	3	4	5	6	sum
3,0	6	13	17	13	3	-	52
	8	18	19	16	8	-	69
6,0	-	1	1	3	14	13	32
	-	0	2	5	12	22	41
6,3	1	2	3	6	1	-	13
	0	0	0	1	2	-	3
1,-	11	1	-	-	-	-	12
	11	2	-	-	-	-	13
2,1	3	2	1	-	-	-	6
	3	2	1	-	-	-	6
6,1	-	1	-	-	-	5	6
	-	0	-	-	-	0	0
2,0	1	2	-	-	-	-	3
	0	0	-	-	-	-	0
6,5	-	-	-	-	-	3	3
	-	-	-	-	-	0	0
3,1	-	-	-	-	1	1	2
	-	-	-	-	0	0	0
6,4	-	-	-	-	2	-	2
	-	-	-	-	0	-	0
5,-	-	-	-	-	-	1	1
	-	-	-	-	-	0	0

(e) Skip — stimulus (3+1+2)

period, phase of response	1	2	3	4	5	6	sum
6,0	1	3	5	8	14	9	40
	0	0	3	11	18	16	48
2,0	4	9	12	3	-	-	28
	10	13	10	4	-	-	37
6,4	1	-	4	6	1	1	13
	0	-	1	2	3	2	8
3,0	4	6	1	1	-	-	12
	3	7	8	5	-	-	23
1,-	10	2	-	-	-	-	12
	9	2	-	-	-	-	11
6,1	-	-	-	2	2	4	8
	-	-	-	0	0	0	0
12,0	-	-	-	-	1	5	6
	-	-	-	-	1	4	5
6,3	1	1	-	-	1	1	4
	0	0	-	-	0	0	0
3,1	-	-	-	2	1	1	4
	-	-	-	0	0	0	0
6,5	-	-	-	-	2	1	3
	-	-	-	-	0	0	0
2,1	1	1	-	-	-	-	2
	0	0	-	-	-	-	0

(f) Cross — stimulus (2+1+1+2)

period, phase of response	1	2	3	4	5	6	sum
2,0	10	11	13	14	6	1	55
	14	19	16	9	2	0	60
1,0	12	8	-	-	-	-	20
	8	2	-	-	-	-	10
6,4	-	-	2	2	4	9	17
	-	-	1	4	6	9	20
4,0	-	2	4	2	5	2	15
	-	1	4	5	8	4	22
6,0	-	-	2	3	3	1	9
	-	-	1	4	6	9	20
6,2	-	-	1	1	3	3	8
	-	-	0	0	0	0	0
6,3	-	-	-	-	1	3	4
	-	-	-	-	0	0	0
2,1	-	1	-	-	-	1	2
	-	0	-	-	-	0	0
6,5	-	-	-	-	-	1	1
	-	-	-	-	-	0	0
7,0	-	-	-	-	-	1	1
	-	-	-	-	-	0	0

FIGURE A.2: Parncutt, 1994, Figure 2(c) - 2(f) on p. 417: "[...] The periods and phases shown on the far left are expressed in nominal 1/4-note beats. Phases are shown only if necessary for the unambiguous specification of pulse responses; otherwise, a dash (-) is marked." (figure caption)

Appendix B

Additional listings of meter classes

Tables B.1, B.2, and B.3 provide additional listings to illustrate some remarks in section 4.2.2.

TABLE B.1: Multiplicative meters up to 24 periods per metric cycle (cf. Härpfer, 2014, p. 1025)

periods	multiplicative meters	time signature example
4	2 x 2	2/4 (elementary pulse: 1/8)
6	3 x 2	6/8 (e.p. 1/8)
	2 x 3	3/4 (e.p. 1/8)
8	2 x 2 x 2	4/4 (e.p. 1/8)
9	3 x 3	9/8 (e.p. 1/8)
12	3 x 2 x 2	12/8 (e.p. 1/8)
	2 x 3 x 2	6/8 (e.p. 1/16)
	2 x 2 x 3	3/4 (e.p. 1/16)
16	2 x 2 x 2 x 2	4/4 (e.p. 1/16)
18	3 x 3 x 2	6/4 (e.p. 1/8 triplet)
	3 x 2 x 3	3/2 (e.p. 1/8 triplet)
	2 x 3 x 3	9/4 (e.p. 1/8)
24	3 x 2 x 2 x 2	4/4 (e.p. 1/16 triplet)
	2 x 3 x 2 x 2	12/8 (e.p. 1/16)
	2 x 2 x 3 x 2	6/4 (e.p. 1/16)
	2 x 2 x 2 x 3	3/2 (e.p. 1/16)

"The numeric notation [in table B.1] shows the relations of period durations of adjacent levels. For example the only possible meter with four pulses has three levels with two periods of level 0 (the pulse-level) combined in level 1 and two periods of level 1 combined in level 2. It has to be made clear in which order the multiplicators relate to the indices of the levels. Here the order is from lower/faster levels to higher/slower levels. Hence, in the first example of the two six-period-meters level 1 contains three periods of level 0 and level 2 combines two periods of level 1. The examples of time signatures just give a hint to relate a certain meter to a possible musical notation. Taking other musical context into account there could be different solutions. A meter with only four pulses in a full cycle could also be notated in $\frac{2}{8}$ time signature and the pulse duration would then be $\frac{1}{16}$."[1]

[1] Härpfer, 2014, pp. 1025 f.

TABLE B.2: Combinatory explosion of grouping possibilities between 13 and 24 periods per metric cycle (cf. Härpfer, 2014, p. 1028)

periods	possible layer structures (with intermediate levels)	examples	number of layers
13	32	2 0 0 1 0 2 0 1 0 0 1 0 0	2
14	34	2 0 0 1 0 2 0 1 0 2 0 1 0 0	2
	(6)	(2 0 1 0 1 0 0 2 0 1 0 0 1 0)	
15	57	2 0 0 1 0 2 0 1 0 2 0 0 1 0 0	2
	(6)	(2 0 1 0 0 2 0 0 1 0 2 0 1 0 0)	
16	87	2 0 1 0 1 0 0 2 0 0 1 0 0 1 0 0	2
	(6)	(2 0 1 0 0 1 0 0 2 0 0 1 0 1 0 0)	
17	149	3 0 1 0 2 0 1 0 3 0 1 0 2 0 0 1 0	2 or 3
18	201	2 0 1 0 1 0 2 0 1 0 0 2 0 1 0 0 1 0	2 or 3
	(18)	(3 0 1 0 2 0 0 1 0 3 0 1 0 2 0 1 0 0)	
19	332	2 0 1 0 1 0 2 0 0 1 0 0 1 0 2 0 1 0 0	2 or 3
20	457	3 0 1 0 2 0 1 0 0 3 0 1 0 0 1 0 2 0 1 0	2 or 3
	(54)	(3 0 1 0 2 0 0 1 0 0 3 0 1 0 1 0 2 0 1 0)	
21	709	3 0 1 0 2 0 1 0 0 3 0 1 0 0 2 0 1 0 1 0 0	2 or 3
	(24)	(2 0 1 0 1 0 0 2 0 0 1 0 1 0 2 0 1 0 0 1 0)	
22	1046	2 0 1 0 1 0 2 0 1 0 0 1 0 2 0 0 1 0 0 1 0 0	2 or 3
	(172)	(3 0 0 1 0 0 2 0 1 0 0 3 0 1 0 2 0 1 0 1 0 0)	
23	1776	3 0 1 0 1 0 0 2 0 0 1 0 3 0 0 1 0 0 2 0 0 1 0	2 or 3
24	2775	3 0 1 0 0 2 0 1 0 0 1 0 0 3 0 1 0 0 2 0 1 0 1 0	2 or 3
	(508)	(2 0 0 1 0 1 0 0 2 0 0 1 0 0 1 0 2 0 1 0 0 1 0 0)	

"To illustrate the combinatory explosion of possibilities with a growing number of periods, the numbers of alternative layer structures in the area between 13 and 24 periods"[2] are listed in table B.2. "All metric cycles with a non-prime number of periods in this range contain layer structures with intermediate levels. Their numbers and correspondent examples are shown in brackets. They form a subset of all possible layer structures the numbers of which are listed first. The rightmost column gives information about possible depths of metric stratification in these additive structures. There could be second-oder and third-order groupings according to the assumption"[3] stated in section 4.2.1.

Table B.3 enumerates "hybrid" metric cycles containing up to 24 periods. "The second column shows the multiplicative metric structures. As in [table B.1] the numeric notation specifies the relations of period durations of adjacent levels. The additive components can be specified with the possible layer structures of certain levels."[4]

[2]Härpfer, 2014, p. 1027
[3]ibid.
[4]ibid.

TABLE B.3: Hybrid meters up to 24 period per metric cycle (cf. Härpfer, 2014, p. 1028)

periods	hybrid meters	structural variants	number of hybrid meters
10	5 x 2	2	4
	2 x 5	2	
14	7 x 2	3	6
	2 x 7	3	
15	5 x 3	2	4
	3 x 5	2	
16	8 x 2	3	6
	2 x 8	3	
18	9 x 2	4	8
	2 x 9	4	
20	10 x 2	4	14
	2 x 10	4	
	5 x 2 x 2	2	
	2 x 5 x 2	2	
	2 x 2 x 5	2	
21	7 x 3	3	6
	3 x 7	3	
22	11 x 2	14	28
	2 x 11	14	
24	12 x 2	20	46
	8 x 3	3	
	3 x 8	3	
	2 x 12	20	

Appendix C

Implementation of the extended indispensability algorithm

The algorithm described in section 4.3.3 is provided in the following as a C++ function which takes two integer arrays of length `n` (for n elementary pulses, given by the first argument). The first array (`gnsm`) represents the GNSM code of the meter which has to be prepared before the function call. When terminated, the second array (`idp`) will contain the correspondent indispensability values for the `n` pulses.

```
Copyright (c) 2015 Härpfer. All rights reserved.

#include <vector>

void getIndispensability(int n, int* gnsm, int* idp) {

    vector<int>  *triplets  = new vector<int>;
    int          strata     = gnsm[0];
    int          topPulses  = 0;
    int          higherPulse;
    int          pulse;

    for (pulse = 0; pulse < n; pulse++) idp[pulse] = -1;

    // get sum of top layer pulses
    for (pulse = 0; pulse < n; pulse++) if (gnsm[pulse] == strata) topPulses++;

    pulse = 1;

    // set basic IDP for top layer
    if (topPulses == 2) {
        idp[0] = 1;
        while (true) {if (gnsm[pulse] == strata) {idp[pulse++] = 0; break;} pulse++;}
    }
    else {
        idp[0] = 2;
        while (true) {if (gnsm[pulse] == strata) {idp[pulse++] = 0; break;} pulse++;}
        while (true) {if (gnsm[pulse] == strata) {idp[pulse++] = 1; break;} pulse++;}
    }

    // get indispensability values iteratively
    for (int focus = strata - 1; focus >= 0; focus--) {

        int rise = 0;

        for (pulse = 0; pulse < n; pulse++)
            if (gnsm[pulse] - focus == 0) {

                higherPulse = pulse;
```

```
                while (gnsm[++higherPulse % n] < gnsm[pulse]) ;

                if (gnsm[higherPulse % n] > gnsm[pulse]) {
                    idp[pulse] = idp[higherPulse % n];
                    rise++;
                }
            }

        for (pulse = 0; pulse < n; pulse++)
            if (gnsm[pulse] - focus > 0) idp[pulse] += rise;

        rise = 0;
        int lowest_idp, current_idp;
        bool firstTriplet = true;

        for (pulse = 0; pulse < n; pulse++)
            if (gnsm[pulse] - focus == 0) {

                higherPulse = pulse;
                while (gnsm[++higherPulse % n] < gnsm[pulse]) ;

                // triplets
                if (gnsm[higherPulse % n] == gnsm[pulse]) {

                    if (firstTriplet) {
                        lowest_idp = current_idp = idp[pulse] = idp[higherPulse % n];
                        firstTriplet = false;
                    }
                    else
                        if (idp[higherPulse % n] < lowest_idp)
                            lowest_idp = idp[pulse] =
                            idp[higherPulse % n];
                        else
                            idp[pulse] = ++current_idp;

                    triplets->push_back(pulse);
                    rise++;
                }
            }

        if (triplets->size() > 0) {

            int triplet = 0;

            for (pulse = 0; pulse < n; pulse++) {
                if (pulse != (*triplets)[triplet] and idp[pulse] >= lowest_idp)
                    idp[pulse] += rise;
                if (pulse == (*triplets)[triplet]) if(triplet < triplets->size() - 1)
                    triplet++;
            }

            triplets->clear();
        }
    }

    delete triplets;
}
```

Appendix D

Comparative data related to metric coherence and metric contrast

Tables D.1 and D.2 provide a selective comparison of metric coherence and metric period contrast. Metric coherence is defined independently of pulse rates (section 4.4.1). Therefore, a representative elementary-pulse rate of $e = 0.25s$ is chosen for the parallel estimation of metric period contrast. Although they are differently scaled and in different ratios between each other, the values of the comparisons between *different* meters can be listed in exactly the reverse order. The different values of metric coherence regarding the comparison of a meter with itself, reflect another peculiarity of this measure. As indicated in section 4.4.1, metric coherence is not yet properly applicable to metric phase shifts: it would be necessary to readjust scaling factors of BARLOW's procedure. The application of metric contrast – including metric phase contrast – in the context of metric malleability is discussed in section 5.2.5.

TABLE D.1: Metric coherence

meter	3 2 2	2 3 2	2 2 3	2 2 2	3 3
3 2 2	0.41454	0.364211	0.207975	0.415752	0.164577
2 3 2		0.41454	0.269584	0.352401	0.217829
2 2 3			0.41454	0.184157	0.388446
2 2 2				0.463821	0.154779
3 3					0.446297

TABLE D.2: Metric period contrast at $e = 0.25s$

meter	3 2 2	2 3 2	2 2 3	2 2 2	3 3
3 2 2	0	0.0642101	1.22335	7.78204e-05	1.18766
2 3 2		0	0.570617	0.0711607	0.585237
2 2 3			0	1.27004	0.00225218
2 2 2				0	1.22897
3 3					0

Figure D.1 shows the corresponding functions of metric period contrast, as currently defined in dependence of pulse rate. As indicated in section 4.4.2, further studies are needed to qualify the amount of influence on metric contrast by rhythmic presentation rates.

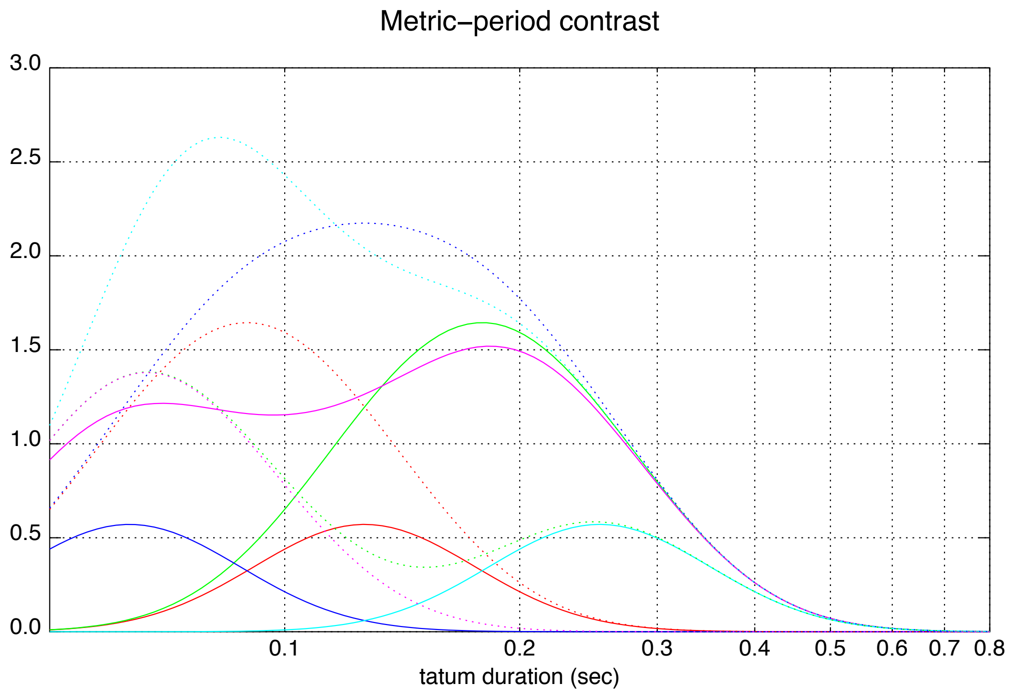


FIGURE D.1: solid red: 3 x 2 x 2 vs. 2 x 3 x 2
solid green: 3 x 2 x 2 vs. 2 x 2 x 3
solid blue: 3 x 2 x 2 vs. 2 x 2 x 2
solid magenta: 3 x 2 x 2 vs. 3 x 3
solid cyan: 2 x 3 x 2 vs. 2 x 2 x 3
dotted red: 2 x 3 x 2 vs. 2 x 2 x 2
dotted green: 2 x 3 x 2 vs. 3 x 3
dotted blue: 2 x 2 x 3 vs. 2 x 2 x 2
dotted magenta: 2 x 2 x 3 vs. 3 x 3
dotted cyan: 2 x 2 x 2 vs. 3 x 3

Appendix E

Instances of malleable necklaces of class $n = 12$

The following figures display additional results of the study presented in section 5.3. Supplementary information about parameter settings is provided in relation to the diagrams, using the notation introduced in section 5.3.

Figure E.1 shows the most contrasting meter constellations for the five necklaces of the 12/5 class which are favored according to one specific parameter setting.

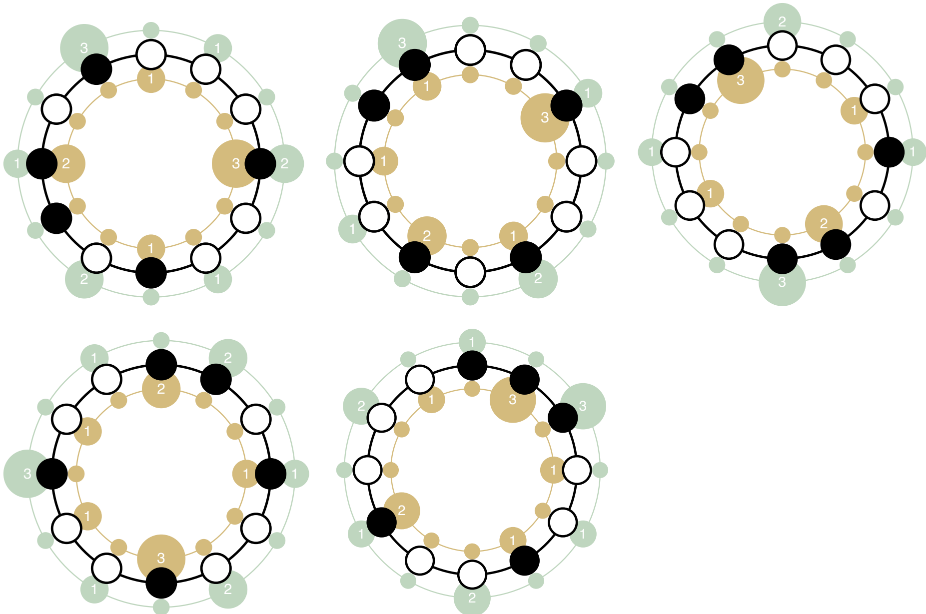


FIGURE E.1: Fife rhythmic necklaces of class 12/5 and their most contrasting salient frameworks (for explanations see text).

The three upper patterns are top-ranked at elementary-pulse period $0.15s$. The pattern [3-2-1-2-4] affords maximal period contrast with $m = 3$, necklace [3-2-3-1-3] allows for maximal phase contrast with $m = 3$, and the pattern [2-1-4-1-4] features maximal phase contrast with $m = 24$. At elementary-pulse period $0.22s$, the pattern [1-2-3-3-3] allows for maximal phase contrast with $m = 24$ (lower-left diagram), and at elementary-pulse period $0.3s$, the pattern [1-1-3-3-4] affords maximal phase contrast, again with $m = 24$ (lower-middle diagram).

Figure E.2 lists the most contrasting meter constellations afforded by the six necklaces of the $12/7$ class which are top-ranked once in respect to a specific parameter configuration.

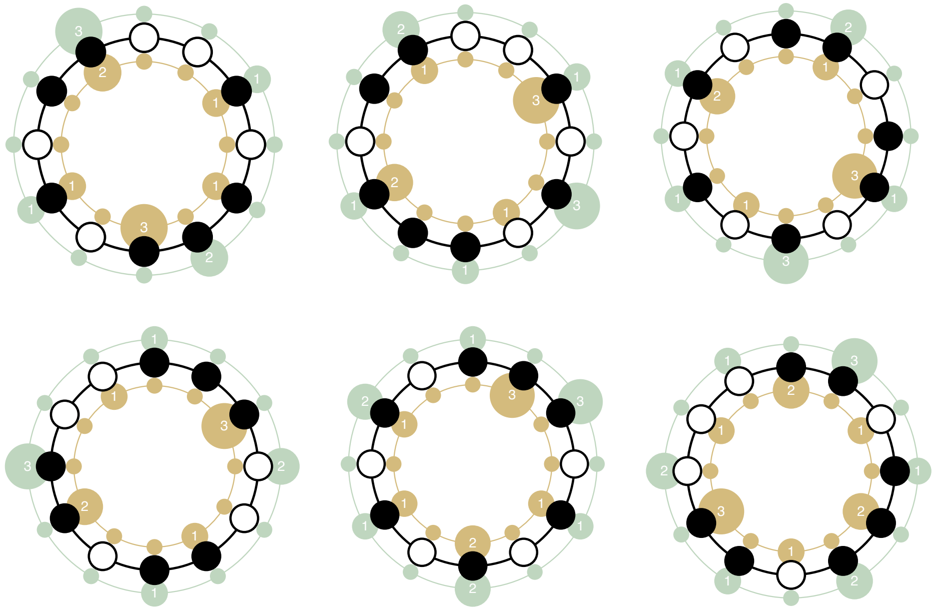


FIGURE E.2: Six rhythmic necklaces of class $12/7$ and their most contrasting salient frameworks (for explanations see text).

Ordered top down and from left to right, the first four instances are related to $e = 0.15s$. Pattern [2-1-1-2-2-1-3] affords maximal period contrast with $m = 3$, necklace [2-2-1-1-2-1-3] features maximal phase contrast with $m = 3$, necklace [1-2-1-2-2-2-2] allows for maximal period contrast with $m = 24$, and pattern [1-1-3-1-2-1-3] represents maximal phase contrast with $m = 24$. Thus, by trend these four patterns are favored when meter salience is a higher-weighted factor than metric contrast. For the remaining instances, top-ranked at elementary-pulse period $0.22s$, these factors are rather balanced ($m = 24$). Under these conditions, pattern [1-1-2-2-2-2-2] allows for maximal period contrast, and pattern [1-2-1-1-2-1-4] for maximal phase contrast.

Bibliography

Andreatta, Moreno (2004). “On group-theoretical methods applied to music: some compositional and implementational aspects”. In: *Perspectives in Mathematical and Computational Music Theory*. Ed. by Guerino Mazzola, Thomas Noll, and Emilio Lluis-Puebla. Vol. 1. Osnabrück Series on Music and Computation. epOs-Music.

– (2011). “Constructing and Formalizing Tiling Rhythmic Canons: a Historical Survey of a "Mathemusical” Problem”. In: *Perspectives of New Music* 49.2, pp. 33–64.

Andreatta, Moreno, Carlos Agon, and Emmanuel Amiot (2002). “Tiling problems in music composition: Theory and Implementation”. In: *Proceedings of the ICMC*. Göteborg, Sweden, pp. 156–163.

Andreatta, Moreno et al. (2006). “Towards Pedagogability of Mathematical Music Theory: Algebraic Models and Tiling Problems in Computer-aided Composition”. In: *Bridges London: Mathematics, Music, Art, Architecture, Culture*. Ed. by Reza Sarhangi and John Sharp. Available online at http://archive.bridgesmathart.org/2006/bridges2006-277.html. London: Tarquin Publications, pp. 277–284. ISBN: 0-9665201-7-3.

Anku, Willie (2000). “Circles and Time: A Theory of Structural Organization of Rhythm in African Music”. In: *Music Theory Online* 6.1.

Baraldi, Filippo Bonini, Emmanuel Bigand, and Thierry Pozzo (2015). “Measuring Aksak Rhythm and Synchronization in Transylvanian Village Music by Using Motion Capture”. In: *Empirical Musicology Review* 10.4, pp. 265–291.

Barlow, Clarence (June 2012). *On Musiquantics*. Ed. by University Mainz Dept. Music Informatics Musicological Institute. Vol. Report. 51. Mainz: University Mainz.

Barnes, Ralph and Mari Riess Jones (2000). “Expectancy, Attention, and Time”. In: *Cognitive Psychology* 41, pp. 254–311.

Bello, Juan P. et al. (2015). “Five Perspectives on Musical Rhythm”. In: *Journal of New Music Research* 44.1, pp. 1–2. DOI: 10.1080/09298215.2014.996572. eprint: http://dx.doi.org/10.1080/09298215.2014.996572. URL: http://dx.doi.org/10.1080/09298215.2014.996572.

Bello, Juan Pablo et al. (2005). “A Tutorial on Onset Detection in Music Signals”. In: *IEEE Transactions on Speech and Audio Processing*. Vol. 13. 5, pp. 1035–1047.

Benadon, Fernando (2004). “Towards a Theory of Tempo Modulation”. In: *Proceedings of the 8th International Conference on Music Perception & Cognition*. Ed. by S.D. Lipscomb et al., pp. 563–566.

Benadon, Fernando (2007). "A Circular Plot for Rhythm Visualization and Analysis". In: *Music Theory Online* 13.3. URL: http://www.mtosmt.org/issues/mto.07.13.3/mto.07.13.3.benadon.html.

Bengtsson, Ingmar (1987). "Notation, Motion and Perception: Some Aspects of Musical Rhythm". In: *Action and Perception in Rhythm and Music*. 55. Stockholm: Royal Swedish Academy of Music, pp. 69–79.

Benjamin, William E. (1984). "A Theory of Musical Meter". In: *Music Perception* 1.4, pp. 355–413.

Benson, Dave J. (2007). *Music: A Mathematical Offering*. Cambridge, UK: Cambridge University Press.

Bernardes, Gilberto, Carlos Guedes, and Bruce Pennycook (2010). "Style emulation of drum patterns by means of evolutionary methods and statistical analysis". In: *Proceedings of the Sound and Music Conference*. Barcelona, Spain.

Berry, Wallace (1976). *Structural functions in music*. Englewood Cliffs, NJ: Prentice-Hall, XII, 447 S. ISBN: 978-0-13-853903-0.

Böck, Sebastian, Florian Krebs, and Gerhard Widmer (2016). "Joint Beat and Downbeat Tracking with Recurrent Neural Networks". In: *Proceedings of the 17th International Society for Music Information Retrieval Conference, ISMIR 2016, New York City, United States, August 7-11, 2016*, pp. 255–261.

Bolton, Thaddeus L. (Jan. 1894). *Rhythm*. University of Illinois Press, originally published in The American Journal of Psychology, Vol. 6, No. 2, pp. 145-238, stable URL: http://www.jstor.org/stable/1410948.

Brochard, Renaud et al. (2003). "The "Ticktock" of Our Internal Clock: Direct Brain Evidence of Subjective Accents in Isochronous Sequences". In: *Psychological Science* 14.4, pp. 362–366.

Buchler, Michael (2006). "The sonic illusion of metrical consistency in recent minimalist composition". In: *Proceedings of the 9th International Conference on Music Perception & Cognition (ICMPC9)*. Ed. by M. Baroni et al. European Society for the Cognitive Sciences of Music (ESCOM). Society for Music Perception & Cognition (SMPC), pp. 697–701.

Burns, James (2010). "Rhythmic Archetypes in Instrumental Music from Africa and the Diaspora". In: *Music Theory Online* 16.4. URL: http://www.mtosmt.org/issues/mto.10.16.4/mto.10.16.4.burns.html.

Chemillier, Marc (2004). "Periodic musical sequences and Lyndon words". In: *Soft Computing* 8.9, pp. 611–616. DOI: 10.1007/s00500-004-0387-2. URL: http://dx.doi.org/10.1007/s00500-004-0387-2.

Chemillier, Marc and Charlotte Truchet (2003). "Computation of words satisfying the "rhythmic oddity property" (after Simha Arom's works)". In: *Information Processing Letters* 86.5, pp. 255 –261. ISSN: 0020-0190. DOI: http://dx.doi.org/10.1016/S0020-0190(02)00521-5. URL: http://www.sciencedirect.com/science/article/pii/S0020019002005215.

Chen, Joyce L., Robert J. Zatorre, and Virginia B. Penhune (2006). "Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms". In: *NeuroImage* 32.4, pp. 1771 –1781. ISSN: 1053-8119. DOI: http://dx.doi.org/10.1016/j.neuroimage.2006.04.207. URL: http://www.sciencedirect.com/science/article/pii/S1053811906005064.

Clampitt, David and Thomas Noll (2011). "Modes, the Height-Width Duality, and Handschin's Tone Character". In: *Music Theory Online* 17.1.

Clarke, Eric (1987). "Categorical rhythm perception: An ecological perspective". In: *Action and Perception in Rhythm and Music*. Ed. by Alf Gabrielsson. 55. Stockholm: Royal Swedish Academy of Music.

Clayton, Martin (1997). "Le mètre et le tāl dans la musique de l'Inde du Nord". In: *Cahiers d'ethnomusicologie* 10. Translated by Georges Goormaghtigh, online since 06 January 2012, connection on 04 March 2016. URL: http://ethnomusicologie.revues.org/888.

Cohn, Richard (2001). "Complex Hemiolas, Ski-Hill Graphs, and Metric Spaces". In: *Music Analysis* 20.3, pp. 295–326.

Colannino, Justin, Francisco Gómez, and Godfried T. Toussaint (2009). "Analysis of Emergent Beat-Class Sets in Steve Reich's *Clapping Music* and the *Yoruba* Bell Timeline". In: *Perspectives of New Music* 47.1, pp. 111–134.

Cuthbert, Michael Scott (2006). "Generalized Set Analysis of Sub-Saharan African Rhythm? Evaluating and Expanding the Theories of Willie Anku". In: *Journal of New Music Research* 35.3, pp. 211–219.

Dadelsen, H.C. von (1994). "Rhythmische Wechselwirkungskräfte – musikalische Indizien der Chaostheorie?" In: *Kulturelle Dialoge: Arbeitsprozesse in Physik und Musik*. Ed. by Akademie der Künste Berlin. Frankfurt a.M., pp. 33–46.

Demaine, Erik D. et al. (2009). "The distance geometry of music". In: *Computational Geometry* 42.5. Special Issue on the Canadian Conference on Computational Geometry (CCCG 2005 and CCCG 2006), pp. 429 –454. ISSN: 0925-7721. DOI: https://doi.org/10.1016/j.comgeo.2008.04.005. URL: http://www.sciencedirect.com/science/article/pii/S0925772108001156.

Desain, Peter (1992). "A (De)Composable Theory of Rhythm Perception". In: *Music Perception* 9.4, pp. 439–454.

Desain, Peter and Henkjan Honing (1992). "Tempo curves considered harmful". In: *Music, Mind, and Machine: Studies in Computer Music, Music Cognition, and Artificial Intelligence*. Ed. by Peter Desain and Henkjan Honing. Amsterdam: Thesis Publishers.

– (1999). "Computational Models of Beat Induction: The Rule-Based Approach". In: *Journal of New Music Research* 28.1, pp. 29–42.

– (2003). "The formation of rhythmic categories and metric priming". In: *Perception* 32, pp. 341–365.

Deutsch, Diana (1982). “Grouping mechanisms in music”. In: *The psychology of music*. Ed. by Diana Deutsch. Vol. 1. New York, London: Academic Press, INC. Chap. 4, pp. 99–134.

Duţică, Gheorghe (2010). “The Modal Palindrome: A Structural Matrix and a Generative Mechanism”. In: *Proceedings of the 11th WSEAS International Conference on Acoustics & Music: Theory & Applications*. AMTA’10. Iasi, Romania: World Scientific, Engineering Academy, and Society (WSEAS), pp. 98–102. ISBN: 978-960-474-192-2. URL: http://dl.acm.org/citation.cfm?id=1863211.1863230.

Eck, Douglas (2001). “A Positive-Evidence Model for Rhythmical Beat Induction”. In: *Journal of New Music Research* 30.2, pp. 187–200.

Essens, Peter J. and Dirk-Jan Povel (1985). “Metrical and nonmetrical representations of temporal patterns”. In: *Perception & Psychophysics* 37.1, pp. 1–7.

Feyerabend, Paul (1984). *Wissenschaft als Kunst*. Frankfurt a.M.: Suhrkamp.

Flanagan, Patrick (2008). “Quantifying Metrical Ambiguity”. In: *Proceedings of the 9th International Conference on Music Information Retrieval*. http://ismir2008.ismir.net/papers/ISMIR2008_153.pdf. Philadelphia, USA, pp. 635–640.

Fleischer, Anja (2002). “A Model of Metric Coherence”. In: *In Procedings of the 2nd Conference, Understanding and Creating Music*.

Fraisse, Paul (1982). “Rhythm and Tempo”. In: *The psychology of music*. Ed. by Diana Deutsch. Vol. 1. New York, London: Academic Press, INC. Chap. 6, pp. 149–180.

– (1987). “A historical approach to rhythm as perception”. In: *Action and perception in rhythm and music*. Ed. by A. Gabrielsson. 55. Royal Swedish Academy of Music, pp. 7–18.

Fripertinger, Harald (1991). “Enumeration in Musical Theory”. In: *Séminaire Lotharingien de Combinatoire*, pp. 29–42.

Gabrielsson, Alf (1982). “Perception and Performance of Musical Rhythm”. In: *Music, Mind, and Brain. The Neuropsychology of Music*. Ed. by Manfred Clynes. New York, London: Plenum Press. Chap. 9, pp. 159–169.

Gomez-Martin, Francisco, Perouz Taslakian, and Godfried T. Toussaint (2008). “Evenness Preserving Operations on Musical Rhythms”. In: *Proceedings of the 2008 C3S2E Conference*. C3S2E ’08. Montreal, Quebec, Canada: ACM, pp. 121–123. ISBN: 978-1-60558-101-9. DOI: 10.1145/1370256.1370275. URL: http://doi.acm.org/10.1145/1370256.1370275.

Gómez-Martín, Francisco, Perouz Taslakian, and Godfried T. Toussaint (2009). “Interlocking and Euclidean rhythms”. In: *Journal of Mathematics and Music* 3.1, pp. 15–30. DOI: 10.1080/17459730902916545.

Gotham, Mark (2015a). “Attractor tempos for metrical structures”. In: *Journal of Mathematics and Music* 9.1, pp. 23–44. DOI: 10.1080/17459737.2014.980343. eprint: http://dx.doi.org/10.1080/17459737.2014.980343. URL: http://dx.doi.org/10.1080/17459737.2014.980343.

– (2015b). "Meter Metrics: Characterizing Relationships Among (Mixed) Metrical Structures". In: *Music Theory Online* 21.2. URL: http://www.mtosmt.org/issues/mto.15.21.2/mto.15.21.2.gotham.html.

Grahn, Jessica A. (2012). "Neural Mechanisms of Rhythm Perception: Current Findings and Future Perspectives". In: *Topics in Cognitive Science* 4, pp. 585–606.

Hall, Rachel W. and Paul Klingsberg (2004). "Asymmetric Rhythms, Tiling Canons, and Burnside's Lemma". In: *Bridges: Mathematical Connections in Art, Music, and Science*. Ed. by Reza Sarhangi and Carlo Séquin. Available online at http://archive.bridgesmathart.org/2004/bridges2004-189.html. Southwestern College, Winfield, Kansas: Bridges Conference, pp. 189–194. ISBN: 0-9665201-5-7.

Handel, Stephen (1984). "Using Polyrhythms to Study Rhythm". In: *Music Perception* 1.4, pp. 465–484.

Handel, Stephen and Gregory R. Lawson (1983). "The contextual nature of rhythmic interpretation". In: *Perception & Psychophysics* 34.2, pp. 103–120.

Handel, Stephen and James S. Oshinsky (1981). "The meter of syncopated auditory polyrhythms". In: *Perception & Psychophysics* 30.1, pp. 1–9.

Hannon, Erin E. and Sandra E. Trehub (2005). "Metrical Categories in Infancy and Adulthood". In: *Psychological Science* 16.1, pp. 48–55. URL: http://www.jstor.org/stable/40064071.

Härpfer, Bernd (2014). "Computing Musical Meter – an Approach to an Integrated Formal Description". In: *Proceedings of the 40th International Computer Music Conference*. Ed. by Anastasia Georgaki and Georgios Kouroupetroglu, pp. 1024–1028.

Hasty, Christopher F. (1997). *Meter as rhythm*. New York: Oxford University Press, Inc.

Holzapfel, Andre and Thomas Grill (2016). "Bayesian Meter Tracking on Learned Signal Representations". In: *Proceedings of the 17th International Society for Music Information Retrieval Conference, ISMIR 2016, New York City, United States, August 7-11, 2016*, pp. 262–268.

Honing, Henkjan (2006). "On the Growing Role of Observation, Formalization and Experimental Method in Musicology". In: *Empirical Musicology Review* 1.1, pp. 2–6.

Honing, Henkjan, Fleur L. Bouwer, and Gábor P. Háden (2014). "Perceiving Temporal Regularity in Music: The Role of Auditory Event-Related Potentials (ERPs) in Probing Beat Perception". In: *Neurobiology of Interval Timing*. Ed. by Hugo Merchant and Victor de Lafuente. Vol. 829. Advances in Experimental Medicine and Biology. New York: Springer, pp. 305–323.

Hook, Julian (2007). "Why Are There Twenty-Nine Tetrachords? A Tutorial on Combinatorics and Enumeration in Music Theory". In: *Music Theory Online* 13.4. URL: http://www.mtosmt.org/issues/mto.07.13.4/mto.07.13.4.hook.html.

Huron, David (2006). *Sweet Anticipation. Music and the Psychology of Expectation*. Cambridge, Massachusetts: The MIT Press.

Jedrzejewski, Franck (2006). *Mathematical Theory Of Music*. Editions Delatour France/Ircam-Centre Pompidou.

– (2014). *Enumeration of Vuza Canons*.

Jedrzejewski, Franck and Tom Johnson (2013). "Mathematics and Computation in Music: 4th International Conference, MCM 2013, Montreal, QC, Canada, June 12-14, 2013. Proceedings". In: ed. by Jason Yust, Jonathan Wild, and John Ashley Burgoyne. Berlin, Heidelberg: Springer Berlin Heidelberg. Chap. The Structure of Z-Related Sets, pp. 128–137. ISBN: 978-3-642-39357-0. DOI: 10.1007/978-3-642-39357-0_10. URL: http://dx.doi.org/10.1007/978-3-642-39357-0_10.

Jones, Mari Riess (1987a). "Dynamic pattern structure in music: Recent theory and research". In: *Perception & Psychophysics* 41.6, pp. 621–634.

– (1987b). "Perspectives on musical time". In: *Action and perception in rhythm and music*. Ed. by A. Gabrielsson. Stockholm: Royal Swedish Academy of Music, pp. 153–175.

– (2010). "Music Perception: Current Research and Future Directions". In: *Music Perception*. Ed. by Mari Riess Jones, Arthur N. Popper, and Richard R. Fay. New York: Springer. Chap. 1, pp. 1–12.

Jones, Mari Riess and Marilyn Boltz (1989). "Dynamic Attending and Responses to Time". In: *Psychological Review* 96.3, pp. 459–491.

Karim, S. et al. (2013). "Generating bracelets with fixed content". In: *Theoretical Computer Science* 475.4, pp. 103 –112. ISSN: 0304-3975. DOI: http://dx.doi.org/10.1016/j.tcs.2012.11.024. URL: http://www.sciencedirect.com/science/article/pii/S0304397512010547.

Kirilov, Kalin (2012). *Rhythmic and Metric Aspects of Contemporary Bulgarian Wedding Music*. Paper at the Second International Conference on Analytical Approaches to World Music (AAWM), http://www.aawmconference.com/aawm2012/papers.htm.

Krebs, Florian et al. (2016). "Downbeat Tracking Using Beat Synchronous Features with Recurrent Neural Networks". In: *Proceedings of the 17th International Society for Music Information Retrieval Conference, ISMIR 2016, New York City, United States, August 7-11, 2016*, pp. 129–135.

Krebs, Harald (1987). "Some extensions of the concepts of metrical consonance and dissonance". In: *Journal of Music Theory* 31.

Large, Edward W., Philip Fink, and J. A. Scott Kelso (2002). "Tracking simple and complex sequences". In: *Psychological Research* 66, pp. 3–17. DOI: 10.1007/s004260100069.

Lee, Christopher S. (1991). "The Perception of Metrical Structure: Experimental Evidence and a Model". In: *Representing Musical Structure*. Ed. by Peter Howell, Robert West, and Ian Cross. London: Academic Press. Chap. 3, pp. 59–127.

Leman, Marc and Pieter-Jan Maes (2014). "The Role of Embodiment in the Perception of Music". In: *Empirical Musicology Review* 9.3-4, pp. 236–246.

Lerdahl, Fred and Ray Jackendoff (1981). "On the theory of grouping and meter". In: *The Musical Quarterly* 67.4, pp. 479–506.

– (1983). *A Generative Theory of Tonal Music*. Cambridge, Massachusetts: The MIT Press.

Lewin, David (1980). “On Generalized Intervals and Transformations”. In: *Journal of Music Theory* 24.2, pp. 243–251.

Locke, David (2011). “The Metric Matrix: Simultaneous Multidimensionality in African Music”. In: *Analytical Approaches To World Music* 1.1, pp. 48–72.

London, Justin (2008). *Cognitive and Aesthetic Aspects of Metrical Ambiguity*. Colloquium talk given at The University of Alberta and The University of Pennsylvania. URL: http://www.people.carleton.edu/~jlondon/MetricAmbiguity.pdf.

– (2012). *Hearing in Time : Psychological Aspects of Musical Meter*. Second Edition. Oxford University Press.

Longuet-Higgins, H.C. and Christopher S. Lee (1984). “The Rhythmic Interpretation of Monophonic Music”. In: *Music Perception* 1.4, pp. 424–441.

Madison, Guy and Björn Merker (2002). “On the limits of anisochrony in pulse attribution”. In: *Psychological Research* 66, pp. 201–207. DOI: 10.1007/s00426-001-0085-y.

Magill, Jonathan M. and Jeffrey L. Pressing (1997). “Asymmetric Cognitive Clock Structures in West African Rhythms”. In: *Music Perception: An Interdisciplinary Journal* 15.2, pp. 189–221. ISSN: 0730-7829. DOI: 10.2307/40285749. eprint: http://mp.ucpress.edu/content/15/2/189.full.pdf. URL: http://mp.ucpress.edu/content/15/2/189.

McAuley, J. Devin (2010). “Tempo and Rhythm”. In: *Music Perception*. Ed. by Mari Riess Jones, Arthur N. Popper, and Richard R. Fay. New York: Springer. Chap. 6, pp. 165–200.

McLachlan, Neil (2000). “A Spatial Theory of Rhythmic Resolution”. In: *Leonardo Music Journal* 10, pp. 61–67.

Moelants, Dirk (1997). “A Framework for the Subsymbolic Description of Meter”. In: *Music, Gestalt, and Computing : Studies in Cognitive and Systematic Musicology*. Ed. by Marc Leman. Springer, pp. 263–276.

Moelants, Dirk and Martin F. McKinney (2004). “Tempo Perception and Musical Content: What Makes a Piece Fast, Slow or Temporally Ambiguous?” In: *Proceedings of the 8th International Conference on Music Perception & Cognition*. Ed. by S.D. Lipscomb et al., pp. 558–562.

Noorden, Leon van and Dirk Moelants (1999). “Resonance in the Perception of Musical Pulse”. In: *Journal of New Music Research* 28.1, pp. 43–66.

Osborn, Brad (2010). “Beats That Commute: Algebraic And Kinesthetic Models for Math-Rock Grooves”. In: *Gamut* 3.1, pp. 43–68.

Palmer, Caroline and Carol L. Krumhansl (1990). “Mental Representations for Musical Meter”. In: *Journal of Experimental Psychology: Human Perception and Performance* 16.4, pp. 728–741.

Papadopoulos, Athanase (2014). *Mathematics and group theory in music*. Ed. by L. Ji, Athanase Papadopoulos, and S.-T. Yau.

Parncutt, Richard (1987). "The perception of pulse in musical rhythm". In: *Action and perception in rhythm and music*. Ed. by A. Gabrielsson. 55. Stockholm: Royal Swedish Academy of Music, pp. 127–138.

– (1994). "A Perceptual Model of Pulse Salience and Metrical Accent in Musical Rhythms". In: *Music Perception* 11.4, pp. 409–464.

Patel, Aniruddh D. (2008). *Music, Language, and the brain*. Oxford: Oxford University Press.

Petersen, Peter (2010). *Musik und Rhythmus. Grundlagen, Geschichte, Analyse*. Mainz: Schott Verlag.

Petitjean, Michel (2010). "Chirality in Metric Spaces". In: *Symmetry: Culture and Science* 21.1-3, pp. 27–36.

Polak, Rainer (2010). "Rhythmic Feel as Meter: Non-Isochronous Beat Subdivision in Jembe Music from Mali". In: *Music Theory Online* 16.4. URL: http://www.mtosmt.org/issues/mto.10.16.4/mto.10.16.4.polak.html.

Polak, Rainer and Justin London (2014). "Timing and Meter in Mande Drumming from Mali". In: *Music Theory Online* 20.1. URL: http://www.mtosmt.org/issues/mto.14.20.1/mto.14.20.1.polak-london.html.

Polansky, Larry (1996). "Morphological Metrics". In: *Journal of New Music Research* 25, pp. 289–368.

Povel, Dirk-Jan (1984). "A theoretical framework for rhythm perception". In: *Psychological Research* 45, pp. 315–337.

Povel, Dirk-Jan and Peter Essens (1985). "Perception of Temporal Patterns". In: *Music Perception* 2.4, pp. 411–440.

Povel, Dirk-Jan and Hans Okkerman (1981). "Accents in equitone sequences". In: *Perception & Psychophysics* 30.6, pp. 565–572.

Pressing, Jeffrey (1997). "Cognitive complexity and the structure of musical patterns". In: *Proceedings of the Fourth Conference of the Australasian Cognitive Science Society*. Newcastle, Australia.

Repp, Bruno H., Justin London, and Peter E. Keller (2011). "Perception–production relationships and phase correction in synchronization with two-interval rhythms". In: *Psychological Research* 75, pp. 227–242. DOI: 10.1007/s00426-010-0301-8.

Roads, Curtis (2001). *Microsound*. The MIT Press.

Rosenthal, David (1989). "A Model of the Process of Listening to Simple Rhythms". In: *Music Perception* 6.3, pp. 315–328.

– (1992). "Emulation of Human Rhythm Perception". In: *Computer Music Journal* 16.1.

Sadakata, Makiko, Peter Desain, and Henkjan Honing (2006). "The Bayesian Way to Relate Rhythm Perception and Production". In: *Music Perception* 23.3, pp. 269–288.

Schulze, Hans-Henning (1989). "Categorical perception of rhythmic patterns". In: *Psychological Research* 51, pp. 10–15.

Sethares, William A. (2007). *Rhythm and transforms*. London: Springer.

Shannon, Claude E. (1948). "A Mathematical Theory of Communication". In: *The Bell System Technical Journal* 27.3, pp. 379–423, 623–656.

Shmulevich, Ilya and Dirk-Jan Povel (2000). "Complexity measures of musical rhythms". In: *Rhythm perception and production*. Ed. by Peter Desain and L. Windsor. Lisse, Netherlands: Swets & Zeitlinger, pp. 239–244.

Sioros, George and Carlos Guedes (2011a). "A formal approach for high-level automatic rhythm generation". In: *Proceedings of the BRIDGES 2011 – Mathematics, Music, Art, Architecture, Culture Conference*. Coimbra, Portugal.

– (2011b). "Automatic Rhythmic Performance in Max/MSP: the kin.rhythmicator". In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. Oslo, Norway.

– (2014). "Sound, Music, and Motion: 10th International Symposium, CMMR 2013, Marseille, France, October 15-18, 2013. Revised Selected Papers". In: ed. by Mitsuko Aramaki et al. Cham: Springer International Publishing. Chap. Syncopation as Transformation, pp. 635–658. ISBN: 978-3-319-12976-1. DOI: 10.1007/978-3-319-12976-1_39. URL: http://dx.doi.org/10.1007/978-3-319-12976-1_39.

Sioros, George, André Holzapfel, and Carlos Guedes (2012). "On Measuring Syncopation to Drive an Interactive Music System". In: *Proceedings of the 13th International Society for Music Information Retrieval Conference*. http://ismir2012.ismir.net/event/papers/283-ismir-2012.pdf. Porto, Portugal.

Tanghe, Koen et al. (2005). "Collecting ground truth annotations for drum detection in polyphonic music". In: *Proceedings of the 6th International Conference on Music Information Retrieval*, pp. 50–57. URL: http://dx.doi.org/1854/4581.

Taslakian, Perouz (2008). "Musical Rhythms in the Euclidean Plane". PhD thesis. Montreal,Quebec: School of Computer Science, McGill University.

Temperley, David (1999). "Syncopation in rock: a perceptual perspective". In: *Popular Music* 18.1, pp. 19–40.

– (2000). "Meter and Grouping in African Music: A View from Music Theory". In: *Ethnomusicology* 44.1, pp. 65–96.

– (2004). "An Evaluation System for Metrical Models". In: *Computer Music Journal* 28.3, pp. 28–44.

– (2007). *Music and probability*. Cambridge, Massachusetts: MIT Press.

– (2008). "Hypermetrical Transitions". In: *Music Theory Spectrum* 30, pp. 305–325.

– (2010). "Modeling Common-Practice Rhythm". In: *Music Perception* 27, pp. 355–376.

Temperley, David and Christopher Bartlette (2002). "Parallelism as a Factor in Metrical Analysis". In: *Music Perception* 20.2, pp. 117–149.

Tenney, James and Larry Polansky (1980). "Temporal Gestalt Perception in Music". In: *Journal of Music Theory* 24.2, pp. 205–241.

Thul, Eric and Godfried Toussaint (2008). "Rhythm Complexity Measures: A Comparison of Mathematical Models of Human Perception and Performance". In: *Proceedings of the 9th International Conference on Music Information Retrieval*. http://

ismir2008.ismir.net/papers/ISMIR2008_125.pdf. Philadelphia, USA, pp. 663–668.

Tomkins, Calvin (2013). *Marcel Duchamp: The Afternoon Interviews*. Ed. by Paul Chan. First edition. Brooklyn, NY: Badlands Unlimited.

Toussaint, Godfried T. (2005). "The *Euclidean* Algorithm Generates Traditional Musical Rhythms". In: *Proceedings of BRIDGES: Mathematical Connections in Art, Music and Science*, pp. 47–55.

– (2010). "Generating "Good" Musical Rhythms Algorithmically". In: *Proceedings of the 8th International Conference on Arts and Humanities*. Honolulu, Hawaii, pp. 774–791.

– (2013). *The geometry of musical rhythm*. Boca Raton: CRC Press.

– (2015). "Quantifying Musical Meter: How Similar are African and Western Rhythm?" In: *Analytical Approaches To World Music* 4.1.

Vazan, Peter and Michael F. Schober (2004). "Detecting and resolving metrical ambiguity in a rock song upon multiple rehearing". In: *Proceedings of the 8th International Conference on Music Perception & Cognition (ICMPC8)*. Ed. by S.D. Lipscomb et al. SMPC.

Volk, Anja (2004). "Exploring the Interaction of Pulse Layers Regarding their Influence on Metrical Accents". In: *Proceedings of the 8th International Conference on Music Perception & Cognition (ICMPC8)*. Ed. by S.D. Lipscomb et al. SMPC, pp. 435–438.

– (2008). "The Study of Syncopation using Inner Metric Analysis: Linking Theoretical and Experimental Analysis of Metre in Music". In: *Journal of New Music Research* 37.4. http://dx.doi.org/10.1080/09298210802680758, pp. 259–273.

Vos, Joos and Rudolf Rasch (1981). "The Perceptual Onset of Musical Tones". In: *Perception & Psychophysics* 29.4, pp. 323–335.

Vuust, Peter and Maria Witek (2014). "Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music". In: *Frontiers in Psychology* 5.1111. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2014.01111. URL: http://www.frontiersin.org/auditory_cognitive_neuroscience/10.3389/fpsyg.2014.01111/abstract.

Witek, Maria A. G. et al. (Apr. 2014). "Syncopation, Body-Movement and Pleasure in Groove Music". In: *PLoS ONE* 9.4, pp. 1–12. DOI: 10.1371/journal.pone.0094446. URL: http://dx.doi.org/10.1371%2Fjournal.pone.0094446.

Wright, Matthew James (2008). "The Shape of an Instant: Measuring and Modeling Perceptual Attack Time with Probability Density Functions. (If a Tree Falls in the Forest, When Did 57 People Hear it Make a Sound?)" Dissertation. Stanford University.

Xenakis, Iannis (1992). *Formalized Music. Thought and Mathematics in Music*. Revised Edition. Harmonologia 6. Stuyvesant, NY: Pendragon Press.

Yeston, Maury (1976). *The stratification of musical rhythm*. New Haven, CT: Yale University Press.

Yu, Minhong, Laura Getz, and Michael Kubovy (2015). "Perceiving the initial note: Quantitative models of how listeners parse cyclical auditory patterns". In: *Attention,* *Perception, & Psychophysics* 77.8, pp. 2728–2739. ISSN: 1943-393X. DOI: 10.3758/s13414-015-0935-0. URL: http://dx.doi.org/10.3758/s13414-015-0935-0.