Formalized Groove

Xavier De Wannemaeker

Bachelor's thesis

Institute of Sonology, 2013

I want to thank my teachers at the Institute of Sonology for their advise and support, and Kees in particular, for the energy he puts in running our department. I thank my classmates for the feeling of togetherness we have shared through the course of our study, as well as all the Conservatory students who have inspired me during those years. I want to give my regards to my high-school friends: I hope our friendship will last for a lifetime. I want to express my gratitude to Marinus, Audrey, Marko, Elisabeth, Matus, Younes and Michel for being such good friends, through good and bad times. Younes deserves a special mention both for his input to this project and for getting me started with SuperCollider. Thank you to Julia for her smile, her energy, her determination, her cooking, her love and for copywriting this thesis. Finally I want to thank my parents and my grandmother, for their patience and unconditional support since always.This wouldn't have been possible without all of you !

Contents

Introduction

The title of this thesis refers to what to me seem to be the two major revolutions Western music has undergone during the course of the twentieth century : On one hand, the endless field of structural and timbral possibilities the use of electronic tools has opened, and on the other hand, the sense of groove common to all musics that have descended from the African diaspora.

I borrow the term 'formalized' from Xenakis, “the single most defiantly original and unconventional composer in history” (McHard, 2001, p. 215), whose continuous importation of new techniques from sciences, the use of stochastic functions in particular, is iconic of the revolution the use of electronic means has represented for music from the post-war avant-garde to this day. In this sense, 'formalized' refers to all progress both institutional and non-institutional electronic musics have achieved through their most innovative creations. 'Formalized' stands for experimental electronic music in general.

In all ages and places, popular music has mostly been aimed at making people move, so that dance is what first comes to mind when evoking music in general. And when it comes to qualifying the quality of music that induces the urge to dance, 'groove' is the most commonly used term. Groove stands as the legacy of all the crossbreeding that has taken place between African-American and Western musical culture during the 20th century, and which, according to McClary (2004), may be its most important feature. Groove stands for this special quality that all of today's popular music genres share, from blues to jazz, funk, rock, rap and electronic dance music (EDM), a quality as easy to recognize as it is hard to describe.

Altogether, both the Avant-Garde and the African-Americanization have greatly influenced the shape of all of today's music forms in one way or another, through a multiplicity of cross-fertilizations between ideas and cultures, and the difficult process of “deconditioning the Western ear in order to allow it to create again, by reestablishing criteria the traditional system has suppressed” (Schaeffer, 1966, p.402, author's translation).

Nowadays, EDM in its broadest sense predominantly carries the torch of innovation in groove, with an output ranging from the most market-oriented productions to the most experimental ones, usually labeled as 'electronica'. Yet, however innovative some electronica artists prove to be, their music doesn't enjoy much recognition from the academic world, just as is the case for groove music in general, and, to my knowledge, no institutional electronic music composer ever took groove seriously. I believe the lack of academic interest in groove has to do with its association with popular music, and therefore is a matter of status.

In her book devoted to experimental electronic music, Joanna Demers provides the following account of the gap between popular music and culture on one hand, and the academic world of scholars and composers on the other hand:

> "There are clear reasons for the relative absence of popular music in academic textbooks, which profile academic music almost exclusively and which only in later editions have acknowledged the existence of nonacademic and popular music. Many people still believe that high culture is more prestigious than popular culture. Popular music often compares itself with art music in order to access some of that prestige, just as art music frequently avoids comparisons with popular music lest it seem trivial or commercial" (Demers, 2010, p. 149).

In an attempt to categorize all experimental electronic music into broad trends, Demers makes the distinction between the following three meta-genres: institutional electroacoustic music, electronica, and sound-art. In analyzing their interrelationship, she makes the following two assessments: "(1) each meta-genre exists in alienation from the other two [...] and from a mainstream culture that barely supports its existence, and (2) alienation from mainstream culture is a sign of aesthetic integrity" (Ibid., p. 147). Each meta-genre thus isolates itself from the others, even though all three are concerned with innovation. Research on electronica is sparse, and composers of electronica proceed in a very intuitive way, building on their practical experience, but often with a serious lack of theoretical knowledge.

As a composer, my intuition is that music would have a lot to gain if groove was promoted as a valid compositional tool by institutional composers, and studied more deeply by academic researchers. My goal is not to provide a general theory of groove, it rather is to dive into literature from different fields (musicology, music perception, cultural studies,...) in order to define a set of rules to generate music I like.

I am trying to program an electronic instrument for live use, in which music unfolds through the interaction between computer and performer, according to the following process: First, the computer generates a sequence of sounds according to the performer's preferences, a set of rules and a fair deal of stochastics. Then, the performer can modify the sequence through a certain degree of control over the rules it obeys to. And, to close the loop, the performer can request the computer to generate a new sequence in the same way, and start piling sequences up by contrasting them with each other.

To my knowledge, there exists no instrument for the live generation and performance of groove. Popular electronic music's most groovy producers usually compose in studio conditions, and what they call a live performance is usually not much more than live mixing of preexisting tracks. The 'IDM'-duo Autechre, though, comes closest to what I am after. Despite the haze they build around their compositional methods, a few of their Max-Msp patches have been circulating the net, indicating their use of generative methods for building their grooves, although their approach is probably intuitive. I enjoy some of their tracks very much for their inventiveness, and beyond the potential discussion on whether their music is to be considered groovy in a traditional sense or not, what matters is their music always sounds new to me. That's what I'm after !

Chapter 1: Groove and meter

1.1. What is groove?

Groove seems to be as difficult to define as it is easy to feel. As a term that originated in popular music slang, it hasn't received much scholarly attention yet, and there exists nothing as a consensual definition yet. Madison, trying to provide a definition in most general terms possible, proposes that groove be "a quality of music that makes people tap their feet, rock their head, and get up and dance. For some music, such as jazz and various kinds of dance music, this is perhaps the most essential feature ..." (Madison, 2006, p. 201). In its widest acceptation, and in popular culture, groove can't be dissociated from the urge to dance.

Yet, when it comes to further identifying the musical qualities that induce this sense of groove, two main trends appear: groove is either considered a matter of repetition, or a matter of expressive musical performance in general, and of 'expressive timing' in particular, a word I borrow from Clarke (2000).

In an attempt to define groove in the context of EDM (and mostly techno), Butler describes the essence of groove as "a short configuration of bass line and percussion that unfolds in continuous repeating cycles" (Butler, 2006, p. 5), and uses borrows the term from the locked grooves, or short loops DJ's use in building up their sets. Yet, given that most EDM is based on a very tight time-grid, his definition almost excludes expressive timing.

Vijay Iyer, a renowned jazz-pianist, gives a completely different account, insisting on the role of expressive timing in groove:

> "In groove contexts, musicians display a heightened, seemingly microscopic sensitivity to musical timing (on the order of a few milliseconds). They are able to evoke a variety of rhythmic qualities, accents, or emotional moods by playing notes slightly late or early relative to a theoretical metric time point [...] In groove-based contexts, even as the tempo remains constant, fine-scale rhythmic delivery becomes just as important a parameter as, say, tone, pitch, or loudness. All these musical quantities combine dynamically and holistically to form what some would call a musician's 'feel' " (Iyer, 2002, p. 398).

In their groundbreaking research on expressive timing, Bengtsson & Gabrielsson (1969) define expressive timing as "systematic variations in duration" from the notated form, variations which indeed are systematic, as they seem to obey to certain rules rather than bare randomness. Sounding rhythm is thus the result of the combination of rhythmic structure, or how a certain music would be notated, and 'expressive timing' (Clarke, 2000).

Expressive timing is a vital part of what musicians call expression, a term that stands for all tonal, timbral, dynamic and temporal variations from the score. Yet, in her introduction to groove in electronic music, Danielsen (2010b) points out that temporal aspects of rhythm can't be dissociated from aspects of sound, such as timbre, dynamics and pitch, grouping all these particular aspects of groove into the larger field of 'microrhythm'.

At this point, and to avoid any terminological ambiguity, I want to mould all previous information into one clear definition: Groove is a musical quality which consists of a certain temporal organization of time in music, a certain kind of rhythm, based on the cyclical repetition metrical patterns, and in which expressive timing, and microrhythm in a larger context, play a central role.

This definition may seem vague at this point, as I introduce many new terms. My goal in the following pages will thus be to provide a clear framework for understanding groove. As a beginning, the very concepts of rhythm and meter should be discussed.

1.2. What is rhythm?

> "The task of those who study rhythm is a difficult one, because a precise, generally accepted definition of rhythm does not exist. This difficulty derives from the fact that rhythm refers to a complex reality in which several variables are fused" (Fraisse, 1982).

The New Oxford American Dictionary proposes the following definition of rhythm :

rhythm |ˈ*riðəm*|
a strong, regular, repeated pattern of movement or sound
• *the systematic arrangement of musical sounds, principally according to duration and periodic stress.*
• *a particular type of pattern formed by such arrangement*
• *a person's natural feeling for such arrangement*
• *the measured flow of words and phrases in verse or prose as determined by the relation of long and short or stressed and unstressed syllables.*
• *a regularly recurring sequence of events, actions, or processes*
• Art *a harmonious sequence or correlation of colors or elements. (New Oxford American Dictionary, 3rd Edition)*

The first three sub-definitions relate to musical rhythm. The last three indicate that the term can also apply in various non-musical contexts. The fifth provides the most general definition, as relating to event, actions, or processes. Whereas in the first the nature of the arrangement of musical events is not mentioned, in the fifth, rhythm is assumed to be recurring. Together these two definitions come very close to my own conception of musical rhythm.

Yet definitions of musical rhythm seem to be as numerous as their scholars, and I will just sum a few to make clear what I mean: Smoll (1973) defines musical rhythm as "the perception of a series of stimuli as a series of groups of stimuli", Radocy defines musical rhythm as "the pattern of organized sounds and silences" (Radocy, 1980), Martin proposes that "the perception of early events in a sequence generates expectancies concerning later events in real time." (Martin, 1972), and Schwanda that rhythm is "movement that is aesthetically pleasing to the observer as well as to the one being observed" (Schwanda, 1969). All definitions are taken from Waadeland (2000, pp. 15-16).

As clearly appears from these examples, existing definitions of rhythm are vague and divergent. In fact, rhythm can only be described in contrast to its perceptual counterpart, meter.

1.3. Meter, a framework for understanding rhythm

When listening to a isochronous sequence of perfectly identical sounds, the human brain starts grouping them in groups of 2 or 3, and that one of them, usually the first, seems to sound more intense and to last longer. This phenomenon, which has been observed more than a century ago, has been termed subjective rhythmization. As soon as a difference is introduced into that isochronous sequence, the human brain automatically organizes them as groups between two repetitions of that difference, a phenomenon by contrast termed objective rhythmization (Fraisse, 1982). Rhythmization is a good illustration of the broader concept of meter, in that meter is a mental construct our brain derives from the rhythm present in the music.

There exist different conceptions of meter, but altogether it is easy to mould them into a broader one, which I will try to achieve now. Large (2008, p. 190), to begin with: "In music perception and cognition [...], pulse and meter refer to percepts. They are responses to patterns of timing and [...] stress in the acoustic rhythm." He doesn't define rhythm in itself, yet makes clear he considers rhythm some objective organization of acoustic information ('acoustic rhythm', 'rhythmical surface') , and pulse and meter mental constructs or tools for us humans to interpret this information. London (2002, p. 531) defines meter as a "stable, recurring pattern of temporal expectations, with peaks in the listeners expectations coordinated with significant events in the temporally unfolding musical surface." Meter is composed of nested internal clocks which we synchronize to the rhythmical surface, in the same fashion that we synchronize to recurring events in our environment in general, the way our attention functions. As Jones notes:

> "Attention is cast from some reference event at one point in time toward a target event scheduled for a later time. [...] We continually cast ourselves forward by rhythmically anticipating future events that may occur within small and larger time intervals. These paths form the patterns of mental space and time and so can establish for us that sense of continuity and connection that accompanies comprehension" (Jones, 1981, p. 571).

This mechanism, which, after London (2012) I will name 'entrainment', is crucial to our survival as it allows us to predict danger by focusing our attention to changes in our environment. More information on the psychological mechanisms and the importance of surprise in music can be found in Huron (2007).

Sound only translates into music when it provides the listener with a comfortable balance between predictability and surprise, which varies from one listener to another, depending on the listener's attentional acuity, which can be trained. In order to perceive a series of sonic events as a rhythm, the listener needs a mental framework that allows for such predictions, and this mental framework is called meter. In the following pages, I will dive deeper into the exact mechanism of meter in relation to our perception, and the discrepancy between classical and groove approaches to meter. As we will see, rhythmical musics are always conceptualized according to a certain model for meter, and models for classical and groove musics show strong divergences.

1.4. The classical theory of meter

The disciplines of musicology and music cognition, as all modern sciences, have originated in the hegemonic West, and through the ages research has mainly focused on Western art-music between the Renaissance and the 19th century, which I will refer to as 'classical music'.

I will start this discussion of meter by introducing Lerdahl and Jackendoff's (L&J) 1983 book, *A Generative Theory of Tonal Music,* which accounts as one of the most influential books in modern music analysis, inspired by Chomsky's universal grammar theory, and whose title clearly limits its scope to classical music as defined in the previous paragraph.

L&J regard meter as a regular, hierarchical pattern of strong and weak beats to which the listener relates events at the musical surface, which they call phenomenal accents. "The listener's cognitive task is to match the given pattern of phenomenal accentuation as closely as possible to a permissible pattern of metrical accentuation. [...] Metrical accent, then, is a mental construct, inferred from but not identical to the patterns of accentuation at the musical surface." (Lerdahl & Jackendoff, 1983, p.18)

Their theory is “generative” in that it provides a clear set of formal procedures to derive a metrical description of a piece of music from the score, a description which classifies all notes according to their structural stress in a hierarchical fashion (Waadeland, 2000; Patel, 2007). Central to the metrical hierarchy lies the pulse, or tactus, “one specially designed metrical level, [...] the level of beats that is conducted and with which one most naturally coordinates foot-tapping and dance steps” (L&J, 1983, p. 71). The speed of the pulse is described by the tempo.

Fig. 1.1 provides an example of a typical analysis of a piece according to L&J's rules :

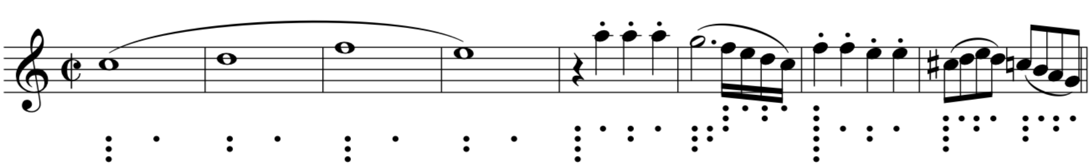

Fig. 1.1: Metric analysis of the opening measures of the finale of Mozart's 41st Symphony IV, after Lerdahl & Jackendoff (London, 2004)

Typically, such an analysis is either represented by a tree structure above the score or, as in this example, by dots under it. The dots associate structural accents to measure subdivisions, on different hierarchical levels, according to above-mentioned set of rules. Two adjacent levels' periods always have simple integer ratio relationships, usually as 2:1 or 3:1, so that more complex ratios will always be brought back into lower components (4:1 for example would be split into two 2:1 levels). Obviously, polyrhythms and complex tuplets, which will be discussed later, are hard to fit into this scheme.

The first four measures in fig. 1 are analyzed as having a metrical structure composed of half, whole and double whole note levels, and the more levels cross on one note, the more the note will be experienced as being accented. The metrical structure is subject to change, as happens several times from measure 5 on. Coinciding with the introduction of new note subdivisions at the musical surface, new metrical subdivisions appear.

In fact, L&J's set of rules is based not only on metrical structure, but also the *grouping structure* that “expresses a hierarchical segmentation of a piece into motives, phrases and sections” (Lerdahl & Jackendoff, 1983, p. 8). Cambouropoulos provides the following summary of the grouping structure:

> “The concept that rhythm relates to cognitive grouping of musical events is a Gestalt-based one. The Gestalt principles of perceptual organization are a set of rules-of-thumb that suggest preferential ways of grouping mainly visual events into larger scale schemata. Two of the Gestalt principles state that objects closer together (Proximity principle) or more similar to each other (Similarity principle) tend to be perceived as groups” (Cambouropoulos, 1997, p. 278).

So, in combination with each other, and according to L&J's rules, the accentual structure of a piece can be derived from its metrical and grouping structures. As we will see, rhythm is always to be interpreted as a combination of both meter and grouping, and these two perceptual structures continuously interact with each other. Grouping will further be discussed in chapter 2.

L&J's represent meter as a time-discrete grid. Even though such a discrete concept of time proves very powerful for modeling meter, it has to be nuanced.

1.5. Meter as a continuous process

Time is continuous, any event can happen at any time, and therefore any time-discrete representation has to be regarded an approximation of reality. Large & Palmer (2002), drawing on L&J's hierarchical model, propose a time-continuous model of meter composed of simultaneous nested internal clocks, of which Fig. 1.2 provides a graphical illustration:

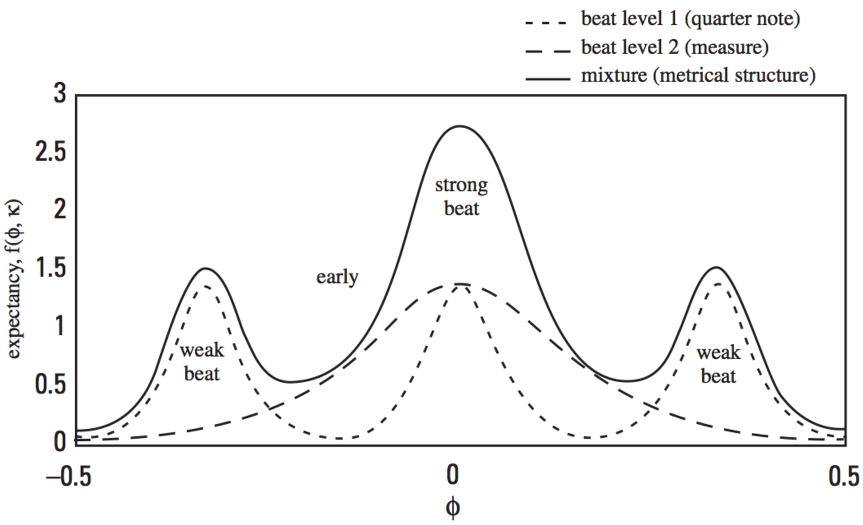

Fig. 1.2: Model expectancies for a ternary meter (3:1 period ratio), based on von Mises distributions, after Large & Palmer (2002), (from London, 2004, p.21)

Each metrical level is modeled by and oscillation, and each individual oscillation is self-sustaining, in that "once established, (it) tends to be continued in the mind and musculature of the listener, even though . . . objective pulses may cease or may fail for a time to coincide with the previously established pulse series" (Cooper & Meyer, as cited in Large & Palmer, 2002). "The job of the oscillator is to synchronize with the external rhythmic signal. However, it does not respond to just any onset as a potential beat; it responds only to onsets in the neighborhood of where it expects beats to occur. Thus, it has a region of sensitivity within its temporal cycle whose peak or maximum value corresponds to where the beat is expected" (Large & Palmer, 2002, p. 7).

Depending on whether an event occurs at (or around) the peak or the valley of one or another level's oscillation, it will be perceived as belonging to and accented according that level, or to simply be out of meter. When sustained deviations from the expected position occur, our brain adapts phase and/or period of the oscillations in order to stay tuned, thereby building a new sense of meter. If the link between the rhythmical surface and the sense of meter is broken, "certain kinds of pathological "enstrangement" from the music can occur, [...] a form of amusia in which the music is subjectively reported to sound like "noise", or be "flat" and uninteresting" (Wilson & Pressing, 1999, in Pressing, 2002). Meter is thus essential to the aesthetic appreciation of music, and its sense should thus be maintained at any price. Of course, some listeners prove better at perceiving meter as others, and similarly some prove quicker at adapting their meter than others. Therefore, Butler (2006) makes the distinction between conservative and radical listeners, depending on their ease at performing such reevaluations or at losing the meter completely.

Yet to challenge the listener's sense of meter proves to be a powerful tool in making music sound more interesting, and such strategies, which Butler groups under the term 'metrical ambiguity', are very common in groove contexts.

1.6. Meter : Perceptual limits and the role of memory

If the perception of meter varies from one listener to another, crucial factors in this perception, such as the music and the language we are confronted with, are often culture-specific. I will just cite a few examples, all taken from London's book on the psychological aspects of meter (London, 2004), to give an idea of possible divergences:

- The ability to detect a pulse is functional at birth, and infants in low show the same capabilities in interpreting music from all cultures. Yet, from the age of 1 year, abilities for other culture's musics decline.
- In comparing several composers' work, correlation was found between rhythmic variation and the composer's native language. Similarly, English listeners tend to group non-isochronous sequences of tones as short-long, and Japanese as long-short.
- Preferred tempos slow down with age, while the range in which we adapt and engage widens.

Therefore, an important degree of subjectivity, both by listeners and performers, is involved in interpreting a same piece of music, and L&J's analyses are merely to be considered a probable interpretation. Virtually, the tempo of a piece itself can be interpreted differently, as any periodicity present in the signal can be heard as a pulse. Yet, the perception of pulse has temporal limitations. London (Ibid.) provides an extensive overview of research on the subject, of which I present the following summary, all values being approximative :

Discerning 2 events:	*> 2 ms*		
Ordering 2 events:	*> 20 ms*		
Meter:	*100 – 6000 ms*	|	*10 – 600 bpm*
Pulse:	*250 – 2000 ms*	|	*30 – 240 bpm*
Preferred tempo range:	*500 – 700 ms*	|	*85 – 120 bpm*

Two sonic events can only be discerned from each other if separated by a minimum 2 ms. In order to perceive the order in which they appear, at least 20 ms are needed.

To perceive a regular stream of events as a metrical level, a minimum of 10 bpm is needed, below that level, the stream's coming events cannot be anticipated, the internal clocks don't work, and our predictive ability becomes less and less efficient. The same happens above 600 bpm, so that sequences occurring at higher tempo's will be perceived as groups and/or related to lower metrical levels (note that 600 bpm involves durations of 100ms, or 10Hz, which is close to audio rate, a temporal domain where periodicities are perceived as pitches instead of durations).

To perceive a pulse, and thus a sense of tempo, the range is much smaller. If the perception of pulse ranges from 30 to 240 bpm, pulses are preferably chosen in the range 85-120 bpm, the range of spontaneous foot-tapping, close to the heartbeat and walking rates, suggesting a link between tempo and certain bodily functions.

One major consequence of these natural limits is that the maximum number of metrical levels above and below the pulse is tempo-dependent. London provides a clear graphical illustration of how all possible sub- and supra-pulse levels fall into these perceptual categories, in this case at a tempo of 92 bpm (Fig. 1.3):

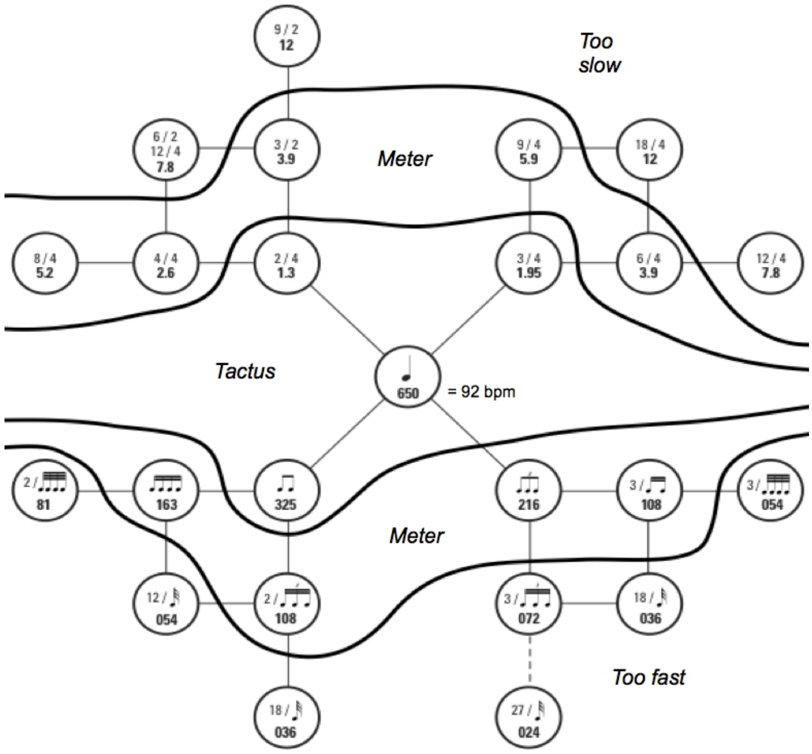

Fig. 1.3: Tempo-dependent graph of multiply related metrical levels, after London (2004)

These perceptual limits of meter are directly related to how our brain works. In his 1993 essay, Brower takes our different types of memory (echoic, short-term and long-term) as a starting point, and describes each memory's contribution to our perception of rhythm.

The echoic memory serves as a precognitive store for sensory information. Echoic memory thus functions as a buffer for sensory information before our brain starts processing it. When listening to metered music, our brain takes the pulse as a reference for processing information, and therefore the echoic memory coincides with small sub-pulse rhythmic groups.

The short-term memory (STM) is a store for 5 to 9 chunks of information, and corresponds to our notion of psychological present. By internalized information, he means information that can be labeled according to former experiences. For example, if 9834 accounts as 4 chunks, 1945 to most people will account as just one. In music, the STM usually coincides with a measure or a phrase, and the chunks can coincide with the sub-pulse groups stored in the echoic memory.

Beyond the temporal span of the STM comes the long-term memory LTM, which corresponds to the past. It is a permanent store for information, and consists of two subtypes: The episodic and semantic memories. As the episodic memory serves as a store for sensory information, about event and their context, its role in appreciating music is to identify and recognize recurring events. The semantic memory is a store for conceptual knowledge, and therefore helps us in understanding large-scale musical structure.

These three types of memory play a crucial role in our perception of time in music, in that our internal clocks are very precise in maintaining the pulse, that sub-pulse timing shows significant distortions both at the perception and the production of rhythm, and that supra-pulse rhythmical periodicities are perceived as concatenations of beats. As Iyer summarizes :

> “ [...] Rhythmic material below [the pulse level] is perceived categorically as combinations of subdivisions of a main regulating pulse, and durations above it are considered to be on the level of metric grouping of pulses. By this division, echoic memory covers the immediate timescale of rhythmic activity, whereas short-term or working memory covers meter and phrases. These different types of memory involve different kinds of processing. We *entrain* to a pulse based on the echoic storage of the previous pulse and some matched internal oscillator periodicity; we *feel* the relationships among strong and weak beats (accentual meter); we *count* times between phrases or bars (metric grouping); and we *recognize* sub-pulse rhythms qualitatively.” (Iyer, 2002, pp. 395-96)

1.7. The dichotomy between rhythmical structure and expressive timing

From the score to an experience, rhythm undergoes the performer's interpretation as well as the listener's perception. It can be difficult to separate these two cognitive and motoric processes, but fact remains that the musical experience doesn't have much in common with the notated rhythm. I now want to show a few results from empirical research to illustrate the gap between the notated and perceived rhythms.

In a summary of previous research, Fraisse (1982), besides introducing the concept of rhythmization I mentioned already in section 1.3, explains that whichever sequence of sounds we are confronted with, we tend to experience their rhythm as a sequence of just 2 time intervals: long and short, the first being perceived as twice as long as the second, in a subjective ratio 2:1. Yet when confronting listeners with sequences made of just two durations, and when trying different objective ratios, he found that listeners had a preference for objective ratios around 1.75:1, even though they would describe them as being 2:1. This clearly illustrates the gap between notated music and music as it is perceived, and gives a good idea of the difficulties that arise in establishing a theory of meter.

Yet interestingly, this ratio of 1.75:1 is analogous to the average swing ratio in jazz as it has been measured many times by several researchers. Swing consists of performing a notated rhythm by delaying every second 8^{th} note, thereby creating the 'feel' peculiar to the genre. In notated form, swing in jazz is either presented as regular eight notes or a ratio 1:1, or as triplets, a ratio 2:1, whereas effectively performed swing usually lies in between these two ratios. *Audio example 1* is a small recording of my instrument with a fixed swing ratio of 2:1.

In line with such empirical research, Honing (2002) proposes to confront listeners to various recurring series of three durations, hereby requesting them to indicate how they perceive the rhythm. The results could then be represented graphically by means of a 'performance space', of which Fig. 1.4 is an illustration:

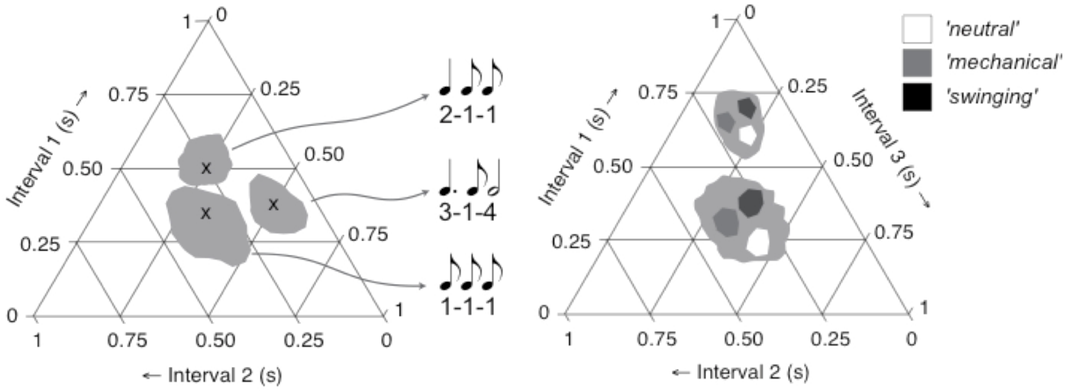

Fig. 1.4: Performance space with (a) structural interpretations of rhythmical sequences and (b) interpretations of feel (Honing, 2002)

The performance space represents all combinations of three successive durations in a flattened three-dimensional space, and each axis expresses on of the three durations, relative to the cycle length. Each point in this space is thus associated with three values that sum up to 1. For example, the values associated with the cross in the middle of the 2-1-1 region are 0.50 for interval 1, 0.25 for interval 2 and 0.25 for interval 3.

Fig. 4(a) shows regions that listeners associate with three different notations. Besides the nominal values, there is thus a wide range of values which would be interpreted as corresponding to a certain notation. Note that the gaps between the zones are narrow, indicating that few combinations are perceived as ambiguous. Fig. 4(b) shows how regions can be subdivided into subregions, according to the feeling they convey to the listener.

It is obvious that the regions' positions and sizes do not correspond from one drawing to another, which makes clear that this representation is not based on systematic empirical results. Yet I want to trust that a plot based on measurements would take a similar shape.

In either case, the close link between preferred swing ratio and swing as it is performed in jazz shows a clear duality between perception and the established music notation system. In fact, if the musical notation comes close to how rhythm is conceptualized, it is far from the actual rhythmical surface. The rhythmical surface can thus be considered a combination of the rhythmical structure, of music as it is notated and perceived, and expressive timing, or variations around that structure.

1.8. Tempo variations vs. event shifts, two paradigms for expressive timing

Having introduced the dichotomy between rhythmical structure and expressive timing, I want to present a model for expressive timing in classical music : the KTH system, as proposed by Friberg, Bresin and Sundberg (2006). There are a few other models, but this one seems to be the most elaborate, despite the fact the results don't always sound convincing. It is relevant to take a closer look at its set of rules, as a source of inspiration for my model of groove.

The KTH system aims at adding expression to the rendering of MIDI-files, as musical pieces in MIDI-format usually come without any expression, have flat dynamics, mechanical timing and are reproduced following a rigid well-tempered scale expressed in integer MIDI note numbers. It therefore proposes a large set of rules, all based on conclusions from past empirical research. The rules range from phrasing and timing rules to tonal tension and intonation rules. I will enumerate a few, limiting myself to the ones directly connected to timing, as melody isn't my primary concern :

- **phrase arch**: apply a time-arch to tempo and level for each phrase
- **final ritardando**: slow tempo down at the end of each section
- **duration contrast**: make short notes shorter, and long notes longer
- **tempo uphill**: accelerate when pitch is rising
- **double duration**: when encountering a duration sequence of 2:1, bypass the duration contrast rule and make the long note shorter and vice-versa
- **inégales**: create long-short patterns for equal note values (swing)
- **punctuation**: introduce a final lengthening and a micropause at the end of each short melodic fragment
- **performance noise**: simulate motor inaccuracies

All the rules are based on the shortening of lengthening of notated durations. Put differently, this means that the model applies expressive timing by constantly changing the performed tempo. Indeed, in classical music, the tempo changes all the time, and composers often indicate intended tempo changes in the score: think of rubato, accelerando, etc. Yet in groove music, the average tempo is more or less constant, and if tempo changes occur, they occur on much longer time-spans as in classical music. Think of slow and progressive accelerandi stretching over several measures (*Audio Examples 2 and 3*, provided by the authors, allow to compare a piece by Bach as performed without and with the KTH rules respectively).

In this context, expressive timing can in fact as well be modeled by event shifting as by tempo variations (Bilmes, 1992), as any variation in the interval between two events can either be considered a tempo change or a shift from the expected position. Honing (2002) provides a clear graphical illustration of both approaches (Fig. 1.5).

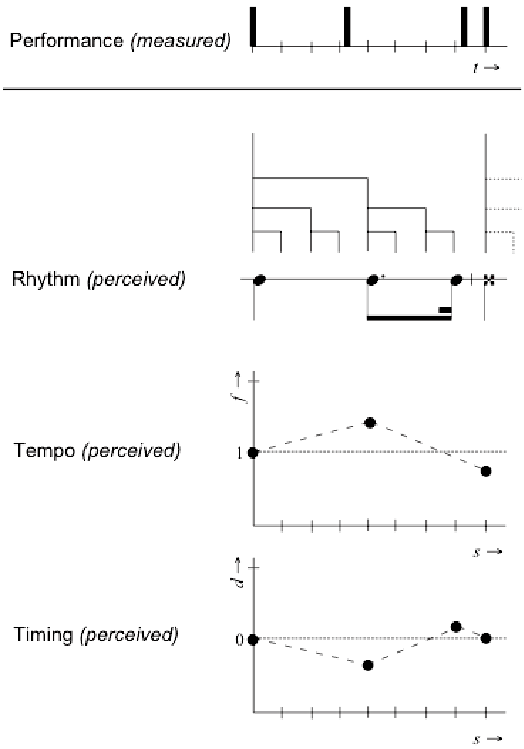

Fig. 1.5: Components of performed rhythm (Honing, 2002)

This illustration clearly shows that, despite its timing deviations, a performed rhythm (first drawing) is interpreted by the listener as a set of durations with integer ratios to each other, or rhythmical structure, as in notated form (second drawing). In this respect our brain performs the same kind of process as the 'quantize' function in a DAW. The question here is whether timing variations occurring around the rhythmical structure will be perceived as tempo variations (third drawing) or as event shifts (fourth drawing). If a tempo-based model such as the KTH-rules comes close to our perception of classical music, an event-shift model proves more adequate for groove-based musics, as the tempo is more stable in these musics. Yet for having tried repeatedly to make sense of music according to one or another paradigm, my feeling is that perception is in fact based on both, so that our internal clocks perform both phase and tempo adjustments during the course of the music. Interestingly, it seems that any, even random, recurring timing pattern tends to be perceived as more and more comfortable with repeated listening, confirming Kvifte's (2007) theory that non-isochronous pulses can also serve as a meter.

Yet, in constructing a model, choices have to be made, and, like many others, I will opt for a fixed-tempo, event-shift model for my grooves, keeping tempo variations and non-isochronous pulses in the back of my mind.

As it has been defined in these pages, groove is based on the cyclical repetition of certain patterns. As we have seen, the repetitive character of groove has repercussions on the choice of a timing model. It is thus time for a brief discussion of repetition in music in a cultural context.

1.9. Groove as repetitive music

Western art-music, since the Baroque period, has often been described and experienced in terms of a teleological process: “The music seems to form large-scale curves of tension, often building up to a climax before reaching a conclusion that is felt to be well prepared and quite natural. However, this is really only one way of shaping musical time, and it is in fact something of a historical parenthesis. Linear, teleological, or “closed” musical forms are in a larger historical and geographical context the exception rather than the rule” (Danielsen, 2006, p.150). Teleology in Western music didn't appear by chance, and many scholars find its roots in the general evolution of Western culture. In an influential essay, Kramer provides a lively illustration of this:

> “Ideas of cause and effect, progress, and goal orientation have pervaded every aspect of human life in the West at least from the Age of Humanism to the First World War. Technologies, theologies, and philosophies have sought to improve human life; capitalism has sought to provide a framework for material betterment, at least for the few; science has been dominated by the temporally linear theories of Newton and Darwin; even our languages are pervaded by words that refer to goals and purposes. [...] Listening to tonal music has become comfortable to Westerners not only because we have learned a complex skill but also because the linearity of tonality neatly corresponds to many goal-oriented processes in life. [...] Balinese music, like Balinese life, is not oriented toward climax. Activities in Bali are understood and appreciated not as means toward goals but rather as inherently satisfying” (Kramer, 1981, pp. 539-40).

Kramer subsequently stresses the distinction between two main kinds of temporality : linear or non-linear, teleological or not, and at the extreme opposite of linear lies vertical time, “a single present stretched out into an enormous duration, a potentially infinite "now" that nonetheless feels like an instant” (Ibid.:549), a description that evokes drone music.

On the axis between these two extremes lies a vast field of possibilities, which Fink (2005), in a fascinating book on minimalism, names 'recombinant teleologies'.

One example of recombinant teleology is what Kramer calls discontinuous time, common in 20th century avant-garde music (e.g. Webern, Feldman or Cage), in which discontinuities segment and reorder linear time, with a loss of large-scale teleology as a result.

Leonard Meyer was one of the first to recognize the profound implications of Cage's chance music, which he considered “anti-teleological”:

> “The music of the avant-garde directs us toward no points of culmination [...] It arouses no expectations, except presumably that it will stop [...] Underlying this new aesthetic is a conception of man and the universe, which is almost the opposite of the view that has dominated Western thought since its beginnings [...] Man is no longer to be the measure of all things, the center of the universe [...] His goals and purposes, his egocentric notions of past, present, and future; all these are called into question. For these artists, writers and composers, *the Renaissance is over*” (Meyer, 1967, in Fink, 2005).

Discontinuous time thus had major metaphysical implications. At the time of Cage's chance music, repetition, another kind of recombinant teleology, had long made its apparition in Western art-music, through the works of composers like Debussy and Stravinsky for example, and was far from being well accepted by everyone:

> “Cultural critic Theodor Adorno, who worked closely with the musicians of the Second Viennese School, explained why Schoenberg fought so vehemently against instances of repetition: if we understand a piece of music as an allegory of personal development, then any reiteration registers as regression [...] Adorno critiqued Stravinsky severely – not for his dissonance or departure from traditional practices, but for the hypnotic effect of his repetitious ostinato patterns, which Adorno heard as seducing listeners into passive acceptance of the most barbarous elements of encroaching totalitarianism. He believed that when audiences give up the admittedly difficult task of critical thinking, then the path is paved for demagogues like Hitler or Stalin” (McClary, 2004, pp. 291-92).

The use of repetition in music has remained controversial in Western art-music ever since, but quickly became the norm in popular music through the process of African-Americanization, and the concept of groove became more and more common.

Groove is “built around the extensive repetition of relatively short musical segments”. (Hughes, 2003, p. 15). This cyclicality of time is indeed a feature all groove-based musics share. Yet “a groove does not stand still. Even though the principles comprising it are supplied in advance [...] and even though the groove is not proceeding toward a definite goal, it is – to the last section – in motion” (Danielsen, 2006, p.155), to the point that, like Fink (2005), one could argue that each repetition is in fact a climax in itself.

The appreciation of repetition in groove music greatly depends on whether the listener is directed toward difference rather than similarity, “not difference in itself but difference stepping forward in relation to the same, to a figure, a formality or convention, perhaps even tradition itself. [...] In other words, repetition in a groove is a sort of micro-level signifying: it is repetition and revision in one and the same maneuver. The aim is two-fold. On one hand, it is important that the same is repeated every time and recognized or categorized as such. On the other hand, it is equally important that this same is different. [...] The difference is a difference within the repeated same" (Ibid.:159). Repetition in groove is thus a framework within which to express emotions by means of subtle changes and small variations, the kind of restriction that enhances creation, and it is not completely alien to teleology after all, as it is always going somewhere. If not at the rhythmical surface, then always in the way it is perceived.

1.10. Rhythm and meter in African music

In a paper which can be seen as a summary of concepts described earlier in *African polyphony and polyrhythm*, Arom (1989) enumerates the basic rhythmic features that prevail in music from Central Africa. Here follows my interpretation, which I contrast with previously developed concepts about classical music:

- **Steady, regular tempo** with no accelerandos, rallentandos or rubatos, *as opposed to phrasing in classical music*.
- The predominance of repetitive, uninterrupted formulae, in which similar material appears at regular intervals, as evidence of **strict metrical periodicity,** *as opposed to L&J's changes of metrical structure.* Yet there is a certain degree of variation of the rhythmical surface in relation to the meter.
- Diagonal ordering of individual parts, crossing, interweaving of individual rhythms, exemplified by the famous **polyrhythms**, *as opposed to the hierarchical (or vertical) ordering in L&J's system.*
- **Absence of a system of regular accentuation**, *as opposed to L&J's systematic accentuation for each level.* Arom therefore advocates the term **period** to replace the old concept of *measure.*

Arom defines meter as "a sequence of equal unaccented values with no tone color differentiation" (Ibid., p. 93), in contrast to rhythm, which appears as soon as one introduces different durations, accentuation, or tone colors (to Arom, tone color is the combination of pitch and timbre, a concept more suitable than simple pitch). Arom's notion of meter boils the complex hierarchical concept of L&J down to just one level, the pulse :

> "The pulse is an isochronal standard, which is used by Central African cultures as the unit of reference for the measurement of musical time. It provides a series of regular reference points for ordering rhythmic events. In polyrhythmic music, the pulse is the common regulator of temporal organization for all the parts. It is thus the basic unit of time according to which all durations are defined." (Ibid,. p. 92)

The pulse can thus be divided in different ways, by 2, 3, 4, 5 and sometimes 6. This divisions results in the smallest operational value, as he calls it, or 'density referent' as I will call it after Jones (1954, In Temperley, 2000). To generate polyrhythm, pulses group into periods of different length and which relate to each other by simple ratios, usually 1:2, 1:3, 2:3 and 3:4. When superimposed with each other, these individual periods result in a 'macro-period', the time it takes for all individual periods to meet in phase again, as fig. 1.6 illustrates it.

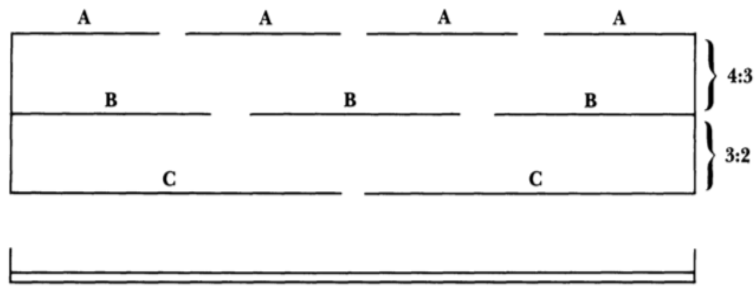

Fig. 1.6: Individual periods and macro-period in polyrhythmic music (Arom, 1989)

According to Arom, "the interweaving of accents and tone colors, together with the absence of a reference system of regular accentuation, creates a feeling of uncertainty and of ambiguity regarding how the subdivision of the period is perceived" (Ibid., p. 91). In fact, to achieve what we will name 'metrical ambiguity' (Butler, 2006) is exactly what sub-saharan music is all about.

1.11. Polyrhythm and polymeter

In his classic book *African rhythm and African sensibility* (1979), Chernoff, who has spent many years learning the art of drumming in Ghana, attempts to explain the differences between Western and African music from an anthropological point of view, stressing the importance of cultural factors in the perception of music. For him, the conflict of rhythmic patterns is the main principle of African music. "There are always at least two rhythms going on" (Chernoff, 1979, p. 42). "The basic organization of rhythms is the essential composition. [...] The coherence of the conflicting rhythms is [...] based upon a kind of tension which gives the music its dynamic power" (Ibid., p. 53), a tension such that "the listener must be actively engaged to make sense of the music" (Ibid., p. 50). For Chernoff, "the only way to hear the music properly [...] is to listen to at least two rhythms at once" (Ibid., p. 51), a faculty

Western ears are less trained for. One should even speak of polymeter instead of polyrhythm, in that performers of conflicting rhythms have different meters in mind: “Polymeter is a mental construct. It could be notated as polyrhythm, but that's not the way the performers think of it” (Ibid., p. 47). The essence of African music is thus the conflict between different meters and the interpretational ambiguity that flows out of it, and the music can only be good if that ambiguity, the balance between conflicting meters is maintained. In fact, once the patchwork of repetitive patterns and conflicting meters is established, the whole music consists of a game of ambiguity, where one or another meter is put forward by the musicians, who “highlight various rhythms to increase the effect” (Ibid,. p. 53). I insist on Chernoff's conception of conflict because it fits my own conception of African music very well, and I want to keep it in mind for later in this text.

Yet, when it comes to choosing between the concepts of polyrhythm and polymeter, I have my doubts. In analyzing microrhythm in a few African music recordings, which analyses I will present further on in this text, I have found that microtiming and accentuation patterns of different parts at one moment of the piece usually point toward one definite pulse, which thereby is assigned a dominant role through performance, consistently altering the ambiguity for the listeners. This does not mean a dancer can't choose for another pulse at any time, but it implies that the way the music is performed always points towards one unique dominant pulse. Leaving this discussion aside, we can for nor consider the terms 'cross-rhythms', 'polymeter', 'diagonal ordering', etc. as all referring to the same phenomenon, which I from now will refer to as 'polyrhythms'.

To conclude this section, it is contributive to show an illustration of how the dilemma between polyrhythm and polymeter can be expressed in Western notation of African music, depending on how notes are grouped (Fig. 1.7).

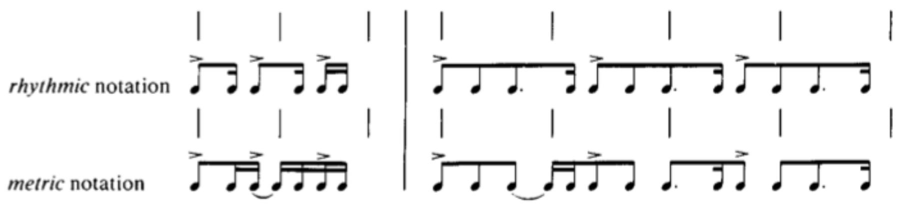

Fig. 1.7: rhythmic vs. metric notation of polyrhythms (Arom, 1991)

The rhythmic notation groups the notes according to accentuation, putting more stress on the individual parts' pulses, highlighting repetition but thereby blurring the main pulse and decreasing vertical readability. The metric notation does the exact opposite. This notational dilemma points to another shortcoming of Western music notation, and one should be aware of it in interpreting scores.

Chapter 2: Making sense of random timing

2.1. What density referent?

Central to most studies on African rhythm is the notion of density referent, the lowest level duration which all other durations in the music are a multiple of, and which remains throughout the whole piece. But what about tuplets then? Western classical music allows for changes in sub-beat divisions by the use of different tuplets, mostly switching duple or triple. Indian music is based on all kinds of tuplets, with many ratios based on prime numbers like 5,7,11 and so on. The concept of density referent, although very useful in considering sub-pulse timing in African rhythms, thus proves deficient for other kinds of rhythm.

It is interesting to note that Lerdahl and Jackendoff's model, which has proven too strict on accentuation patterns for musics outside the Western tradition, does allow the metrical hierarchy to change over time, so that some levels may appear or disappear according to changes of density in the rhythmical surface, as well as they can change, and adapt to new pulse subdivisions.

A discussion of timing variations based on the dichotomy between syntactical and sub-syntactical structure, between rhythmical structure on one hand and expressive timing on the other hand, is not possible without clearly defining what the structure exactly is. If on beat level the structure is clearly defined in relation to the pulse, how to model sub-pulse timing in absence of any clearly defined density referent?

2.2. Beat grouping

The first question that came into my mind when starting this research on microtiming was: and what if the whole concept of rhythm based on integer ratios between durations is wrong? What is any random division of pulse can prove efficient?

It is easy to generate random non-integer (float) timing patterns, and to my knowledge no researcher has ever tried this otherwise obvious approach. I have thus generated 12 sequences of 7 identical sounds (which means they are identical in all their attributes, be it length, pitch timbre and dynamics), recurring every 1.17s, and separated in time by random periods (SuperCollider code: *({3.0.rand}!7).normalizeSum*), have listened to them, and taken conclusions. I am the only person who has listened to them at this point, so my conclusions are only to be regarded as indicative. I provide the full series of sequences as *audio examples 4 till 15*, in the order they were generated, without discarding any of them.

It is striking at first hearing that these sequences are perceived as rhythm. Some are more confusing than others, but with repeated listening they seem to make more and more sense, seeming more and more comfortable with time.

Two perceptual phenomena obviously take place: 1) A sense of pulse appears, and 2) Although they appear gradually through a slow fade-in, each recurring sequence as it is perceived is assigned a beginning point (note that it is possible to focus on different events as potential beginning point).

Even though each sequence induces a sense of pulse, the tempo of these pulses is not common to all sequences. At first hearing, I spontaneously classified the sequences according to different integer period subdivisions, being :

- 5 beats: sequence 2
- 3 beats: sequence 12
- 4 beats: sequences 1,3,4,5,6,7,9,10 and 11
- 7 beats: sequence 8

At closer hearing though, I found I can also create a sense of 4 beats pulse on sequence 2, and a sense of 3 beats on sequence 1.

To me, this demonstrates that we spontaneously seek to organize all temporal sequences of sounds according to some pulse, demonstrating the theoretical elements exposed in previous pages. If at first, this pulse is perceived as irregular, it seems more and more regular with time, in an attempt to seek regularity at any price. In opposition to Kvifte, this provides evidence that the perceived pulse is always regular, even though we tend to bend our expectations to make the rhythmical surface fit. Indeed, even though the resulting rhythms may feel more or less comfortable compared to the pulse, and even though comfort comes with repeated listening, there always appears a sense of regular pulse.

When considering the different beat grouping categories, and thus the different pulse tempos, it is remarkable that most sequences were classified as having 4 beats per period, and rarely by groups of 3, 5 or 7. Obviously, I, if not everyone, preferably attempt to perceive beats by groups of 4, trying to assign as many events as possible to beats according to this scheme, and only consider other beat groupings in case this attempt fails. Yet individual factors are not to be excluded, and in this context it is important to mention that I mostly grew up with techno music, which is predominantly grouped by 4, and usually runs at relatively fast tempos.

Nevertheless, from this experiment I conclude that there is a strong natural tendency to synchronize attention to a pulse that groups beats by 2 or its multiples, and then by 3 or its multiples, and that higher prime numbers only come into consideration in case this strategy fails. There is thus ground for the predominance of duple and triple meter in classical music, in that it expresses a natural preference.

If, from an ecological point of view, one considers entrainment as a mechanism originally aimed at predicting events in our environment, as exposed in section 1.3, it is a judicious strategy to focus on duple division of time, as it is the most likely to occur in nature, duple division, similar to oscillatory movements such as walking, breathing or the succession of day being predominant (London, 2012).

Soon after classifying the random sequences, I wondered what would happen when combining multiple sequences that I categorized according to the same pulse. Depending on the point of view, the results can be considered either chaotic or complex. As, besides the first, none of the events in one sequence coincides with the other sequence's events, combining both has to result in more complexity, blurring the fragile sense of pulse. To maintain a clear sense of pulse, it is thus important to keep a certain synchronicity of events on and around beat locations.

2.3. Sub-pulse Gestalts

The loops in the previous sub-section recur every 1.17s, a value which fits into the perceptual limits for pulse (see section 1.6). Nevertheless, I tend to assign them faster pulses, subdividing them into several beats, thereby perceiving these loops as periods. Possible explanations for this may include the influence of personal tempo preferences or the influence of the density of events on the tendency to perceive faster tempos, but in either case I decided to generate shorter loops which I unmistakably perceive as consisting of 1 beat.

The experiment on longer loops has provided me with some insight on beat induction, but the question of beat subdivision, or sub-pulse timing, remained open. I expect the present experiments on shorter loops will draw some light on that very question.

The 19 loops I present now have been generated in the same fashion as the preceding ones, only with fewer events and with a shorter cycle-length. There are 4 loops with 2 events, 5 with 3, 6 with 4, 2 with 5, 1 with 7 and 1 with 11 events. The repartition and the choice of the number of events is not systematic. I again just generated them and present all without discarding any, ordered by increasing density. *Audio examples 16 till 34.*

120 bpm being a comfortable tempo, and because the period length coincides with it, my pulse synchronizes with the loop length. I always spontaneously synchronize its phase with one of the events. In fact, it soon appears it is virtually possible to synchronize the phase to any of the events, with a bit of focus and by starting the loop at different points. Yet some

events feel more comfortable to synchronize the phase with, so that one can organize the events according to the preference they receive as a beat.

Figure 2.1 represents the 19 loops on a time plot, and I have marked the preferred beat position by arrows. Arrows above the signal indicate the most comfortable beat positions, and arrows under the signal represent the most likely alternatives.

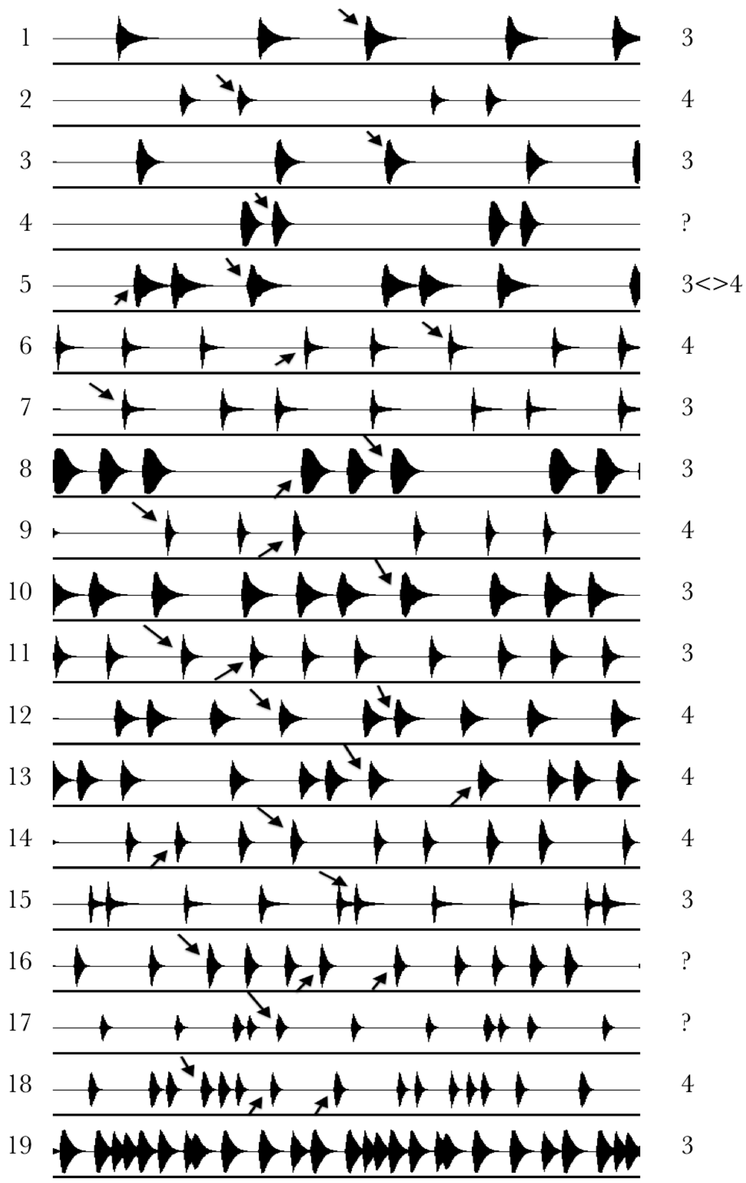

Fig. 2.1: 19 random event loops at 120 bpm

Loop n° 12 has two arrows above the signal because I simultaneously perceive both events as beats, with a double tempo. Loop n° 19 has no indication at all, because it proves very hard for me to identify a precise beat in such a dense succession of events.

The numbers on the right of the figure give an indication of how I feel the subdivision. I have found out that I only discern two kinds of feel: either 3 or 4 subdivisions. Many of these random sequences are ambiguous though, and if my preference one of both subdivisions, it is often easy to imagine it the other way. In some cases, as in n° 4, 16 and 17, I can't define the feel at all, yet the beat is still interesting to follow, even though it is more challenging.

I conclude that subdivisions of single beat loops are perceived as either duple (2 or 4), triple (3), or undefined, and that subdivisions by other numbers are never perceived. In either case the pulse is easy to follow and the result is aesthetically pleasing. The sense of pulse in these short loops is more stable than in the previous section's longer loops, because the beats are isochronous. To sound pleasing, sub-pulse events thus require a much less tight timing than pulse-level events. They can in fact be totally random, as long as the pulse is clearly defined.

I have slowed down n°16 and 17 4 times, trying to find out why they don't induce a clear feel *(Audio examples 35 and 36)*. When listening to these slowed down versions, I find them still confusing, yet the most comfortable loop division to me is 7, in both cases. I am sure it is pure coincidence that both come down to 7 and not 5 for example, but it indicates that none of both is close to actually being duple or triple. When listening to them as being divided by 7, I find I perceive them as sequences of 7 beats, meaning the change of tempo has promoted another temporal level as a pulse. If beats can be grouped by any number to form a period, sub-pulse divisions can only be felt as duple or triple. This dichotomy depends on the chosen pulse, which depends on the speed at which the sequence is presented.

What these experiments reveal to me, in absence of a clear theory from literature, is that the human brain attempts to perceive any recurring sequence of temporal events according to the sub-pulse / supra-pulse dichotomy exposed in previous sections, and does so following a certain method:

First the brain searches for the most viable pulse inside or close to the preferred tempo range. It does this by searching for series of more or less isochronously spaced events inside the sequence, picks one of them according to how isochronous it actually is, and thus how good a candidate it makes for synchronizing the pulse. In searching for the pulse, a certain hierarchy of preferred beat groupings applies, first 2 & 4, then 3 and 6, and finally 5, 7 etc.

Once the pulse is well-defined and entrainment takes place, the brain classifies sub-pulse divisions by a certain feel, i.e. duple, triple or more or less ambiguous, and whichever this feel is, microtiming variations below the pulse level will never be perceived as aesthetically displeasing, at the contrary of microtiming variations occurring above the pulse.

Timing in groove music thus has to be treated differently on sub- and supra-pulse levels : At the pulse level, the timing has to be tight in order to facilitate entrainment, and below the pulse, any timing will do. Therefore I now introduce the concept of sub-pulse Gestalts, stressing that sub-pulse groups of events are felt rather than counted, and classified as either duple or triple.

2.4. "Bonker Brains" by Hrdvision: an example of conflicting Gestalts

I now want to provide an example of music that plays with contrasting rhythmic Gestalts: *Audio example 37* is a piece of EDM which starts in strictly triple timing (measurements at the rhythmical surface show 0.33-0.33-0.33) and ends by contrasting this strictly triple Gestalt with another Gestalt close to quintuple division (0.3-0.3-0.2-0.2). The two Gestalts are represented schematically in figure 2.2.

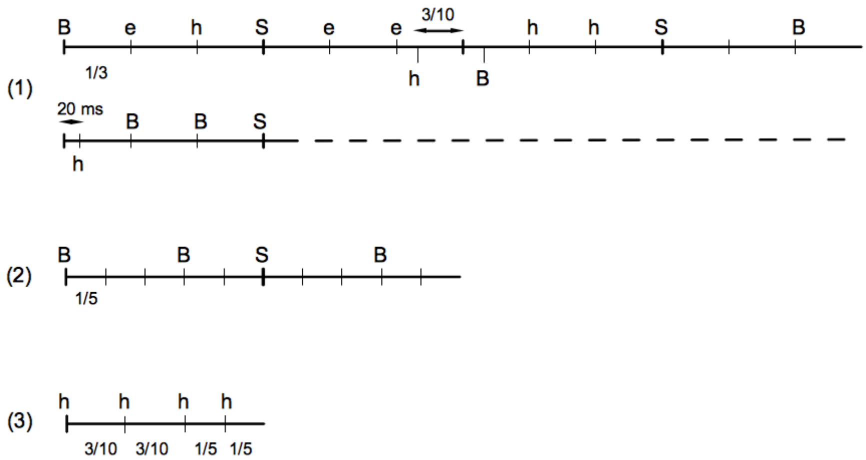

Fig. 2.2: Different Gestalts in "Bonker Brains" by Hrdvision. 'B' stands for bass drum, 'S' for snare, 'h' for hat and 'e' for any other sonic event.

(1) represents the time division of the period during the first part: 4 beats per period, and triple beat division, the only exception being that the hi-hat just before and the bass drum on the third beat are both delayed by 0.3 times the beat duration, which is a hint towards the further antagonisms to come.

(2) represents the bass drum and snare pattern in the middle part, from 0'30" in the sound example. Every second bass drum falls on 0.6 times the beat duration, in contrast with the triple division of the melody (0.67*beat duration). The hats are temporarily absent.

(3) from 0' 43" on, new hats, different from the ones in the first part, appear, in a 0.3-0.3-0.2-0.2 Gestalt, accentuating the antagonism with the melody that stays triple throughout the song.

In my opinion, it is this contrast between these two antagonistic Gestalts that makes the piece unique and fascinating. As the piece has been composed by electronic means, the timing patterns are strictly repeated from one period to another, which amplifies the antagonism even more, and the strictly triple and almost quintuple subdivisions are indeed precise to sample level.

When listening closely, I can perfectly feel the triple division, from beginning till end. Yet as the quintuple division slowly takes over the scene, no quintuple feel appears, the resulting pattern just feels more complex, more ambiguous. From the beginning till the end, the triple feel remains preponderant.

2.5. Choosing a downbeat

If we have found clear indices of how perceived timing is organized around the pulse and of what factors intervene in establishing the perceived tempo, we haven't delved into the factors that influence the choice of one event as a beginning of period or as the position of a beat, or put differently, how we organize events in groups.

In section 1.4, in our discussion of L&J's set of rules, we have seen that besides the metrical structure, L&J consider the organization of notes according to some grouping structure, so that besides its relation to the hierarchical meter, each note is also perceived as part of a melodic group or phrase. This mechanism can in fact be extended to any, not only pitched, sounds and many researchers try to find correlation between grouping and any parameter that defines sound, be it pitch, timbre, intensity, length, panning etc. In fact, grouping is essential to our segregation of sounds from the sonic landscape, as any sound has to be perceived as a separate entity before any identification and processing can occur. Grouping is thus the mechanism that allows us to perceive different sounds as separate in the first place. Different and contrasting series of several similar sound events tend to be perceived as separate auditory streams, meaning that we perceive them as groups of individual sounds, and that one can focus on one or another stream. Segregation of streams depends on similarity of some sounds in pitch, timbre and other sonic characteristics. For more information, see Bregman (1990).

Now to return to our grouping of randomly timed sequences of identical events, interestingly, literature (see Deutsch, 1999, for an overview) doesn't provide many explanations. If studies of grouping in function of pitch and melodic structure are countless, I could only find one study of grouping in function of timing in random sequences of identical sounds. Indeed, Povel & Okkerman (1981) have subjected listeners to recurring sequences of two identical sounds, separated by various short-long intervals, and found that listeners tend to place the accents differently in function of the proportion between the two intervals: they place the accent on the first sound of a group if the difference between the intervals is about 5 to 10%, and place the accent on the second sound of a group if the difference between the intervals is made bigger. This seems to confirm my accentuation in the first 4 1-beat loops.

Unfortunately, I couldn't find any similar study on groups of more than two sounds (there definitely seems to be a good opportunity here for new empirical research), but will try to make my own observations nevertheless.

When taking another look at fig. 8, a few clear tendencies become apparent:

(1) preferred beginnings are generally placed on the event preceding the longest interval of the loop.

(2) Second choices usually fall on events successive to the longest interval.

(3) When a certain number (>2) of events are in close succession to each other, the beginning falls either on the last of the first event of the group, never on events in the middle of the group.

(4) When two events are close to each other compared to all other durations, it is likely that the accent falls on the second of both.

Yet these tendencies seem to be in competition with each other. Loop n° 12 shows an apt illustration of such a dilemma, as the accent is falls either just before the longest duration of the group, or either just after the shortest duration, reinforcing the duple feeling in this loop. In loop n° 17, it seems like the longest interval overrules the shortest one in placing the beginning. I have no exact explanation of how one tendency can overrule another, but generally one could combine these 4 tendencies into one rule: the accent usually falls on an event preceded by a high density of events and followed by a low density of events.

Now that we have taken a few conclusions from the analysis of random sequences, it is time to return to performed music and investigate how performers make use of microtiming in achieving groove.

Chapter 3: Microtiming in groove music

3.1. Overview of existing models

There is no such thing as a clear theory of microtiming in groove music, and to my knowledge no one has ever provided any efficient generative model. Literature on the subject is scarce, and very often authors describe the groove in qualitative rather than quantitative terms, in what often seems to be pure speculation. There are a few interesting studies and theories, but their scope is usually limited to one particular aspect of one particular genre.

Regarding the production of groove by electronic means, commercial music software and hardware manufacturers have introduced 'groove templates' in the 80's, and these templates have been widely used in electronic music ever since. Groove templates provide an efficient model for adding microtiming to rhythmical sequences, but they are static as they implement fixed variation for each step in a sequence of *n* steps. For example, a sequence of 4 beats each divided into 4 density referents comes down to a sequence of 16 steps. A groove template is thus a series of 16 microtiming variations, one variation for each step, and the microtiming recurs with the sequence. The values are usually taken from measurements on recordings of live performances, so that the results sound realistic, but there is no place for playing with the template in the course of the music. The expressivity of groove templates thus limits itself to the periods, presenting the same microtiming pattern over and over again.

Some commercial systems are equipped with a shuffle control, which allows the user to swing the successive notes, and why not change the swing ratio over time. Yet this approach encourages the user to choose a fixed swing ratio throughout the whole, and the microtiming variations are then limited to subpulse timing, each pulse thus providing the same Gestalt.

In academic research, the KTH rules constitute the most convincing model I could find, yet its scope is restricted to classical music. For groove music, the only generative theory I have found is Waadeland (2000), who proposes to model microtiming by what he calls 'rhythmic frequency modulation', or RFM, a model that fails to convince me and which I therefore won't elaborate on in these pages. The road thus seems wide open for a new model.

In what I consider the most relevant account on the subject of microtiming in groove, Iyer (2002) summarizes other scholar's research and classifies microtiming as it occurs in performed grooves into 5 categories: asynchrony, streaming, spreading, backbeat delay and swing. His view has been very influential on my thinking, and I therefore want to keep his categorization while giving my own interpretation or them.

In the following pages, I will take Iyer's 5 categories of microtiming as it occurs in performed groove musics as a starting point for building my model, in relation with my own findings about random sequences.

3.2. Asynchrony

Asynchrony deals with slight spreading over time of a structurally synchronous event. Asynchrony thus escapes notation, as Fig. 3.1 makes clear.

Fig. 3.1: Structural representation of possible asynchrony (Iyer, 2002)

A performing ensemble can never achieve perfect synchrony, so that events intended to be played synchronously will always be separated by a small amount of time. Only a computer can achieve perfect synchrony.

Performers can also consciously make use of asynchrony to stress the different timbral constituents of the attack or exploring different temporal successions of these constituents.

I think the point here is that asynchrony occurs below the perceptual limit for meter, so that together they are perceived one event slightly stretched over time rather than two metrically distinct events. Iyer cites 30ms as a typical value for asynchrony. Relating to our previous discussion of sub-pulse Gestalts, asynchrony has to occur significantly below the 16th notes to be perceived as structurally synchronous. At 100 BPM, a 16th note equals 150ms, and a 32d note 75ms, already below the 100 ms threshold. Smaller and smaller time intervals induce more and more a feel of blending together and one can argue that a group of events smaller than 50ms will be perceived as just one structural event. As mentioned in section 1.6, it takes a minimum 2ms to discern to structurally synchronous events, and 20ms to discern their order of occurrence, therefore asynchrony can play with this sense of succession without interfering with the Gestalts. Furthermore, the audio range starts at 20Hz, so that any asynchrony below 5ms will induce a sense of pitch. If performing in this range of timing is virtually impossible for humans, it surely is a potential tool for electronic musicians who can rely on computers for precise timing.

Audio example 19 provides a good example of asynchrony. The loop length is 500ms, so that 16th notes equal 125ms. There occur two events per loop, separated by 60ms, significantly below the limit for meter and slightly below the interval of 32d notes at this tempo. The two events provide a clear sense of succession and of unfolding in time, but they don't give any hint about sub-pulse division. This loop is as easy to imagine in duple as in triple meter, or quintuple with a bit of focus, because the interval between the two events can not be perceived as meter, leaving the sub-pulse division undetermined.

To close the subject of asynchrony, and as Iyer states it : "in some musical situations in which blending is preferred, [...] multitimbral asynchrony may be undesirable, but it is often a valued musical trait in groove-based music" (Ibid., p. 400). For having performed many experiments with such small-scale microtiming, I can only confirm: in some cases, asynchrony can be very desirable, especially on and for inducing perceived beat locations, and in others perfect synchrony may be favored. To contrast different degrees of asynchrony is thus an important means of creating variation in groove.

3.3. Streaming

Bregman (1990) assesses that we tend to group successions of sonic events according to their similarity, perceiving them as separate streams. For instance, a sequence of two different sounds can be perceived as two separate auditory streams, and this perceptual phenomenon gets stronger with rising speed as well as difference pitch or timbre (by timbre, Bregman means any other sonic quality that can differentiate a sound from another). Similarly, the same can occur if a sequence of different sounds gives the opportunity to classify them according to similarity. For example, if a rhythmical sequence is composed of bass drum and bass guitar on one hand and of closed and open hats on the other hand, the closed and open hats will be united as one stream, and the bass drum and guitar will be perceived as another stream due to the different frequency bands one group and another occupy. Yet other factors such as brightness or attack time can play a role in streaming.

Iyer argues that timing can be used as a tool for generating streaming by contrasting some sounds to some others, typically by moving them forward and backwards in time (Fig. 3.2).

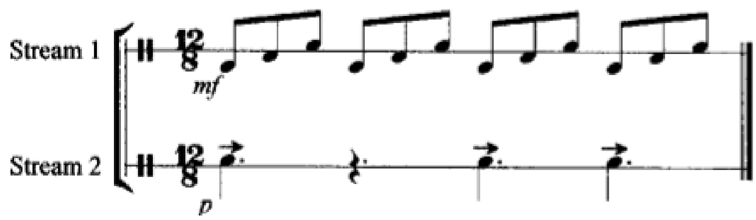

Fig. 3.2: Delay as a means of contrasting two streams (Iyer, 2002)

When considering a certain rhythm composed of various sounds, a certain selection of sounds can thus be brought into the attention of the listener by means of a shared time delay.

There seems to be no systematic rule in the order in which different parts of an ensemble are streamed. As Prögler reports, Ron Malo, one of Elvis Presley's producers has once explained “he liked splicing Ron Tutt's (Elvis Presley's drummer) parts because Ron always put the kick drum right on the beat, the snare a little behind the beat, and the toms and cymbals trailing after that. Conversely, cutting to John Guerin (a great jazz-rock player) was tough because John staggers his drums the other way: kick on the beat, snare a little bit ahead, and the hi-hat and some other drums leading that” (Prögler, 1995, p. 22). We can assume that the order is peculiar to the performer and/of the musical genre and that, in general, any stream can be streamed forward or backwards.

3.4. Spreading

Spreading is related to the idea of embodiment, or that microtiming feels the most comfortable to the listener when it sounds as if produced by a performer, which has to deal with his/her own body parts' inertia in hitting the drums. As Iyer puts it, “It was not until the advent of automated machinery that human ears were ever treated to inhuman rhythmic precision. Perhaps by virtue of their familiarity, the sonic traces of temporal constraints imposed by the body are in certain contexts perceived as aesthetically pleasing, while inhuman rhythmic response often is not” (Iyer, 2002, p.402).

Imagine one arm is supposed to perform a certain score, then the shorter the time is between 2 events, the harder it becomes for the performer to fight his own arm's inertia. S/he will thus tend to smooth his movements by making structurally short times a bit longer and structurally long times a bit shorter. This smoothing of successive intervals is called spreading, and one can an equivalent in the KTH double duration rule.

Spreading can also be seen as a tool for creating 'categorical ambiguity'. As we have seen previously that sub-pulse gestalts are categorized as either duple or triple, spreading is one way of interpolating from duple division to triple division. For instance, fig. 3.3 provides an example of interpolation from 16^{th} notes to 8^{th} triplets.

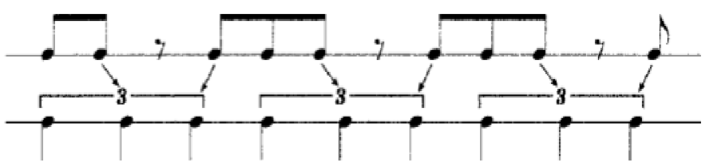

Fig. 3.3 : Interpolation between quadruple and triple subdivision by means of spreading (Iyer, 2002)

Audio example 38, “Kenu Dugu Fanga” by Kenestar provides a nice example of spreading in African music. The bell pattern in this track is constantly hesitating between duple and triple feel. Between 0'10” and 0'20” of the audio example, it follows the exact progression depicted in Fig. 3.3.

I am skeptical about the idea of embodiment. If it is obvious that inertia can play a role in the production of rhythm, and that spreading does occur in performed music, I assume that spreading is appreciated more for the categorical ambiguity it induces rather than for its relation with embodiment. The explanation of spreading due to the inertia of one limb is not convincing as drummers use several limbs in their playing, which allows them for more freedom, by means of asynchrony for example, which is harder to explain in relation to embodiment.

In either case, spreading plays an important role in making music more aesthetically pleasing. In a study on the subject, Cambouropoulos, Dixon, Goebl & Widmer (2001) have subjected listeners to differently smoothed versions of 5 musical excerpts, asking them to rate each version for its aesthetic qualities. They have found that smoothed versions gain better ratings, especially if intervals are smoothed according to the two (1 preceding and 1 following) adjacent intervals, rather than higher order smoothing functions (more than 1 preceding and 1 following intervals).

3.5. Swing

As we have seen in section 1.7, listeners tend to classify time durations in any sequence in two categories: long and short. And when confronted with different sequences made of two durations, they have a preference for a long-short ratio of about 1.75:1.

It is thus no surprise that swing plays a prominent role in groove music. In jazz, it usually occurs at 8th note level, and in jazz-rock, funk and many other genres it mostly occurs at 16th note level.

Swing can be seen as exerting two functions in perception of rhythm : 1) According to Iyer, it facilitates the perception of higher level rhythmic structure as 2 swung 8th notes are perceptually grouped into quarter notes, or beats. And 2) I believe swung 8th notes serve categorical ambiguity, as their ratios lie between duple and triple division.

Several studies have shown the swing ratio to depend on tempo, so that jazz performers tend to swing the 8th notes more at low tempo's. Fig. 3.4 provides the results of measurements Friberg & Sundström (2002) have performed on several jazz recordings of 4 famous drummers.

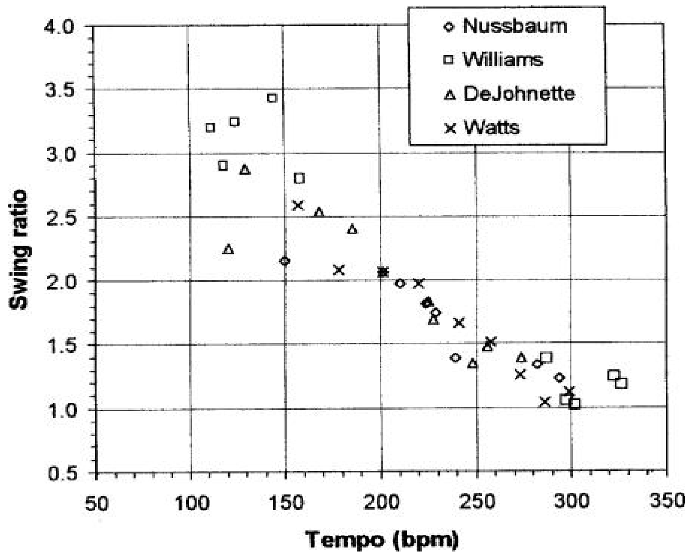

Fig. 3.4: Drummers' swing ratio as a function of tempo, measured on the ride cymbal (Friberg & Sundström, 2002)

The plot seems to indicate a swing ratio of more or less 3.5:1 at 100 bpm and of 1:1 at 300 bpm. Fig. 3.5 shows a plot of hypothetical swing ratios assuming a second note duration of 100ms:

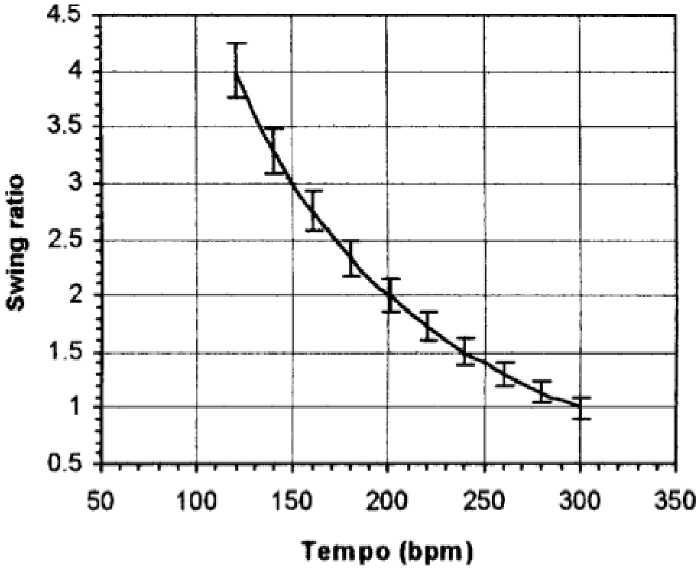

Fig. 3.5: Swing ratios assuming a constant second-note duration of 100ms (Friberg & Sundström, 2002)

The analogy is striking. Even though each performer has his own way of dealing with swing, the 100ms limit for meter perception seems to be respected on the average, as values below that limit would loose the sense of sub-pulse division. The 100ms threshold thus marks the edge between two perceptually time domains: above 100ms times are perceived in proportion to the pulse and its tempo, and below 100ms times seem to be perceived independently of tempo, just in a certain succession. The 100ms threshold thus draws a clear limit between asynchrony and swing. Values above it would be notated as notes occurring at different structural positions, and values below it would thus be notated as synchronous. Values above it participate to the sub-pulse duple, triple or ambiguous feel, and values below don't.

In performed music, the swing ratios evolve over time, according to what I will name 'swing phrasing'. In a paper devoted to jazz 8th notes, Benadon (2006) focuses on the subject, and puts the swing phrasing in relation to melodic phrasing. As melody is out of the scope of this text, I will not delve deeper into the subject, but i imagine the concept of swing phrasing can prove useful in a model for groove, be it centered on non-pitched sounds.

Fig. 3.6 shows one of Benadon's analyses. Measured time intervals between successive 8th notes are represented above the notated score, and successive swing ratios are represented below. The vertical lines that cross the score represent separation between melodic phrases, which are polyrhythmic, and are thus not to be confused with measures, as the time signature is indeed 4/4.

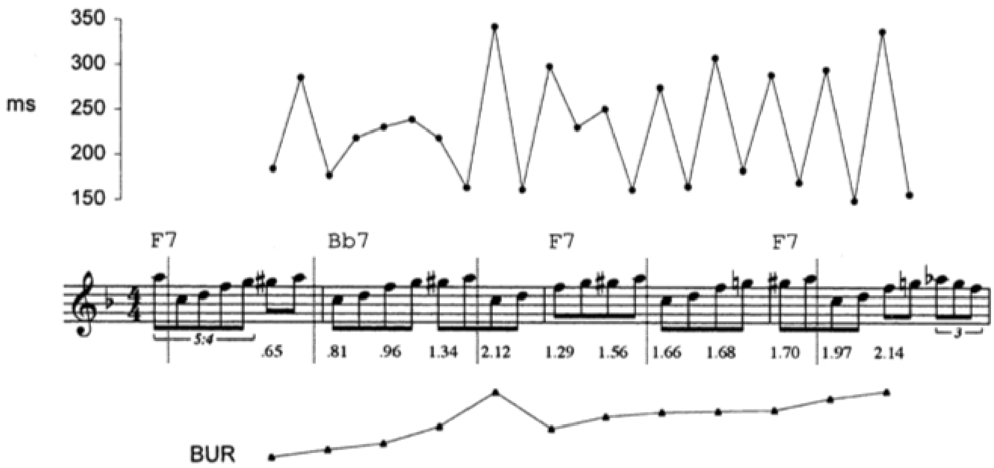

Fig. 3.6: swing phrasing in piano part of "Walkin" (132 bpm) by George Shearing (Benadon, 2006)

From the fifth value starting from left, the swing ratios (marked 'BUR', under the score) are greater than 1, which means that the offbeat durations are shorter than the beat durations, just according to Friberg & Sundström's average values. Yet the first four swing ratios are smaller than 1, meaning the swung 8^{th} notes are played before their structural position instead of after. If on the average, swing ratios are usually greater than 1 during a piece, the occasional use of inverted swing van be used to create contrast in the music, which is good to remember for our model.

There is no reason why performers in a same ensemble would phrase their swing synchronously. Friberg & Sundström give the following account of their findings in comparing different musician's swing ratios: “The interaction of the musicians in a jazz ensemble can be characterized in the following way: The drummer plays with a rather large swing ratio, with values up to 4 for slow tempi. The soloist has a smaller swing ratio, typically between 1 and 2. This difference is possible because the soloist is delayed at the downbeats and synchronized at the off-beats” (Friberg & Sundström, 2002, p. 344), a feature which may be assumed to be characteristic of jazz in particular. I would expect the soloist and the drummer's timing to coincide on the beat instead of the offbeat, guaranteeing a stable pulse. But according to this account the soloist is delayed, creating a tension with the drums by means of streaming.

The swing ratio also varies from one performer to another, so that each soloist can be assigned a certain histogram performed swing ratios peculiar to his/her style. Benadon (2006) has performed this task in his paper, providing such histograms, which for instance indicate that Bill Evans practices higher swing ratios (around 1.4:1) than Dexter Gordon (around 1:1).

To close this section on swing, I want to present the first out of 2 analyses I made of 2 tracks by Brandy, from her album “Full Moon”. Both tracks are produced by Rodney Jerkins. Rodney Jerkins is commonly considered one of the finest producers of electronic groove. His productions usually show very complex timing patterns. My assumption is that he performs most of his rhythms on the keyboard, recording MIDI data in his DAW, and then fine-tunes the timing to chisel his grooves. If it wasn't for Brandy, Rodney would figure among my favorite musicians.

It may seem paradoxical to provide an electronic groove as a sound example to a section further dedicated to swing in performed groove and mainly focused on jazz. Yet electronic grooves have the double advantage of being easy to analyze and to be identical at every period recursion. The two analyses are thus valid for all periods in the sound examples, providing a lot of clarity in transposing the graphical representations into sound and vice versa. Furthermore, they both provide good illustrations of swing.

The first example is taken from the track “I thought” (*Audio example 39*). Figure 3.7 provides a graphical representation of the 4-beat loop in this song.

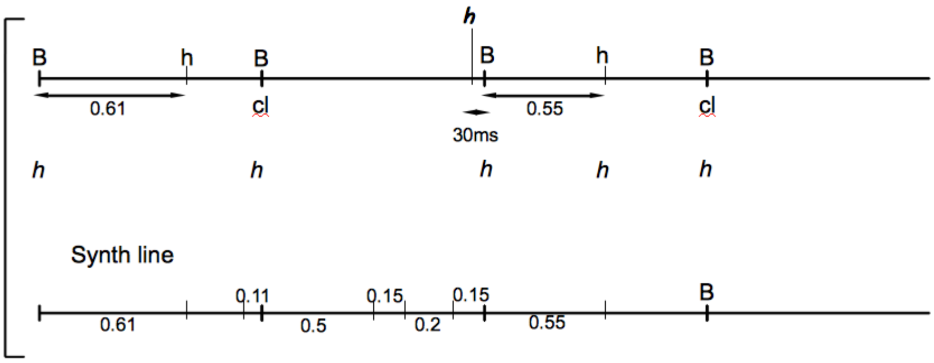

Fig. 3.7: The 4 beat groove in "I thought" by Brandy

The bass drums that coincide with the beats in this track are evenly spaced through the periods, accompanied either by a hat on beats 1 and 3 or a clap on beats 2 and 4. Two hats appear between the beats in the period, the first between beats 1 and 2, the second between beats 3 and 4. Note that the absence of hats between the other beats is a generally rare feature in groove music, as in most grooves all second 8th notes coincide with a hat. A third hat is placed 30ms before the 3d beat, focusing the energy on that beat and destabilizing the first, in what can obviously be described as an instance of asynchrony (I from now on take the habit of expressing values under 100ms as absolute times and values above 100ms as ratios of the beat duration).

Yet the most interesting feature consists of the different swing ratio that apply to the 2 first-mentioned hats: the first is placed at 0.61*the beat duration, and the second one is placed at 0.55. When listening closely, this variation in swing ratio does induce a rather triple feel in the first case, and a rather duple feel in the second case. This gives a good idea of the perceptual effect of swing phrasing.

The synth line is very irregular, but remarkably, structurally corresponding notes are synchronous to the two above-mentioned hats. Other beat subdivisions in the synth line can be interpreted as a combination of 8th and 16th note swing ratios, so that the synth line in all its structural subdivisions points towards the next beat.

3.6. Backbeat delay

As Iyer introduces the concept:

> "The notion of a backbeat is indigenous to the modern drum kit, an instrument pioneered by African Americans in this century. It consists of a strongly accented snare drum stroke or handclap on beats two and four of a four-beat metric cycle, where the beat is typically a moderate tactus rate. [...] The backbeat appears to have arisen in the middle of the 20th century, as the popular swing rhythm yielded to the even more popular, more bombastic rock and roll rhythms of artists such as Little Richard and Chuck Berry." (Iyer, 2002, p. 405)

The backbeat indeed is predominant in most today's popular music genres, including rock, funk, hiphop and EDM, in which it is often associated to the bass drum.

In performed groove, the backbeat is often placed slightly before or after its structural position compared to the bass drum. As Iyer formulates :

> "If we consider the downbeat to be exactly when the bass drum is struck, then the snare drum is very often played ever so slightly later than the midpoint between two consecutive pulses. [...] Often musicians are aware of this to some degree, and they have a term for it: the drummer is said to play 'in the pocket'. Although perhaps unaware of the exact temporal details of this effect, a skilled musician or listener in this genre hears this kind of expressive micro-delay as 'relaxed' or 'laid back' as opposed to 'stiff' or 'on top'."

As appears from this quote, groove musicians find the backbeat delay so important that they have a whole terminology for their different types. Stewart (1987, in Prögler, 1995) has attempted to quantify these different terms as precise timing variations, providing the "graphic display of feel in relative milliseconds at 130 bpm" represented in figure 3.8:

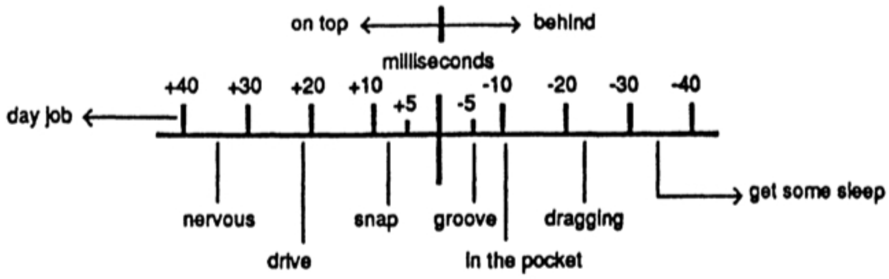

Fig. 3.8: Stewart's 'feel spectrum', in Prögler (1995)

The popular terminology is significant of how backbeat delays are perceived: early backbeats are considered "stiff", "nervous" or "on top", and late backbeats are labeled "behind", "dragging", or "laid-back". This representation clearly indicates that backbeat delays occur well below the 100ms, and more precisely in a range of +/- 40ms. Even though Stewart indicates a certain tempo for his graph, hence I assume that the feel these delays create is similar at all tempos.

Audio example 40, that comes with Butterfield's (2006) paper on microtiming, presents a basic rock rhythm in which different backbeat delays are contrasted with an otherwise strictly quantized rhythm, and comes in 4 segments separated by a pause. In the first segment, the backbeats come with no delay. In the second segment, they come 10ms behind ("in the pocket"), in the third, 20ms on top ("drive"), and in the third, 30ms on top ("nervous"). The examples give a good idea of the effect, event though other typical microtiming features are missing. For instance, one would expect a real drummer to stream the hats in the direction of the backbeat delay, as in Ron Malo's account, or to slightly swing the 16th notes, as is commonly done.

Iyer considers backbeat delays as a category of microtiming *an sich*, which applies specifically to snare drums structurally affiliated to beats with even numbers (2 and 4 in a 4-beat sequence). Yet such systematic delays can occur on other beat positions and to other instruments, such as in funk, where the first beat is systematically played earlier (Danielsen, 2006). I therefore want to reintroduce the notion of 'event shifts', and use it specifically for delays applied to events occurring on beat positions.

Now that I have introduced the notion of backbeat delay, we can take a look at my analysis of the Brandy's "Can we". *Audio example 41* & Fig. 3.9

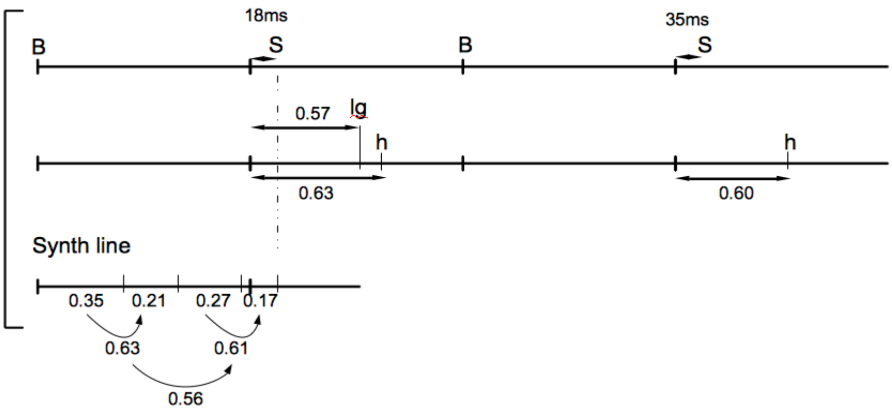

Fig. 3.9: 4-beat groove in Brandy's "Can we"

Again, "B" stands for bass drum, "S" for snare and "h" for hat, yet there is this typical sine sweep sound which I call 'lasergun', or "lg", in reference to its common use in science fiction movies. Even though I have only indicated it between the second and third beats, it actually appears between all beats, and thereby exerts the role of a hi-hat. It occurs with a fixed beat ratio of 0.57, between duple and triple feel as in the first brandy example. Yet a hat occurs every two beats, first with a beat ratio of 0.63, then 0.60, in either case consistently late compared to this lasergun, pushing toward triple feel where it occurs.

In this example, the backbeats are delayed by different values, appearing late compared to their presumed position and thereby pushing towards the next beat. The synth line, compared to the first beat position and the actual delayed backbeat, again shows a combination of 8^{th} and 16^{th} note swing ratios. Interestingly, they seem to be organized hierarchically: The interval between bass drum and snare is first divided according to beat ratio 0.56, and consequently each of the resulting intervals is further divided according to ratios 0.63 and 0.61, which are very similar values. It is interesting to think that this Gestalt could well be approximated by applying two successive swing ratios: first apply 0.57 on 8^{th} note level, then 0.62 on 16^{th} note level.

3.7. The downbeat in anticipation: Groove in James Brown's music

In a book dedicated to funk grooves, Danielsen (2006) points out to one essential element in funk music: "the One". The expression "the One", which emanates from James Brown himself, stands for the first beat of the pattern (or period; for information: funk music is mainly organized around the cyclical recurrence of 4-beat patterns).

According to Danielsen: "The One is on top of the basic unit; it is the first beat of the pattern and as such a focal point of the groove. James Brown's slogan "the rhythm on One" also points to the importance of handling this aspect of the funk groove the right way. As part of this, [...] the One should be played on top, or, in other words, as a *downbeat in anticipation*" (Danielsen, 2006, p.73).

Audio example 42, taken from the track "Hot Pants", which runs at 100 BPM, clearly shows this effect. Each pattern of 4 beats seems to be pushing towards the downbeat, and the downbeat always seems slightly in front of the pulse.

Yet, this effect is not simply achieved by playing the downbeat earlier than expected. In fact, all parts in the groove point forward towards the first beat. Fig. 3.10 shows the spectrogram of 1 period of 4 beats taken from the track.

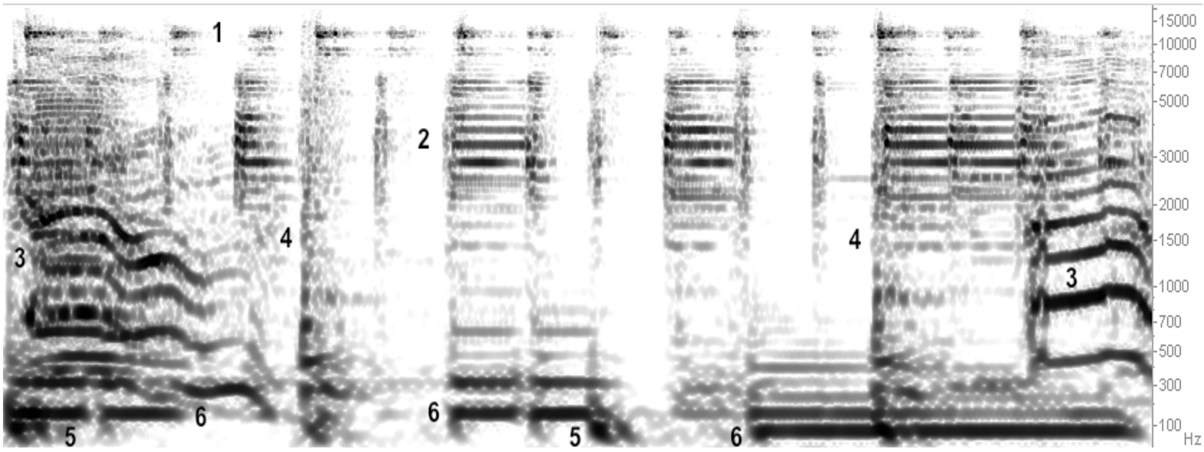

Fig. 3.10: Spectrogram of "Hot Pants" by James Brown, 2'04 to 2'08, 100 bpm

I have added the numbers to help identify different instruments.

The vertical placement of the 4 beat positions corresponds roughly with the position of the first (3), the first (4), the second (5) and the second (4).

(4) indicates the placement of the backbeat, and (5) the bass drum. The pattern of successive placements of bass drum and backbeat is roughly 630 – 610 – 605 – 595 ms, for an average value of 610ms. There are two different ways of interpreting these values:

- The tempo accelerates over the course of the period and resets on the downbeat
- Assuming a steady tempo, and taking the downbeat as reference, beats 2, 3 and 4 are late, so that beat 1 is in fact early compared to them.

(1) indicates the frequency band where the hi-hats are most easily identifiable. They play 16th notes with a swing ratio close to 1:1, with distinctly higher ratios in front of beats 2, 4 and 1, and are placed an average 20ms late compared to the other parts.

(2) indicates the guitar riff. It also plays a regular pattern of 16th notes, but the accents are placed according to a polyrhythmic figure, so that the accentuation pattern roughly sounds like 0..0..0..0..00.. , the 0's indicating stronger accents. This kind of figure is very common in funk grooves. It can be seen as a repeated syncopation or as a polyrhythm that resets every period. In fact, this syncopated figure of accents suggests a different meter (every 3 16th notes, or 125 bpm), and can thus be seen as an polyrhythm, creating metrical ambiguity inside the period. Yet, the polyrhythm is reset every 4 beats, thereby accenting the downbeat. Compared to African music, funk thus makes the step from strict polyrhythms to reset polyrhythms, loosing much of the initial metrical ambiguity but focusing all attention on the downbeat.

(3) indicates the placement of James Brown's shouts. On the first beat, the shout is a bit late compared to the bass drum (number 5, left of the number), or inversely the bass drum is in front of the shout. At the end of the period, James Brown's shout is a bit late compared to

the hi-hats. Given that the hi-hats are late themselves compared to the beats, the shout is thus strongly swung between beats 4 and 1.

(6) indicates the placement of the bass guitar. The first stroke is more or less synchronous with the first bass drum. The second stroke follows about one 16th later, yet more in synchrony with the hi-hat. It is thus late. This placing of the bass can thus be seen as an instance of spreading. The third stroke is a bit early compared to its structural placement on the 8th note between beats 2 and 3. The fourth stroke is synchronous with the hi-hats, so quite late compared to the beat. Again the bass seems spread. The fifth stroke is a bit late compared to its structurally associated 8th note, and the sixth, playing the same note, is a bit late compared to the snare drum on the 4th beat. Of all parts, the bass guitar shows the least consistent microtiming patterns, being once early and once late. Yet this seems to be a common phenomenon in groove, though, as reported by Prögler (1995) for example.

In analyzing this excerpt, we see that Iyer's categorization proves useful:

- The excerpts shows a clear instance of streaming, as the hi-hats are constantly late compared to other parts.

- The different parts show clear and systematic asynchrony, especially when it comes to the first beat. In fact, for having performed a few brief analyses of other excerpts of the same track, I can confirm the different parts usually appear in the same order: Bass drum, bass guitar, guitar, and hi-hats. Only the voice seems less stable in regard to this scheme. Furthermore, all parts are usually spread over a span of +/- 50ms.

- The alternate placing of the bass guitar before or after its structural position usually indicates spreading.

- The occasionally late placement of several parts can often, if not always be interpreted as an instance of swing with a ratio above 1. Only once on the bass line was the swing ratio smaller than 1.

- Backbeat delays were not significantly present. Instead, the downbeat was systematically performed early.

Danielsen's concept of downbeat in anticipation becomes more clear in analyzing the groove in “Hot Pants”. Not only the downbeat is performed early, but the whole rhythm appears to push toward it: James Brown's shouts show the highest swing ratio just before the downbeat, the guitar riff's polyrhythmic accentuation resets on the downbeat, the hi-hats are streamed, pushing the music forward, and the overall density of events is the highest just before the downbeat. This recalls our conclusion about the placing of the beginning of random loops. In fact, it is the combination of all these elements that gives the music its forward drive towards the downbeat.

3.8. Anacrusis and metrical ambiguity

In a 2006 paper, Butterfield insists on what he calls anacrusis in achieving groove. In a classical sense, anacrusis is a short melodic fragment that precedes the beginning of a section or a piece. Anacrusis prepares the listener for what is to come. Butterfield sees anacrusis in a much broader sense, as any element that anticipates a certain event to come. He doesn't provide a clear definition of the qualities that make some events or structures anacrustic, but in light of our analysis of "Hot Pants", we can argue that most of Iyer's categories, as well as structural tools such as the resetting of a polyrhythm or the increase of density just before a beat, or any change in intensity, timbre or duration occurring before a beat, essentially serve anacrusis towards that beat.

If anacrustic tools can be used to point towards one beat in particular, reinforcing the perception of that beat as a downbeat, as is the case in James Brown's music, it can also be used to point at several beats in a period, creating tension between possible meters by suggesting a different phase or a different tempo for the pulse. Anacrusis is thus a tool to induce metrical ambiguity and "highlight contrasting structural elements", in Chernoff's words.

The creation of metrical ambiguity is thus the most important feature of a groove, and anacrusis stands for a number of tools that help achieving it. I distinguish three kinds of ambiguity:

- *Ambiguity of beginning* (Butler, 2006) is the kind of ambiguity we have experienced in our random sequences as several events could be interpreted as the beginning of a period, depending on anacrustic hints. Ambiguity of beginning occurs either between two events that are perceived simultaneously as beats, with anacrusis pointing either towards the one or the other, which to me sounds very pleasing. Or it can occur between one event that is perceived as a beat and another event that is currently not perceived as a beat, with the anacrusis (typically a swing ratio below 1) pointing towards the latter. I perceive this type of ambiguity of beginning as much more disturbing than the first type. *Audio example 43* is a nice example of ambiguity of beginning. When the beat disappears, we mentally place the beginning one 16^{th} note later, which creates confusion at the beat's return.

- *Ambiguity of metrical type* (Ibid.) regards the relation between beat and period, suggesting either division of a same period into different possible beat durations, or the grouping of beats into period of different lengths. Polyrhythms can suggest both. Performers can point towards one or another metre by means of anacrusis, suggesting one or another event as a beat.

- *Categorical ambiguity* regards the perception of sub-pulse Gestalts, as we have seen in relation to spreading. In this case anacrusis doesn't occur as sub-pulse divisions are concerned, and not beats. Yet an anacrusis can point towards an event between two beats, breaking the Gestalt, and suggesting a change of phase in the pulse, which is an instance of ambiguity of beginning and can better be avoided.

In light of all previous discussion, I want to reintroduce the notion of 'event shifts', which I consider small (<100ms) microtiming variations occurring on beat positions.

It once for all appears to me that Iyer's categories can be classified in two main groups in relation to the perceptual limit for meter:

1) Event shifts are small (<50ms) variations that occur around beat positions. They do not influence the sense of sub-pulse division and their role is to give more energy to a certain beat. Additionally they can provide clarity by spreading structurally equal events in time around the beat location. Early event shifts tend to accentuate the beat, whereas late event shifts tend to put the beat on the background. Early event shifts facilitate the perception of a certain beat as a downbeat, whereas late event shifts tend to work as anacrusis towards the next beat. Asynchrony, streaming and backbeat delays are to be regarded as event shifts. I choose to limit event shifts to the range of +/- 50 ms as values above this value tend to belong more to the rhythmic Gestalts.

2) Rhythmic Gestalts instill a sense of duple of triple subdivision between beat positions. Their timing doesn't need to be as tight as the timing of event shifts. Their timing pattern can take any shape (above or under 100ms) as long as it points towards the existing pulse, reinforcing it. The major mean of doing this is to concentrate event density before the beat rather than after it. Swing and spreading mostly contribute to Rhythmic Gestalts. Swing ratios above 1 have an anacrustic function towards the next beat, and both swing and spreading contribute to creating a duple or triple feel, or categorical ambiguity. I use the term rhythmic Gestalt both for the configuration of events between beats at the rhythmical surface and for the perceived feel.

3.9. "Mach Rider" by Cupp Cave, pushing the groove to its limits

Now that the microtiming according to its anacrustic function and Iyer's 5 categories have been defined, it is time to take a closer look at a few more sound examples.

The first example I want to present is "Mach Rider" by Cupp Cave (*Audio example 44*), from the album "Garbage Pail Beats". Completely broken rhythm is a characteristic feature of this album. If microtiming is essential in building grooves, Cupp Cave's electronic grooves push the microtiming to the edge where it almost destroys the groove. This album completely fascinated me, and this is in fact the trigger that made me realize the power of microtiming and decide to make it my central research subject.

The results of my analysis are represented in figure 3.11:

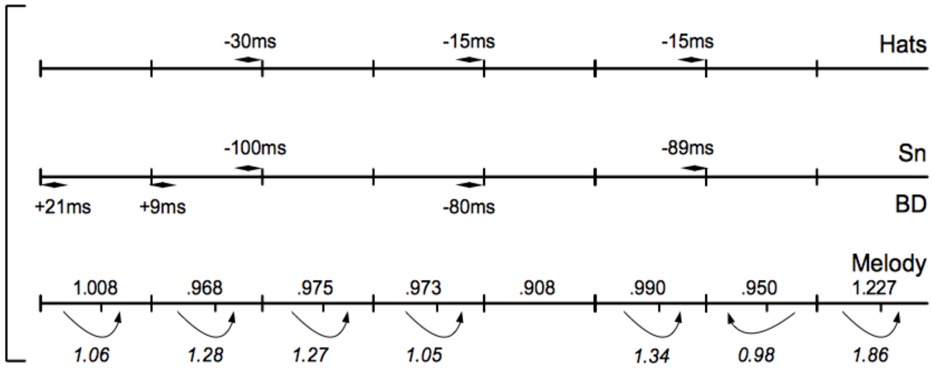

Fig. 3.11: The 4-beat loop in "Mach Rider"

The groove presents itself as a classic hip-hop groove based on the interaction between bass drum and backbeats, and suggests duple beat division, as is usual to the genre. The average tempo is 80 bpm. The hats are placed on all eight notes, and besides three event shifts, are isochronous. I therefore choose them as a reference point for the measurements.

Many elements are timed far from their structural position: The snare comes about 90-100ms early and the bass drum occurs sometimes before, sometimes after. The first backbeat appears 100ms early compared to the hats, and is perceived as a backbeat delay: event though the delay is much higher than the usual +/- 40ms interval, the snare drum is still structurally perceived on the second beat. Yet, in fact, a delay of 100ms (or 121ms compared to the bass drum) comes close to the 16th note value at this tempo (187ms), so that one would expect the snare drum to be perceived as a swung 16th note before the beat. In fact, this demonstrates the importance of enculturation in identifying rhythms: the classic hip-hop break is so familiar to us that we try to fit this groove in it, beyond evidence for a different, yet unfamiliar, time structure. As a result, events with different structural affiliations (in this case, the early snare drum and the hat on the first beat) can be perceived as synchronous. Carlsen & Witek (2010) have provided more examples of this phenomenon. Cupp Cave thus challenges us by pushing well-known timing patterns to the edge of recognition, playing with our tendency to seek for familiarity. My immediate experience has showed that even people accustomed to avant-garde music could become totally disoriented by Cupp Cave's music. This phenomenon is reminiscent of Pressing's amusia, which we was introduced in section 1.6.

The melody is perceived as a series of swung 16th notes with all swing ratios greater than 1 (see values under the timeline), but for one exception. Cupp Cave sticks to conventions in this respect, every second 16th note pushes towards the next 8th note, and the swing ratios are phrased. Yet, the 8th note intervals themselves are irregular (ratios to average value above the timeline), running out of phase compared to the hats. In fact, the longest 8th note value is the last one, the upbeat. Cupp Cave thus does exactly the opposite of James Brown, playing the downbeat late. The music seems to run more or less towards the downbeat, but once it's gotten to the upbeat, the combination of a longer 8th note duration in the melody on one hand, and of early backbeat - late bass drum combination on the other hand, gives the impression that time stands still for a while, missing the One as it is expected, which sounds very disturbing. Whether the listener appreciates this unusual effect or not, the composer surely did this intentionally, challenging our expectations about how a good groove should be.

3.10. Expressive timing in African music

Expressive timing in performed music is harder to analyze than in electronic grooves, as every period shows a different timing, mostly due to phrasing, yet often due to bad synchronization and motoric noise. Therefore, to perform a complete analysis of a performed piece of music would require that all periods be analyzed and compared with each other, so that general trends can appear out of statistical data. I haven't performed any analysis in such a systematic way, but I have looked at a few recordings of music from African I enjoy very much, trying to extract a few tendencies in microtiming.

The first piece of this selection is a recording of Omar El Knitri, downloaded from Awesome Tapes from Africa, and whose title and release name remain unknown (*Audio example 45*). It is a typical, and particularly good piece of Chaabi, which impresses me very much for the energy it emanates. When analyzing it, I find that the tempo is very steady on the short term, and presents a progressive accelerando on the long term. The beat grouping is quadruple, and the beat division is triple.

Fig. 3.12 represents the period's structural organization:

```
HoOoOOOoOooo    hats
..O..O..O..O    claps
...O......O.    high tom
O.....O..o..    low tom

|..|..|..|..    pulse
```

Fig. 3.12: Rhythmical structure in Omar El Knitri

Dots stand for silence, o's for soft strokes, O's for loud strokes, and |'s for the pulse. The H stands for a figure of three hats spread over the density referent.

Beats 1 and 3 are clearly marked by what I call a low tom (as a matter of simplification, descriptions relate to components of the Western drum kit). Beat 2 is marked by an accented hat and the high tom, yet beat 4 is soft and the high tom is displaced one density referent (or one triplet eight note, depending on the point of view) late. This late syncopation sounds confusing as it suggests a possible beat position without being one. The meter is challenged by the claps, which are all placed one density referent early, suggesting a different pulse placed on the claps, to which the second high tom could serve as an anacrusis. Even though this groove is ambiguous in regard to beginning, the pulse is never really challenged, especially due to the voice's anacrustic role.

The timing on beat positions is the tightest I have ever seen, between +/- 10ms, and one can assume that tight timing is an appreciated feature in Chaabi music. Following this assumption, one can conclude that the few timing variations present in this piece are unintended, suggesting 10ms as a possible range for motoric noise. Beat subdivisions are also very tight, tending to be strictly triple, with measured event positions between beat ratios 0.30 and 0.33, and 0.60 and 0.67. As the tempo in this excerpt runs around 170 bpm, this also comes down to variations below 10ms, confirming the motoric noise hypothesis.

The second piece of African music I want to present is “Baya Ba Koni” by the New Star Orchestra, a Yoruba chant from Nigeria, taken from the compilation-CD “Yoruba Street Percussion” (*Audio example 46*).

The timing in this piece seems less tight as in the previous example. The beat division in this piece is duple, the beats are grouped by 4, and the pulse position is absolutely unambiguous. Events on beats 1 and 3 are early, promoting them as more important than beats 2 and 4, and many structural anacrustic figures point towards beat 1. The shakers play two 16^{th} notes before each beat, pointing towards it, and are streamed between 10 and 25ms early. Two groups of two salient drum strokes are played around beats 2 and 4. The ones preceding the beat are placed in ratio 0.72...0.76, and the ones succeeding the beat in ratio 0.21...0.22 of the beat interval. They thus come a bit early compared to presumed 16^{th} note positions, pulling the music backwards, yet the first beat of each period is clearly played early. The low drums are streamed 15 to 30ms late. Tension is thus created between early and late elements, some of which drive the groove forward, while others pull it backward. Iyer (2002) reports a similar construct in a piece by Ahmad Jamal, in which an inverted (0.85) swing ratio is put in balance with many anacrustic elements, hereby challenging the pulse. Furthermore, Gerisher (2006) has pointed to the importance of inverted swing ratios in Brasilian percussion.

Last in our selection is “Giriama Spirit Dance”, from Kenya, as it appears on the compilation-CD “East Africa Witchcraft & Ritual Music” (*Audio example 47*). This piece is a typical example of polyrhythmic music, and Fig 3.13 shows my interpretation of its structure.

```
............  bell
|.....|.....  period
|.|.|.|.|.|.  pulse
.o..o..o..o.  high drum
o.....o.....  middle drum
..oo.o..oo.o  low drum
```

Fig. 3.13: Cross-rhythms in "Giriama Spirit Dance"

The bell lays a clear density referent: the piece is completely perceived as made of 8th notes, so that the feeling is duple. It strongly stands out of the rhythm, and is swung with ratios varying irregularly between 1.3:1 and 1.4:1. This swing clearly stresses the placement and the tempo of the pulse. The drums perform a figure of 3 beats, marking the period length, and the middle drum the period beginning, and the low drums perform an anacrustic figure pointing towards it, with a swing ratio around 1.3:1. Yet the high drums stand in conflict with the pulse, suggesting duple instead of triple beat grouping. But this possible duple meter is completely overruled by the swing in the bells, so that the first meter stays firmly. Returning to Chernoff's polymeter model, this makes me believe that polyrhythm constitute a better model. If structurally, polymeters can seem to exist, expressive timing in performances always point towards one periodicity as the main pulse.

Chapter 4: My instrument

4.1. The importance of the referent

I discovered the power of what I later got to call polyrhythms as a teenager. I found that, starting from a single density referent, one can build up complex rhythmical structures by combining periods of different lengths. *Audio example 48* is a fragment of a piece I composed back at the time, based on the combination of different streams each consisting of one sound and a fixed interval, each sound being assigned a different interval (3,4,5,6,7,9,11,13 etc.). Note that the perceived pulse groups the referents by 4, instilling a duple feeling, according to our hypothesis from chapter 2. In building such music, it is important that all periods be a multiple of the same density referent in order to maintain a sense of pulse. If the ratios between the period lengths are not integer, the result would sound like phasing, a technique extensively explored by Steve Reich, and which I am not interested in.

Now if I am to consider my concept of sub-pulse Gestalt, which can include any number of events in any configuration between two beats, the results would be very successful when combining loops with identical tempo, but for having tried, I can state that such combinations are not as successful in a polyrhythmic fashion. In fact, and as is the case in African music, the sense of pulse should be preserved at any price, and microtiming should point towards one unique pulse.

The only solution for combining Gestalts in a polyrhythmic fashion is thus to return to the concept of density referent, around which to organize the expressive timing.

I have made some experiments with first-order random functions in generating rhythmical structure. *Audio example 49* is taken from a piece of mine in which each stream is generated by means of a Travesty generator applied to a certain array of duration-timbre pairs. Suppose events occurring in one stream can only consist of 5 possible percussive sounds, which I name timbres (t1, t2, t3, t4 & t5) and durations that are multiples of a certain density referent d, then a possible rhythmical sequence would look like [[t1,2*d], [t3,4*d], [t2,2*d], etc.]. To apply a Travesty generator to such arrays is exactly what I did in generating the rhythms for this piece (I have made [preceding duration, timbre] pairs, but of course one could make use of [following duration, timbre] pairs or higher order states).

A Travesty generator is similar to a Markov chain in that the probability for a certain event to occur depends on one or more preceding events. It proves very powerful in generating music in general and rhythm in particular as it gives a certain structural framework to the randomness, getting closer to how human performers improvise. Yet to function properly, both need a finite number of events, which is why I restricted myself to 5 timbres and a few multiples of a density referent.

My point is that, to make first- and higher order random processes possible, one has to limit oneself to a finite (and small) number of durations, and that the use of a common denominator seems to be the most obvious solution for preserving the synchronicity between different streams. Streams of durations without a common denominator would fall out of phase, just like periods of a polyrhythm would.

4.2. Gestalt interpolation

So in order to allow for non-zero order random processes and for polyrhythms, my model has to be based on a density referent, and in order to allow for a wide variety of sub-pulse Gestalts, large event shifts should be possible around each structural position. They should be larger than the density referent itself and smaller than the space between two beats. This inevitably raises the obligation to establish a fixed relationship between beat and referent, and therefore I opt for 4 referents per beat, coming back to the usual 16th notes. I thus organize my rhythms with 4 events per beat, according to a fixed referent. If this excludes the possibility for more than 4 events between two beats, it still allows for less than 4, as events could also be silent. Each event can be shifted, and any Gestalt we have seen throughout these pages can be performed this way, as long as it can be represented by maximum 4 events. For example, the synth-line's swing in Brandy's "Can We" could be modeled by a [0.35, 0.21, 0.27, 0.17] Gestalt (fig. 3.9, section 3.6. Ratios given in proportion of the beat interval), or put differently, as an array of event shifts [0, +0.10, +0.06, +0.08] from a density referent of 0.25 (the first element corresponds to the event structurally affiliated to the beat). Similarly, Hrdvision's hi-hat part from section 2.4 would constitute the Gestalt [0.3, 0.3, 0.2, 0.2], and triple beat subdivision could be obtained through a Gestalts like [0.33, 0.33, 0.17, 0.17] or [0.33, 0.33, 0.22, 0.11]. In this case one could decide to silence the last event for strict triplicity, or keep it audible for extra anacrusis towards the next beat. Gestalts such as [0.17, 0.17, 0.33, 0.33] should better be avoided though, or handled with care at least, as they may blur the sense of pulse and indicate a new beginning.

Audio example 50 is a recording of a live performance I did with my instrument, where different streams of different period length (polyrhythmic) are combined, and where each stream's sub-pulse timing is separately interpolated between Gestalts [0.25, 0.25, 0.25, 0.25], [0.3, 0.3, 0.2, 0.2] and [0.33, 0.17, 0.17, 0.33]. The piece successfully plays with the ambiguity between duple and triple feel, with a convincing transition occurring around 2'00".

Yet around 5'50” the pulse is lost and takes another beginning, an effect I want to avoid. This is caused by the choice of Gestalt [0.33, 0.17, 0.17, 0.33], which points its anacrusis towards the third event of the Gestalt.

4.3. Gestalts as swing ratio arrays

The model I will propose in this section yet has not been implemented yet, but at the term of my research for writing the current text, I came to conclude that the best way to organize Gestalts is to rely on swing.

As we have seen in section 3.5, 8th notes with swing ratios above 1 point towards the next beat, and 16th notes with swing ratios above 1 point towards the next 8th note. Swing ratios above 1 always stabilize the pulse, whereas swing ratios below 1 challenge it. Therefore, swing ratios below 1 should be used with parsimony, and the swing phrasing should include a majority of swing ratios greater than 1, just as it appears in performed jazz (remember Fig. 3.6, section 3.5) and in Cupp Cave's music (Fig. 3.11, section 3.9).

In fact, any Gestalt of 4 events can be defined by the following 3 parameters: 8th note swing ratio, average 16th note swing ratio, and 16th note swing ratio balance. The first defines the position of the third event compared to the first, which is synchronous with the beat. The second then places events 2 and 4 in a fixed ratio between events 1 and 3, and the third performs fine-tuning.

To demonstrate this by means of a practical example, I take the random Gestalt : [0.42, 0.19, 0.13, 0.26], and convert it into these three parameters:

(0.42+0.19)/(13+26) = **1.56**	is the 8th note swing ratio
0.42/0.19 = 2.21 0.13/0.26 = 0.5	are the two 16th note swing ratios
√(2.21*0.5) = **1.05**	is the average 16th note swing ratio
0.5/1.05 = 1.05/2.21 = **0.48**	is the 16th note swing ratio balance

Our Gestalt can now be defined unambiguously as **[1.56, 1.05, 0.48]**. Yet, as we treat ratios here, it is logical to use a logarithmic form. After all, swing ratio 0.5 is perceived as the contrary of 2, and 1 is considered neutral. If we take the log10 value of our swing ratio array, swing ratios between 0.1 and 10 convert into values between -1 and 1, and the neutral swing ratio 1 converts into 0. The choice of base 10 thus seems an adequate choice, as swing ratios outside the range 0.1...10 result in small (<100ms) values at most tempos, and the use of logarithms provides the right scale for further interpolations (it was not used in section 4.2).

Log10([1.56, 1.05, 0.48]) = **[0.19, 0.02, -0.32]**

is thus the final form of our Gestalts.

Here follows a series of other examples obtained by applying the same calculations:

[.25, .25, .25, .25]	–>	[0, 0, 0]	(neutral swing)
[.3, .3, .2, .2]	–>	[.18, 0, 0]	(Hrdvision hats)
[.33, .33, .22, .11]	–>	[.3, .15, .15]	(close to triple Gestalt)
[.5, .17, .17, .17]	–>	[.29, .23, -.24]	(close to sextuplet division)
[.35, .21, .27, .17]	–>	[.10, .21, -.01]	(Brandy synth-line)

We can also compare a few permutations of a same array, showing the clear correlation between both representations:

[.17, .17, .33, .33]	–>	[-.3, 0, 0]
[.33, .33, .17, .17]	–>	[.3, 0, 0]
[.33, .17, .17, .33]	–>	[0, 0, -.3]
[.17, .33, .33, .17]	–>	[0, 0, .3]
[.17, .33, .17, .33]	–>	[0, -.3, 0]
[.33, .17, .33, .17]	–>	[0, .3, 0]

Any Gestalt should thus be converted into what I will call a 'swing ratio array', and interpolations between Gestalts should be done at this level for best sounding results. Positive elements in swing ratio arrays always reinforce anacrusis towards the current pulse, whereas negative elements challenge it.

I consider it to be a good start to phrase the three swing ratios per 4 beats and to make the swing phrasing period common to all streams, since microtiming should always point towards one unique pulse, even for polyrhythms, as we have seen in the “Giriama Spirit Dance”. I will thus experiment with different phrasing arrays and take conclusions. I expect that all streams better share a certain common (global) swing phrasing pattern, from which local deviations would occur, according to local subordinate swing phrasing patterns. Global controls would thus define global patterns for swing phrasing between which local controls would allow to interpolate. Hence streams with the same interpolation setting would be perfectly synchronous to each other, and would change accordingly when their global superiors change. Each streams phrasing would thus place itself somewhere between a few (2 or 3, not more) global presets, and the presets would be modified at global level. Yet this is fiction, the swing phrasing model hasn't been implemented yet and I will for sure learn a lot when trying it out.

4.4. Model for event shifts

Now that we have defined a clear model for sub-pulse Gestalts, we can take a closer look at event shifts, where 'event shifts' stands for shifts occurring around beat positions specifically, in contrast with the account in section 1.8.

If the Gestalts provide a wide variety of microtiming patterns between beats, they keep events occurring at beat positions perfectly synchronous. We thus need a model for backbeat delays and the One, for asynchrony and streaming, deviations that all occur on beat positions. As we have seen in section 3.5, these value should be treated as absolute times instead of in proportion to tempo. Event shifts can thus best be modeled by patterns of delays between -30 and +30 ms, that recur every period. In most grooves we encountered, the period length is usually 4 beats. Therefore, I will adopt 4 beat periods as a fundament for all event shifts, and won't consider any other pattern lengths, suspending them as future possible developments.

In first approximation, we can consider the possible patterns qualitatively, representing them as arrays of 4 binary elements, in which -1 stands for early and 0 for synchronous. I take the decision to neglect late occurrences, as they do not seem to play a role as essential as early ones.

In fact, the patterns that seem the most relevant to me at the term of this research can be ordered so that from one pattern to another, only one element changes:

`[ 0,  0, -1,  0]`	the third beat is early, suggesting it as alternative beginning
`[-1,  0, -1,  0]`	the first and third beats are early, comforting the the first higher metrical level
`[-1,  0,  0,  0]`	the One, common in funk
`[ 0,  0,  0,  0]`	neutral configuration
`[ 0, -1,  0,  0]`	the first backbeat is earlier than the second, as is common in hip-hop
`[ 0, -1,  0, -1]`	both backbeats equally early
`[ 0,  0,  0, -1]`	the second backbeat earlier than the first, also common in hip-hop

I assume that interpolation between such arrays can prove very efficient, as all patterns are then linked to one control. If one combines these patterns with a delay, they transform into patterns of synchronous and late event shifts, and a control for the amplitude of the deviations would make the model complete.

For use in a live instrument, it is important to reduce the number of controls. Therefore, in first approximation, I make all streams share the same event shift pattern, reducing the interpolation to one global control, and combine amplitude and delay into one parameter per stream, on local level. This leaves the performer sufficient space to achieve all kinds of asynchrony and streaming, while significantly reducing the number of controls.

I still have to try this configuration out, but for instance the patterns could be scaled according to a function like *y = 40*tanh((0.5x)^3)*, as represented in Fig. 4.1:

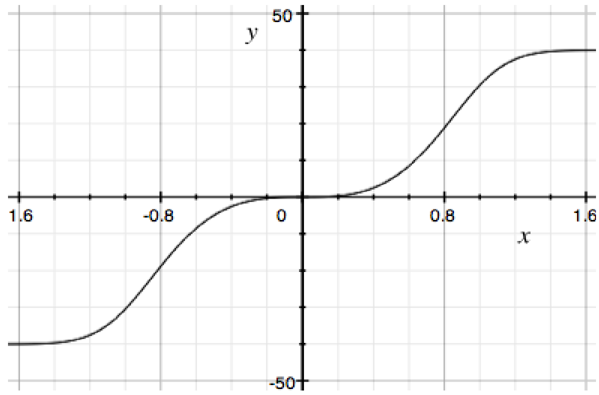

Fig. 4.1: Proposed scaling function for local control of event shifts

If above-mentioned patterns, with values between -1 and 0, are added various offsets in the range -1.5...+2.5, different event shift configurations take place. At offset -1.5, all events are streamed 40ms early, independently from the event shift pattern. At offset +0.5, all events are more or less synchronous with the beat, and at offset +2.5, they are streamed 40ms late. Offsets between these values provide different early-synchronous and synchronous-late expressions of the event shift patterns. For instance, at offset -0.3, the contrast between early and late events in the pattern is maximal, ranging from about 35ms early to quasi-synchronous.

4.5. Melodies of timbre

The principle of my instrument is to combine different streams with each other, by exerting control on the stream's characteristics. If this thesis' main focus is on expressive timing, it is time to reveal how the instrument deals with other aspects of rhythm, namely structural and timbral aspects. By structural aspects I refer to the rhythmical structure as discussed until now, and by timbral aspects I mean all the sound qualities that make the identity of an event, whether intensity, envelope, spectral contents, degree of harmonicity, etc. As it is very difficult to establish a clear correlation between synthesis parameters and perceived sound qualities, I have decided to synthesize all sounds with the same synthesis engine, so that every sound is unambiguously determined by a certain combinations of values in the same array of synthesis parameters, applying to the same synthesizer. This way, I can treat my sounds as decoupled from the synthesis parameters that originally defined them, and organize them on a two-dimensional field, according to perceived timbral similarities. This makes fluid transitions possible from one sound to another, so that a rhythm can potentially be made of an infinite rather than a finite assortment of sounds. Without delving into the precise synthesis techniques I use in producing my sounds, the different sounds that result from certain combinations of synthesis parameters can be given a name after acoustic drums they evoke, and organized in a two-dimensional timbral field (more-dimensional fields are not excluded, but two seems a good compromise between flexibility and simplicity). Figure 4.2 gives an illustration of such a timbral field:

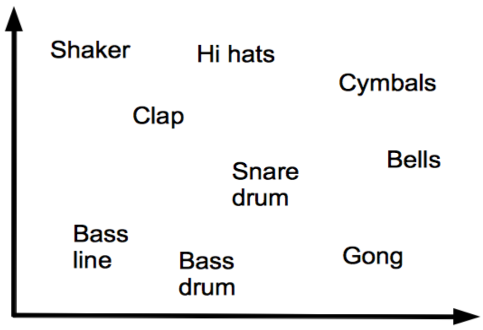

Fig. 4.2: A timbral field

A bass drum is essentially characterized by a brief attack and a low-frequency mode, a snare drum is essentially a higher pitched bass drum plus some band-filtered noise, and hi-hats can be modeled as band-filtered noise at a higher frequency. Therefore, possible interpolation between these classes of sounds is evident. A shaker is essentially different from a hi-hat in that its attack time is longer, as is the difference between a bass drum and what I call a bass line (a bass drum with a longer attack time). Cymbals, bells and gongs occupy similar frequency bands as hi hats, snare drums and bass drums, are all percussive, and what differentiates them from the latter is that they are composed of more modes, and contain less noise. Finally, a clap comes close to a hi-hat with a slightly lower center frequency and a more complex attack. We see that, according to this classification, the y-axis correlates more or less with pitch, and the x-axis with changes in envelope and harmonic content. The important point here is that, if all combinations of synthesis parameters can not be fit into a timbral field, it is nevertheless possible to make a wide selection of different and good sounding sounds, and place them in the field, so that in the end, timbre boils down to just two parameters.

I am currently working on new synthesis techniques that allow a wider variety of sounds, yet timbre has been organized in a x-y field in both *Audio examples 49 and 50*. In both pieces, different combinations of modes are associated with x-y coordinates, and the obvious timbral transitions in piece AUDIO EXAMPLE were indeed obtained by adding certain offsets (directly linked to a MIDI control) to these x-y coordinates, thereby interpolating between different timbres (the interpolation curves are all sine segments). Both axes are wrapped around their limits, so that one can endlessly interpolate in the same direction, by virtually any offset values.

All my events' timbres are thus unambiguously determined by the following parameters: timbre-x and timbre-y. For each stream and each new density referent, the sequencer picks these 2 parameter values from predetermined arrays. I have until now always used arrays of 64 random values (linear distribution) for each parameter. Each stream is assigned different arrays for each parameter, and the arrays can be generated on demand each time the performer want to add a new stream to the rhythm. 64 values equal 16 beats, the maximal period length, yet each stream has a control for period length, allowing to read any smaller number of values and thus to combine polyrhythms.

In *Audio example 49*, arrays of timbre values were still drawn by hand, so that the performer had to determine the rhythmical structure before igniting a stream. Yet, in *Audio example 50* all the streams are based on randomly generated arrays, and the results sound convincing enough to keep this way of working. If the performer doesn't have control anymore on which stream will be introduced next, random arrays offer a wider variety of possibilities and I like the idea that each new stream is a surprise the performer has to deal with, by means of the controls. Randomness thus forces the performer to be creative, and guarantees that each performance be unique.

I plan to replace the x and y translations around fixed positions by polar coordinates. Each sound in the array would be associated with random radius and angle values, forming a timbral Gestalt, which three controls would scale, rotate and translate respectively : 1) scaling: the radius would allow to make all sounds in a stream sound more or less similar, as de facto at radius=0 they would be identical. 2) rotation: a control for the angle would allow to change all timbres, and exchange their qualities (for example, when rotating around the snare drum, a cymbal would become a bell while a bass drum would become a bass line). 3) translation would be a combination of translation on both x and y axes, according to a diagonal function, such as *y = (5x).wrap(1)*, providing many possible centers for the gestalt before returning to its starting point, thanks to the wrapping functions on both axes.

4.6. Dealing with intensity

A good rhythm shouldn't be based on the sole variation of timbre, and therefore intensity (or amplitude) can't be left aside. As we have seen, performed rhythms often show accentuation patterns on certain instruments, and moreover, it rarely occurs that a certain instrument plays on each density referent, and the alternation of silent and sounding events is the very core of the rhythmical structure.

Yet, practically, an event that doesn't sound can be considered silent, so that the accentuation and onset patterns can be combined in one unique structure of intensity. Intensity, whether it comes down to bare amplitude, or amplitude combined with other sonic qualities, constitutes an essential parameter next to timbre, and I have decided to articulate it in three different ways: 1) 'appearance probability', 2) 'appearance category' and 3) 'beat-level accentuation'. I will develop each of these concepts separately in the present chapter.

1) Appearance probability is, as its name suggests, the probability for a certain event to appear. Practically I generate random arrays of 64 such probabilities in the same fashion as the timbre arrays, and the sequencer extracts a probability for each step. The power of this technique is that it brings variation from one period to another, by making a different selection of events silent at each recursion. This allows for certain figures to appear and disappear in the rhythm. If most events have low probabilities, I make sure that a certain number of events appear all the time, giving the groove a clear backbone around which the variations occur. This feature has not been implemented yet.

2) In *Audio example 49* I was drawing sequences of events in a binary fashion as is common in analog sequencers (1's for sounding events, 0's for silent events), I lately have decided to generate random patterns to determine the structure just as is now the case for timbre. If the appearance probability is one tool for determining a stream's structure, I also came to a musically interesting concept: the 'appearance categories'. Each event is assigned a random appearance category value between 0 and 1, and what determines if it is sounding or not is whether this random value falls between two thresholds or not, the two thresholds being directly linked to two global controls. For example, the category array [.45, .75, .29, .12] translates into pattern [1,0,0,0] for thresholds .3 and .5, pattern [1,0,0,1] for thresholds .2 and .5, and [0,1,1,0] for thresholds .5 and .2. Even though the category arrays are generated randomly at first, the performer thus has a strong structurally expressive power by means of just two controls, making the structure as dense or sparse as he wants by means of a range control, and modifying the structure by means of an offset control. For instance, a range of 0 means all events are silent, a range of 1 means all events are sounding, and supposing a range of 0.5, an offset change of 0.5 would invert the pattern (the values are wrapped between 0 and 1).

I have implemented a smoothing function around the thresholds, so that events don't appear and disappear brutally. Slow changes in the threshold controls rather make the sounds slowly fade in and out.

After experimenting with separate controls for each stream, I have finally decided to make both controls globals, as this approach can provide the most drastic changes in the resulting rhythm. In an instrument that mostly relies on slow fine-tuning of parameters, the threshold range and offset controls are the most efficient for instant expressivity.

3) Both the appearance probability and appearance categories are linked to the period length, which can be different for each period. Yet, due to the importance of accentuation patterns in comforting the pulse, intensity should also be articulated inside the beat. One local control, the beat-level accentuation, would allow to stress one referent or another, much in the same fashion as the event-shift patterns. A typical interpolation sequence would be

[0,1,0,1] – [0,0,0,1] – [1,0,0,1] – [1,0,0,0] – [1,0,1,0]

allowing for separate accentuation of 8th note anacruses, beat anacrusis, beat and beat anacrusis, beat and 8th note events respectively.

All three parameters: appearance probability, appearance category and beat-level accentuation would combine (by dB addition) to the resulting accentuation pattern, and events below a -80dB threshold would be discarded to safe CPU power, as is already the case. What is clear is that these implementations will drastically increase the number of controls, making live performance more difficult. I will thus implement them, appreciate their efficiency and make choices accordingly, placing them as global or local or simply discarding them, in order to restrict the number of controls.

4.7. Final remarks

In the previous sections, I have proposed to phrase different parameters according to different clocks: beat-level accentuation would depend on beat-length patterns, swing phrasing and event shifts on 4-beat patterns, and timbre, appearance category and appearance probability on period lengths, different for each stream. I thus consider to organize my music according to three different clocks. Different choices could be made, but that will depend on experience when implementing the complete model.

An important further development for my instrument would be to reintroduce the first order random functions I left aside between composing *Audio examples 49 and 50*. A simplified version of these functions would consist of a small Markov chains linking different event positions inside a pattern. Each position could be given a certain probability of switching to certain other position rather than to switch to the next. For example, a pattern of four elements [1,2,3,4] could be associated with the pattern of probabilities [.2, .6, .1, .7] for transition to positions [3,1,4,2]. This example equals the following transition matrix (input states above, output states on the left):

	1	2	3	4
1	.8	.6	0	0
2	0	.4	0	.7
3	.2	0	.9	0
4	0	0	.1	.3

Fig. 4.3: Transition matrix for simplified Markov chain

Such a simplified model excludes many possible transitions, and the larger the matrix (in reality, the period length would be the size), the more zeroes appear. Yet it does allow for a fair amount of variation, and its implementation is very compact.

A control could define the amount of randomness by multiplying the pattern of probabilities by a number between 0 and 1, allowing for more or less variations. In practice, this stochastic process would apply to all period-length parameters simultaneously, so that at any point in time, timbre, appearance category and appearance probability would always be linked together. One last remark on the use of first order random functions: Even though I am satisfied with my piece *Audio example 49*, I assume that a higher degree of musicality could be obtained by resetting the random process back to the first state at every new period, providing a more stability to the groove.

In the end, I have seen my instrument evolve from more deterministic to more and more probabilistic, relying on linear random distributions for most parameters. The results sure sound convincing, but at some point of its development, the need to return to more determination may appear. For instance, besides polyrhythms, there is a wide variety of common structural patterns that have not been discussed in these pages and can easily be produced by means of algorithms. Euclidian rhythms, and more generally additive rhythms, stay at the back of my head. For more information on these, please refer to Toussaint (2005), Arom (2004) and Kvifte (2007). Structural elements in African-American music, as exposed by Wilson (1974), including call-and-response patterns and anacrustic figures, also open new doors to model improvements. For example, a propensity to anacrustic figures could be obtained by introducing phrasing on the thresholds of the appearance categories, and call-and-response patterns suggest that a same array be read simultaneously by different streams, with an offset. All kinds of possible improvements can be imagined, but for now it is time to return to the code.

Conclusion

This thesis recounts the genesis of my instrument for building grooves live from algorithms.

Starting from the assumption that a good groove can not be made without deep knowledge about how rhythm is perceived, I have provided an overview of relevant literature from perceptual psychology research on the subject, complementing the account with my own experiments and remarks when necessary, and focusing on one musical tool that appears most crucial in creating groove: expressive timing.

In chapter one I have attempted to define groove, defending a certain conception of meter in the end, as I found what fundamental role the choice of a certain paradigm for meter plays in any discussion of groove.

In chapter two I have given an account of my own perceptual experiments with random sequences, attempting to fill a gap that researchers have left open.

In chapter three I have presented Iyer's five categories as the most elaborate account of microtiming in groove, adapting them according to findings from my own research.

Chapter four is dedicated to my instrument, presenting my approach to rhythm as a combination of timbral, structural, microtemporal aspects and the use of phrasing patterns of different lengths. Some ideas have been implemented already in my instrument, whereas others are yet to be. Therefore I have adopted a both more personal and conditional style.

Groove is an important feature of most music we listen to today, and has mostly been looked over by academic circles. This thesis is my attempt to fill this breach, between psychology, musicology and composition.

References

Anku, W. (2000). Circles and time: A theory of structural organization of rhythm in african music. *Music Theory Online, 6*(1). Retrieved from http://www.mtosmt.org/issues/mto.00.6.1/mto.00.6.1.anku_essay.html

Arom, S. (1989). Time and structure in the music of Central Africa: Periodicity, meter, rhythm and polyrhythmics. *Leonardo Music Journal, 22*(1), 91-99.

Arom, S. (1991). *African polyphony and polyrhythm.* Martin Thom, Barbara Tuckett, and Raymon Boyd, trans. Cambridge: Cambridge University Press.

Arom, S. (2004). L'aksak: Principes et typologie. *Cahiers d'ethnomusicologie, 17,* 11-48. Retrieved from http://ethnomusicologie.revues.org/413

Benadon, F. (2006). Jazz eight-notes as expressive microrhythm. *Ethnomusicology, 50*(1), 73-98.

Bilmes, J. (1992). A model for musical rhythm. In *International Computer Music Conference 1992.*

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Butler, M. J. (2006). *Unlocking the groove: Rhythm, meter, and musical design in electronic dance music.* Bloomington: Indiana University Press.

Butterfield, M. W. (2006). The power of anacrusis: Engendered feeling in groove-based musics. *Music Theory Online, 12*(4). Retrieved from http://www.mtosmt.org/issues/mto.06.12.4/mto.06.12.4.butterfield.html

Brower, C. (1993). Memory and the perception of rhythm. *Music Theory Spectrum, 15*(1), 19-35.

Cambouropoulos, E. (1997). Musical rhythm: A formal model for determining local boundaries, accents and metre in a musical surface. In M. Leman (Ed.), *Music, Gestalt and computing - Studies in systematic and cognitive musicology* (pp. 277-293). Berlin: Springer Verlag. Retrieved from http://users.auth.gr/emilios/englishpage/pub.html

Cambouropoulos, E., Dixon, S., Goebl, W., Widmer, G. (2001). *Human preferences for tempo smoothness*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.2.5522

Carlsen, K. & Witek, M. A. G. (2010). Simultaneous rhythmic events with different schematic affiliations: Microtiming and dynamic attending in two contemporary R&B grooves. In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 51-68). Farnham: Ashgate Publishing Limited.

Chernoff, J. M. (1979). *African rhythm and african sensibility: aesthetics and social action in african* *musical idioms.* Chicago: University of Chicago Press.

Danielsen, A. (2006). *Presence and pleasure: The funk grooves of James Brown and Parliament.* Middletown: Wesleyan University Press.

Danielsen, A. (2010a). Here, there, and everywhere: Three accounts of pulse in d'Angelo's 'left and right'. In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 19-34). Farnham: Ashgate Publishing Limited.

Danielsen, A. (2010b). Introduction : Rhythm in the age of digital reproduction. In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 1-16). Farnham: Ashgate Publishing Limited.

Demers, J. (2010). *Listening to the noise: The aesthetics of experimental electronic music.* New York: Oxford University Press.

Deutsch, D. (1999). Grouping mechanisms in music. In Deutsch (Ed.), *The psychology of music (second edition)*. New York: Academic press.

Falk, R. & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review, 104*(2), 301-318.

Fink, R. (2005). *Repeating ourselves: American minimal music as cultural practice.* Berkeley: University of California Press.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149-80). New York: Academic Press. Online version without page numbering retrieved from http://www.zainea.com/rhythmandtempo.htm

Friberg, A., Bresin, R. & Sundberg, J. (2006). Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology, 2*(2-3), 145-161.

Friberg, A. & Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: evidence for a common rhythmic pattern. *Music Perception: An Interdisciplinary Journal, 19*(3), 333-349.

Gerisher, C. (2006). O suingue Baiano: Rhythmic feeling and microrhythmic phenomena in Brazilian percussion. *Ethnomusicology, 50*(1), 99-119.

Hasty, C. (1997). *Meter as rhythm.* New York: Oxford University Press.

Honing, H. (2002). Structure and interpretation of rhythm and timing. *Tijdschrift voor Muziektheorie, 7*(3), 227-232. Retrieved from http://www.mcg.uva.nl/index.html?abstracts/mmm-TvM.html&target

Hughes, T. (2003). Groove and flow: Six analytical essays on the music of Stevie Wonder (Doctoral dissertation, University of Washington). Retrieved from http://www.steviewonder.org.uk/bio/life-stories/groove&flow/T.Hughes%20-%20Groove%20And%20Flow%20(S.Wonder).pdf

Huron, D. (2007). Sweet anticipation: *Music and the psychology of expectation* (paperback edition). Cambridge, MA: MIT Press.

Iyer, V. (2002). Embodied mind, situated cognition, and expressive microtiming in African-American music. *Music Perception: An Interdisciplinary Journal, 19*(3), 397-414.

Iyer, V. (1998). Microstructures of feel, macrostructures of sound: Embodies cognition in West African and African-American musics (Doctoral dissertation, University of California). Retrieved from http://archive.cnmat.berkeley.edu/People/Vijay/%20THESIS.html

Jones, M. R. (1981). Only time can tell: On the topology of mental space and time. *Critical Inquiry 7,* 557-76.

Kramer, J. D. (1981). New temporalities in music. *Critical Inquiry, 7*(3), 539-556.

Kvifte, T. (2007). Categories and timing: One the perception of meter. *Ethnomusicology, 51*(1), 64-84.

Keil, C. (2010). Defining 'groove'. Retrieved from http://www2.hu-berlin.de/fpm/popscrip/themen/pst11/pst11_keil02.html

Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In S. Grondin (Ed.), *Psychology of time.* Bingley, Emerald Group.

Large, E. W. & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science, 26,* 1-37.

Lerdahl, F. & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

London, J. (2002). Cognitive constraints on metric systems: Some observations and hypotheses. *Music Perception, 19*(4), 539-550.

London, J. (2012). *Hearing in time* (2d ed.). New York: Oxford University Press.

Madison, G. (2006). Experiencing groove induced by music: Consistency and phenomenology. *Music Perception, 24*(2), 201-208.

McClary, S. (2004). Rap, minimalism, and structures of time in late twentieth-century culture. In C. Cox & D. Warner (Eds.), *Audio Culture: Readings in modern music* (pp. 289-297). New York: The Continuum International Publishing Group Inc.

McHard, J. (2001). *The future of modern music: a philosophical exploration of modernist music in the 20th century and beyond.* Livonia: Iconic Press.

Patel, A. D. (2007). *Music, language, and the brain.* New York: Oxford University Press.

Prögler; J. A. (1995). Searching for swing: Participatory discrepancies in the jazz rhythm section. *Ethnomusicology, 39*(1), 21-54.

Risset, J.-C. (2010). La fabrique du son musical: Calculer et composer le son. In A. Soulez & H. Vaggione (Eds.), *Manières de faire des sons* (pp. 17-35). Paris: L'Harmattan.

Schaeffer, P. (1966). *Traité des objets musicaux.* Paris: Seuil.

Sherburne, P. (2004). Digital discipline: Minimalism in house and techno. In C. Cox & D. Warner (Eds.), *Audio Culture: Readings in modern music* (pp. 319-326). New York: The Continuum International Publishing Group Inc.

Temperley, D. (2000). Meter and grouping in African music: A view from music theory. In *Ethnomusicology, 44*(1), 65-96.

Toussaint, G. (2005). The Euclidian algorithm generates traditional musical rhythms. In *Proceedings of BRIDGES: mathematical connections in art, music and science.* Banff, Alberta, Canada.

Waadeland, C. H. (2000). Rhythmic movements and moveable rhythms: Syntheses of expressive timing by means of rhythmic frequency modulation (Doctoral dissertation, Norwegian University of Science and Technology). Retrieved from http://ntnu.diva-portal.org/smash/get/diva2:123059/FULLTEXT01

Wilson, O. (1974). The significance of the relationship between Afro-American music and West-African music. *The Black Perspective in Music, 2,* 3–22.

The sound examples can be provided on written demand to my e-mail address x.de.wannemaeker@gmail.com. Furthermore, any questions or remarks on this text and its subject would be highly appreciated.