8.5 Timing and rhythm

It's not only (apparently) world-weary John Rowles who lamented "If I only had time" – this is also *the* cri de coeur of every bad drummer. To stay in time is difficult, and even just to define time is not an easy feat. Usually, one refers to Augustine (a monk, not a drummer); his contemplations about time are widespread, and they are readily abbreviated to: "when nobody asks me what time it is, I do know it; as I seek to explain it to an asker, I do not." It is something like that when it comes to rhythm, groove, and timing. Even modern books on rhythmics♣ make do entirely without defining rhythm – and, yes, it is a challenge. If indeed someone tries, it reads something like this: *rhythm is the regular (periodic) repetition of accents that are pooled together.*

A pattern arises out of grouped accents – with an accent being a distinctive feature, i.e. for example the beats on a bass drum. Already now we can think of examples where this does not fit … in any case: if everyone can criticize this definition because he/she knows anyway what rhythm is, then an extensive definition is indeed superfluous. At least it is in the present context where the focus is on auditory perception, and not on teaching rhythmics. **Fig. 8.29** gives a brief outline on the hierarchical processing of continuous time, its discretization into basic beats, and the latter's grouping and accentuation. Based on this ordering scheme is the individual pattern that repeats within one bar in this example. In a two-bar pattern, two different patterns would alternate (Bossa) – but, again, the emphasis here is on the hearing system and not on the music.

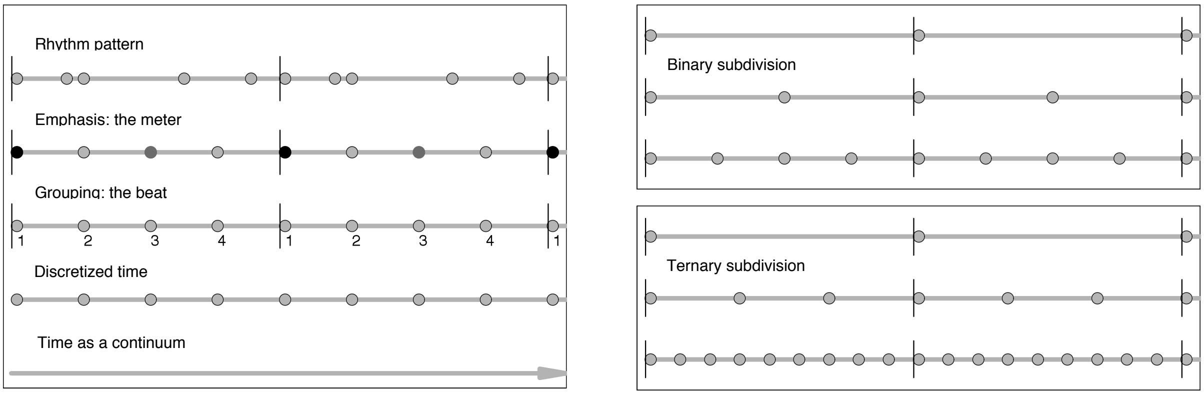

Fig. 8.29: Beat, meter, rhythm (left), binary and ternary subdivision of the beat (right).

In Fig. 8.29, the dots mark the start of individual notes – and the layperson believes that a musician with good "timing" needs to reproduce these starting points as precisely as possibly in order to receive the "playing like a machine"-distinction. Checking whether this is actually true shall be postponed to the next page; first, the focus is on the analysis: what is the accuracy that the hearing system can muster to analyze fluctuations in rhythm? Practical recording-studio experience provides the barely contested orientation value of **10 – 20 ms**. Timing errors of less than that quickly become meaningless. So, after all: the pro should hit his/her notes with a precision of about 1/100th of a second, and hard- and software in music needs to react very quickly in order not to make the ever-present signal-processing delays subjectively noticeable. Also: it must not be overlooked that an effects device with a basic delay of 7 ms is uncritical but 4 of them in series are not tolerable. For this reason some processors offer two settings: little time-lag (low latency) for playing live, and more time-lag (high latency) – but also much effect – for off-line processing.

♣ Marron E.: Die Rhythmik-Lehre. AMA 1991.

But now on to the actual topic of this chapter: **depending on circumstances, a good musician may need to do an objectively inaccurate performance so that it sounds correct subjectively**. The listener does the subjective assessment; the analyzer delivers the objective data. A good musician will not generally play his notes at the time as written but with a slight offset that is in part deliberate (determined) and in part unintentional (stochastic). Both (!) offsets are desirable and as such generic. This is why the sentence in bold above is not good as an excuse for the beginner – objectively wrong playing may indeed simply sound very wrong. But then: what is would be "correctly wrong"? Not in the sense of an aggravation or emphasis of "wrong", but rather: which deviation from the objective click of the metronome leads to a rhythm subjectively perceived as "good"?

Let’s look at the determined (deliberate) deviations first. We have three main aspects here: the tone generation (transient, onset, attack), the tone perception (the auditory event), and the interpretation. Regarding the tone generation: 20 ms may have passed by the end of the tone onset (attack phase) of a wind instrument – or as much as 100 ms for low-volume-notes. Of course, it is not the moment when the player’s lips open that counts as the onset of the tone, but a somewhat later point in time that can only be defined via the perception. Therefore the wind-instrument player needs to the start blowing before the tone is supposed to sound. When exactly the played note is considered to be existent – that is a decision made by the hearing system i.e. it is an act of the **tone perception**. In his book on psychoacoustics [12], Fastl lists eight examples for tones with different time-envelopes (TE), and determines the corresponding subjective start of the note. Only if a note has an abrupt onset of tone and immediately decays again (decaying TE) do objective and subjective starting points practically coincide. Given an increasing TE, the subjective start of the tone is up to 60 ms later than the objective one. For these special sounds, such numbers are of course dependent on the special experiment. However, even if all notes have a steep attack, we find astonishing differences in terms of seemingly equally long pauses: for the hearing system to assess a tone-duration as equally long as a duration of a pause, the objective duration of the tone needs to be considerably shorter relative to the objective duration of the pause! In Fastl’s example, first an allegretto eighth-note and then an eighth-pause (both of 240 ms duration) are to be played. In order for tone and pause to sound equally long subjectively, the tone must not be played for 240 ms but for a shorter 100 ms, while the pause needs to be lengthened to 380 ms! For quarter-note and quarter-pause (nominal length 480 ms), the ratio is not quite as dramatic: tone duration = 260 ms, pause duration = 700 ms. The explanation of these indeed considerable discrepancies is found in the auditory transient processes (attack and decay) that extend the subjective length of a note (relative to the objective length) and thus shorten the subjective length of a pause. On top of these discrepancies (caused by the processing), the individual interpretation of the musician also needs to be considered. For example, given special stylistics, the “one” (1st quarter) will deliberately be played a bit early, or the “three” might arrive a tad later that it nominally should. This is not for a lack of exercise but to demonstrate individual virtuosity. If that weren’t the case, all those 1000s of “Elises” celebrated on the pianoforte would have to sound identical.

It is in fact exactly this deviation from absolute rules that marks the virtuoso, the person “in the know” who understands not only where to deviate from the strict formula, but also how much. This knowledge often exists only implicitly i.e. without explicit awareness of it. If a virtuoso is asked to play the same passages, we will recognize always (almost) the same deviation. It is an expression of the personal style and not random at all. However, if we ask at which points he or she has shifted the “one”, the artist will have difficulties supplying a complete list, and if we inquire about the degree of shifts, the answer is likely to be: *just as I feel it, I don’t check the clock.*

In their work on swing-rhythmic, Lindsay/Nordquist♣ investigate an example regarding maintaining rhythm: **"Fever"** by Ray Charles (2004). The piece is dominated by a finger-snapping that puts the emphasis on the elsewhere often less accentuated even-numbered quarter notes (backbeat, more or less: iambus rather than trochee). These snaps are incredibly precise, as shown in **Fig. 8.30**: here, the envelope over the time reference is depicted (894 ms, corresponding to 67 bpm), and reference is hit with merely ±2 ms deviation. There is no further info but we can only surmise that some kind of assistance system was involved – it is hardly imaginable that a freely playing musician can after 3 minutes still be within 2 ms of the original time. That "Fever" – despite this machine-like precision – still never gets boring is to the credit of the continuously changing pattern played by other instruments. The bass basically plays half-notes, but already the first 4 bars reveal some of the bass notes locking into a ternary grid (splitting the fourths into thirds). The congas, as well, cannot do without the ternary splitting in their "da-dub-da". If merely accents on the four basic beats were allowed, the whole charm of the piece would be gone; it would be life-less and without that "swing".

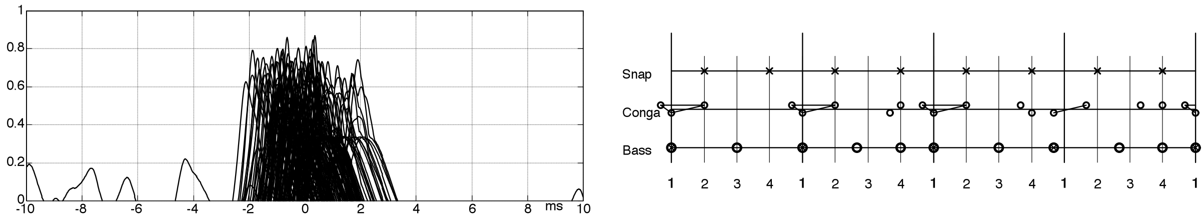

Fig. 8.30: R. Charles / N. Cole: *Fever*. Left: timing-analysis; right: four bars at the beginning of the piece.

"Hit the Road Jack" is another piece of Ray Charles' that provides aid to answering the question how precise the pro keeps time. In this version, Ray C. plays the first 1,5 minutes without accompaniment, and presumably also without click-track. He cranks up the tempo from 91 bpm to more than 95 bpm (beats per minute) – in a dance contest with rigid tempo-specifications (often as little as ±2%) that would be quite borderline. But hey, this is Brother Ray, and it ain't no dance contest, either: that piece needs to be exactly how he plays it.

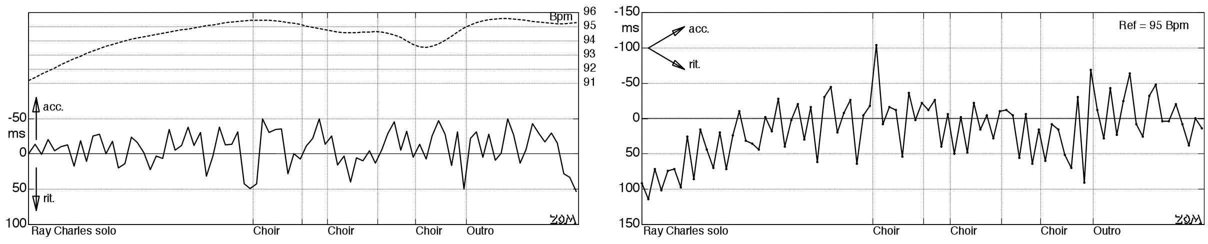

Fig. 8.31: Ray Charles 1981, *Hit the Road Jack*. Absolute (left) and differential-tempo (right) deviations. On the left, the (absolute) deviations are referenced to the smoothed tempo-model (---).

If we subtract out the slow tempo changes (as presented on the right of the figure), short-term fluctuations of a maximum of ±50 ms remain (there is some arbitrariness in that – of course other deviations will result when choosing another bpm-curve). A maximum of 50 ms in a bar of 2,5 seconds – that's all right … especially considering that the larger deviations are a good match to the form of the tune. The first entry of the choir is preceded by a minimal delay that in no way sounds off, but rather expresses the individual interpretation. Two hands on the piano and the voice generate a very lively rhythm that subjectively is perceived as correct – irrespective of what any timepiece says.

♣ Lindsay K.A., Nordquist P.R.: A technical look at swing rhythm in music; http://www.tlafx.com/jasa06_1g.pdf

Fig. 8.32 indicates that "Hit the Road Jack" may also be interpreted in a different manner. Same tune and same singer, but recorded 18 years earlier. Relative to the reference defined at 90,8 bpm, the tempo first minimally lags (-0,4 bpm), then catches up, and drops again towards the end. The deviations rarely cross the 15-ms-mark (it needs to be considered here that in contrast to Fig. 8.31 now quarter-notes are the reference). This is possible because the drums play along from the beginning, and their accents are easier to more precisely measure compared to the start of a piano chord. Neither video footage nor sound documentation reveals whether a metronome was put into service during the recording. Thus hypothetically: a metronome clicks along, the drummer realizes after about 18 s (towards the end of the second chorus) a slight time lag, counteracts and achieves perfect time again at the end of the 2nd chorus. We find larger deviations in the middle of each chorus – which fits the structure because each chorus consists of two halves: this would be a justification for that little swerve in their middle. It's a wrap – next take.

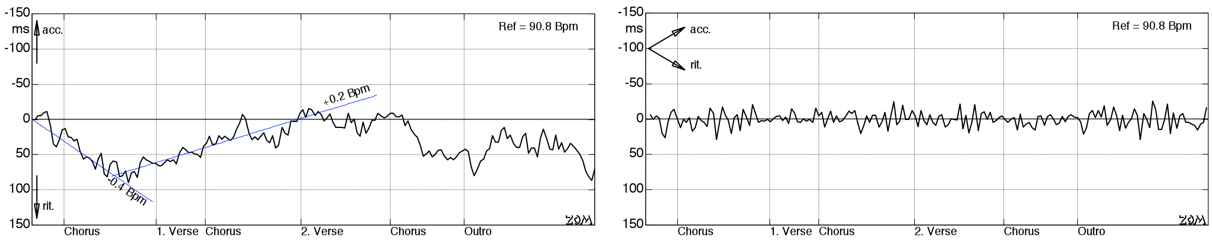

Abb. 8.32: Ray Charles 1963, *Hit the Road Jack*. Absolute (left) and differential (right) quarter-deviations.

Now on to the Eagles, the members of which are taken to be musicians that have a lot experience in the studio. At the start, their **"Heartache tonight"** exhibits a pronounced backbeat, i.e. an emphasis on the even-numbered quarter notes, generated by handclaps. In this tune, the handclap is not always present, therefore the snare drum was analyzed: **Fig. 8.33** shows only very small deviations in the quarter notes following each other at a distance of one second, with the extremes correlating with the structure. Here it can be assumed that a click track was used: both to achieve high precision but also because by now this is usual studio practice (post-processing simply becomes that much easier). This assumption does not at all seek to deny that Mr. Henley does precision work. Indeed, the presence of a click is no guarantee for a precise rhythm.

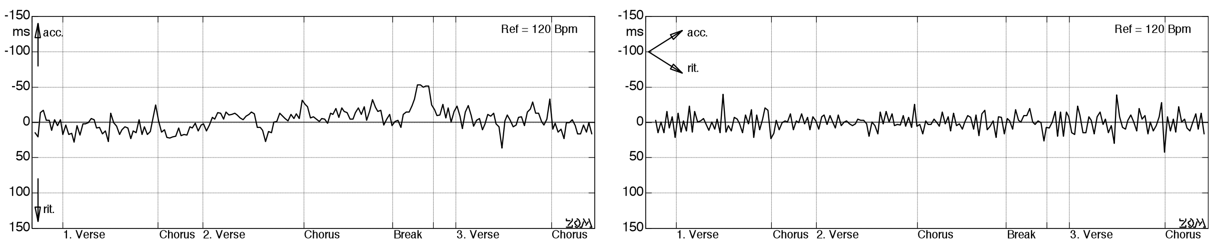

Fig. 8.33: Eagles, *Heartache Tonight*. Absolute (left) and differential quarter-deviations.

"Heartache Tonight" was analyzed as a 4/4-beat: the quarter notes are emphasized, and the counter-beats of the snare drum (offbeats) are on the even-numbered quarter notes. The subdivision of the quarter notes is ternary – it's a shuffle. Each 1st and 3rd quarter is given a grace note that sounds not on the eighth note before but is (due to the shuffle-partitioning) slightly delayed. In triplets notation, the second quarter would be divided in three parts; the grace note would then be located on the last third (eighth triplet). The same would be done ahead of the start of the bar.

Moving to the Rolling Stones, the "biggest Live-Band ever". Active since about 1962, and still rocking the scene♣. Somehow, anyway. When they recorded the live version of their **"Honky Tonk Women"** in 1969, they did have studio experience as well – but somehow in a different way compared to the Eagles[1]. The piece starts at about 103 bpm, speeds up mightily, and ends with about 120 bpm. Not that the guitar had accidentally started off too slowly – no, it's all supposed to be that way. Taking off with the heavy, earthy guitar riff, it ramps up and reaches operating temperature with the first chorus. Towards the second verse, the tempo eases off a bit, and then it's pedal-to-the-metal to the conclusion. Is that wrong? No way – it does groove. The deviations that, after all, are much larger compared to Fig. 8.33 do not stand out much, because the live recording offers a charming "togetherness" in particular in the chorus – it sounds like live recordings from that time simply were. Lively – but not wrong.

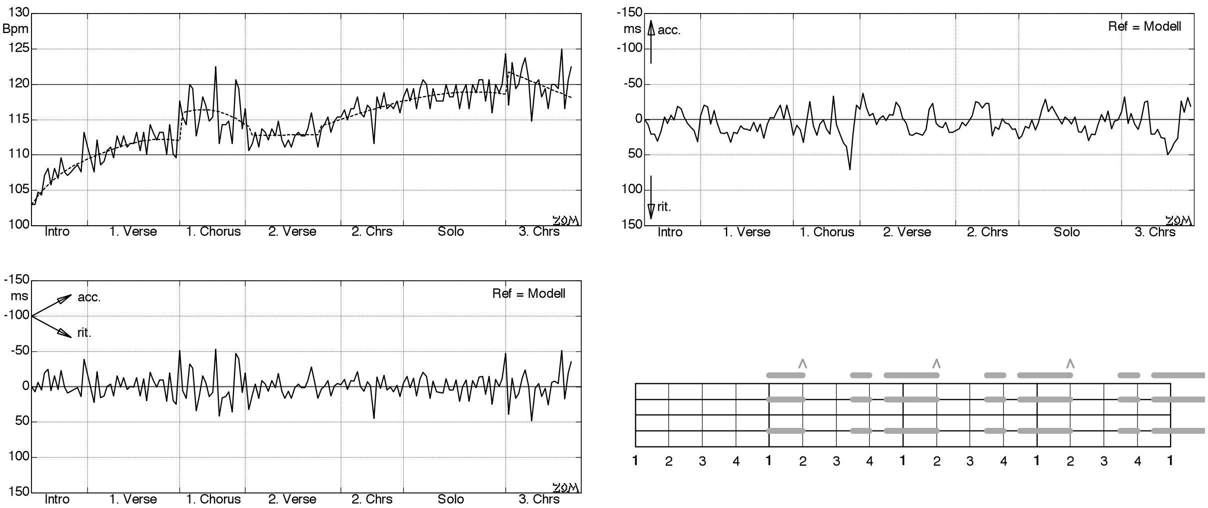

Fig. 8.34: Rolling Stones, *Honky Tonk Women* ("Get Yer Ya-Ya's Out!"-LP, 1969). Bpm-tempo (upper left), and absolute half-note deviations (upper right), referenced to the smoothed tempo model (---). The differential half-note deviations (lower left) also relate to the tempo model. Lower right: rhythm pattern (intro).

The world famous intro-riff is a nice example for the ability of the listener to detect the basic beat even if that is not even played at all. Only at the very beginning is the "one" emphasized in the intro, from then on the power chord is tied to the "four-and" and across the measure line to the following quarter note. The "three", elsewhere the other accent in the standard 4/4-beat, is not emphasized, either. Due to the accent on the "two" that immediately is interpreted as offbeat, the beat is found without any effort. In this example, the *quarter notes* with their time-distance of half a second are interpreted as basic beat; it is this rhythm that the listener's head synchronizes to as he/she "grooves along". To nod with the head two times per second is a very atypical movement. It would be possible to perceive the *eighth note* as basic beat but the corresponding head movement would already be too fast. However, the eighth-note tempo is a great match for drumming along with your fingers. You probably would not want to shake your whole human body in this tempo - but then that's a quite subjective decision. Customarily, 120 bpm is the best "groove along" tempo, which is why it is found often in dance music (moderato – allegretto). The dimensions and masses of the members of grown-up people specify – in conjunction with the spring stiffnesses – the natural frequencies of this "body"-system. And again we find: the tendency to oscillate is especially strong at resonance.

♣ Two musicians talk: „I've read that cockroaches presumably could survive an atomic war" – „Maybe – but Keith Richards would, in any case."

[1] *(Translator's note: about "live": the Eagles' live-version of "Heartache Tonight" is amazingly similar to the studio version – the point where one might ask how live the live-recording actually was …)*

Not everything was entirely in time with the Rolling Stones, as shown by an early recording that (embarrassment-city!!) actually is called **"Time is on my side"**. The tempo remains at a rather constant 67,4 bpm during the first half of the song, and then suddenly drops to 65.4 at the start of the guitar solo (**Fig. 8.35**). That would not be a tragedy because we can see a motivation for this change. The short-term fluctuations that were analyzed using the beat of the snare drum are not really problematic, either (graph on the upper right). What is really off is the tambourine♣ that does get out of (time-) line again and again.

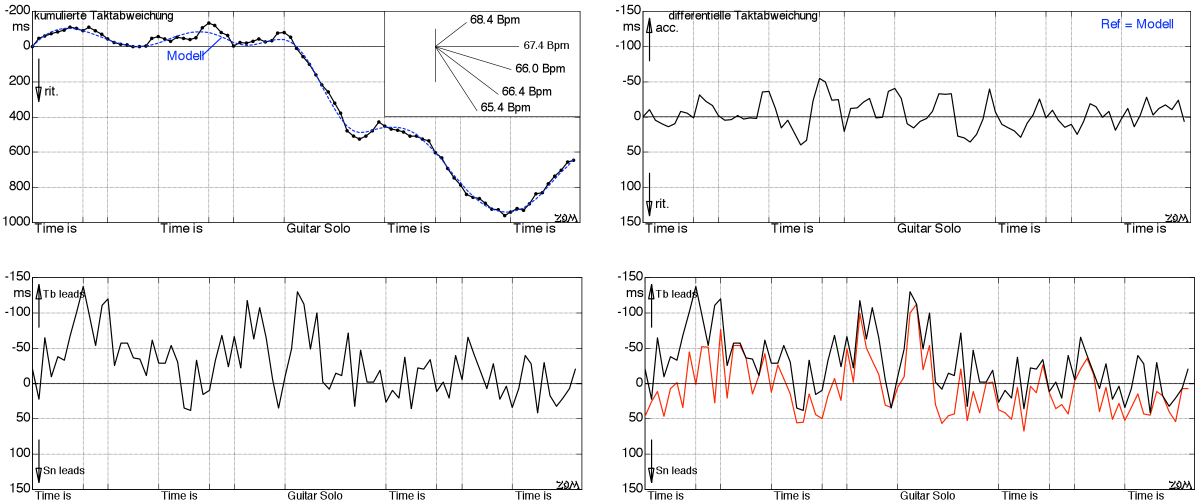

Fig. 8.35: Rolling Stones, *Time is on my side,* 6/8-beat; (Rolling Stones No. 2, 1964).

In the graph on the lower left, the time lag between tambourine and snare drum is depicted; it often surpasses 50 ms, and several times 100 ms – and that is audibly wrong. If the (still) tambourine is not hit *with the (one) hand* but the (other) hand is used to hit the (one) hand *with the tambourine*, then two sounds are created: one when accelerating the tambourine, and the other when it impacts/decelerates. Apparently, the latter playing mode was employed because often two tambourine hits are heard quickly following each other. The first one (shown in the graph on the left) is almost always early while the second one floats around the reference time (shown in the graph on the right as a bold line). So: is time on my (or rather the Stones') side? Not really, that was "out of time" even as early as 1964. Occasionally the two tambourine hits follow each other so quickly that they "melt together", and at other instances it will in fact have been only one single hit. Frequently, the time distance is more than 50 ms, and that's where (as an orientation value) the limit of the appearance of echoes lies. From a time distance of about 50 ms, single echoes become audible; repetitions with a smaller delay are pooled together by our hearing system into *one single* event. The haphazard change from single- to double-impact playing and the strongly fluctuating timing in this example – it is not perceived as stimulation anymore, but only and simply as timing errors.

A slightly delayed tambourine might be still acceptable because it is at least partially masked by the snare drum (accessory masking). However, the early tambourine beat (in the range above about 4 kHz mostly the only event) attracts much attention. This all the more because its percussive character marks a *point* in time – contrary to the slightly open hi-hat, the sizzle of which marks a *range* in time. Having the hi-hat ahead by a sixteenth (150 ms) – yes: that might also have worked. But then it should have been consistently ahead, and not: now on the "three", then on the "four" (together with the snare), and then again on the "three-and-three-quarters". Charley would surely have known all that – someone else must have Jaggered this.

♣ A head-less tambourine is what is meant – the sound of the jingles is discussed here.

As a résumé, let us jot down a few numbers: the limit of perception regarding group-delay distortion is around **2 ms**; for your usual music performance, this value is of no significance. From about **5 ms**, timing errors may be noticeable in individual cases – but only from about **10 ms**, the delay-range relevant for work in the studio starts. To avoid any misunderstanding: we are talking tone/note-onset here, and not a time-shifted superposition! Phaser, flanger, and similar effects devices (Chapter 10.8.3) do generate audible effects already at very short delay times, but these effects are noticeable because of the variations in the spectral envelope, and not as time-related in the first place. Given short repetition-periodicity, the perception threshold for timing drift may be as small as 10 ms (depending on the circumstances), but for rhythms that are not as fast it will extend to about **20 ms**. And that's it, then – even greater delay may indeed become problematic and sound wrong. Still: it may – but doesn't have to.

If the rules of timing could be summarized in a few lines, there would be great-sounding drum-machines only. Exceptions start already in a simple **¾-beat**: playing all three quarter-notes exactly on the regular time will sound wrong. Playing the second quarter note just a tad early creates a fit, making the listeners feel that all quarter notes have the same time-distance. A corresponding shift of as little as about $1/6^{th}$ of a quarter note will suffice, depending on the tempo. However, even with this shift of the second quarter note, the rhythm will start to become monotonous – for example if it's always the same drum pattern that is repeated. A drummer would indeed never repeat exactly the same beat; the drumstick will hit the skin at different positions leading to similar beats that still differ in the details. He/she would also introduce small variations: on top of the sound-color, he/she would vary the volume, and – yes – the timing, as well, in order to optimally support the musical piece. All this is alien territory for a drum-machine on its simplest programming level, and so it sounds just a primitive apparatus. Which it in fact is.

The **shuffle**, that galloping groove onomatopoetically described with "dumm-da-dumm-da-dumm", is an example for large time-shifts. Here, every eighth-note (if a 4/4-beat can be taken as a basis) is played later than the binary notation would require it. How much later – that is left to the artistic interpretation, and it determines whether the musicians in the ensemble play with or against each other. The musician ostentatiously playing a shuffle written "evenly" (i.e. binarily) just as evenly will kill the whole number although he considers him/herself as doing the right thing. Switching to ternary notation (triplets) helps only to some degree because this does not express, either, how strong the "shuffeling" actually is. That is decided by the "feeling", the experience, and expression of the musicians, and by their ability to empathize. That's empathizing with the interpretation of the co-musicians, and with the given piece of music. As a band "grows together" by frequent rehearsals, each musician acquires experience about how the others interpret the music, and in the end everyone "shuffles" so that they all match. That does not necessarily mean that they all play with exactly the same deviations.

Table: tempi (bpm)

Largo	40 – 60	Larghetto	60 – 66	Adagio	66 – 76
Andante	76 – 108	Andantino	100 – 108	Moderato	108 – 120
Allegretto	120 – 132	Allegro	120 – 168		
Presto	168 – 200	Prestissimo	from 200		

Table: meter

Iambus:	– o – o – o – o	Trochee:	o – o – o – o –
Daktyl:	o – – o – – o – – o – –	Anapaest:	– – o – – o – – o – – o

