Proceedings of the 8th International Conference on Music Perception & Cognition, Evanston, IL, 2004
ICMPC8
S.D. Lipscomb, R. Ashley, R.O. Gjerdingen, & P. Webster (Eds.)
Adelaide, Australia: Causal Productions

TOWARDS A THEORY OF TEMPO MODULATION

Fernando Benadon
University of California, Berkeley

ABSTRACT

Tempo modulation (TM) is defined as a change of tempo by means of a shared durational unit. Since it follows a simple pivotal premise, TM provides an efficient way to achieve rhythmic complexity without imposing a high degree of performance difficulty. This kind of rhythmic activity operates on the level of subjective rhythmization. TM functions on two additional levels of temporal perception: the sub-100ms region permeated by new possibilities in beat subdivision, and the longer temporal spans governed by proportionally related tempos in a tempo network.

1. INTRODUCTION

Even though tempo (or metric) modulation (TM) is an often-used technique, its theoretical foundations have been addressed only indirectly. In what is perhaps the most thorough essay on the subject, Nicolas (1990) discusses the relationship of TM to meter and tempo, laying out some compositional applications such as tempo "spirals" and "paradoxes." He argues that "tempo must from now on attain an emancipation previously unknown to it, with the possible exception of the distant past when tempo was not even notated" (p. 75, my translation). In an early theoretic essay, Perkins (1965) explores the rich multiplicity of note values offered by "conventional" notation, using an ingeniously devised logarithmic slide ruler to illustrate possible TMs via different note values. Analytical discussions of TM abound whenever the music of Elliott Carter is involved (e.g., Bernard 1998), but such analyses are confined -- understandably -- to TM function within specific works. Elsewhere, Schuller (1968, p. 117) invokes TM to analyze Louis Armstrong's "unorthodox [rhythmic] procedure" in his *West End Blues* introduction. Though not an article on TM per se, Schick (1994) is worth mentioning here because he explains how a performer can simplify complex polyrhythms by reinterpreting them using tempo changes. These writings reflect a widespread interest in TM and the need for a thorough exploration of the subject. By addressing some of the main properties of TM, this paper will help to formulate a general theory that integrates music-theoretic issues within a cognitive framework.

2. DEFINITIONS

TM is defined as a change of tempo by pivoting on a common durational unit. Figure 1-a shows an example of a TM where the sixteenth-note serves as the pivotal value. This operation transforms the tactus speed and all its subdivisions by a factor of 4/7; expressed in ratios, the tempo relationship from old to new is 7:4.

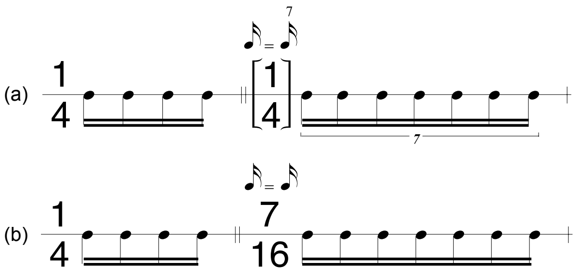

Figure 1: Two different versions of the same modulation.

The same modulation can be renotated as in Figure 1-b. In this case, the use of the septuplet is avoided in favor of a time signature change. That is, the tempo change is expressed as a metric modulation. Since any TM can be notated using two formats, it is not surprising that the terms "tempo modulation" and "metric modulation" have been used interchangeably. The former term is employed here for the reason that while a modulation *can* be expressed as a metric shift, it is *always* a tempo shift. The downside of this choice is that "tempo modulation" has been used to describe such things as continuous tempo fluctuations (including ritardandos and accelerandos, as well as microtemporal activity in the context of expressive timing) and metronomic proportions between movements or sections of a composition (Epstein 1995). TM will be used here to describe a modulation as shown in Figure 1-a. We should also clarify that the term "modulation" will denote a modulation of tempo, and should not be confused with other analogous modulatory processes such as harmonic modulation or frequency modulation.

The term "tempo" should be clarified also. Tempo is usually defined as "the speed of the music," and is sometimes broken down into two interacting components: pulse- and activity-tempo (e.g., Berry 1987, p. 305). Both categories refer to perceived onsets (isochronous or not) which the listener organizes according to different grouping mechanisms. This leads to what might be described as *emergent tempo*, which is different from what we might call *conceptual tempo*: the family of durational values resulting from different subdivisions of a fixed reference value. The relationship between emergent and

ISBN 1-876346-50-7 © 2004 ICMPC

conceptual tempo may be stated as follows: conceptual tempo provides an abstract durational grid over which different emergent tempos may be manifested. While these are usually one and the same in common practice and popular music, composers such as György Ligeti and Conlon Nancarrow are known for exploiting the malleable relationship between the two tempo types. Precisely because emergent tempo is a cognitive/compositional construct, any forthcoming usage of the term "tempo" will refer to the conceptual form.

3. SUBDIVISION SPACE

A given complex rhythm can be notated in different ways. Usually, increased rhythmic precision means increased performance difficulty. Proportional notation (where horizontal note-placement corresponds to onset time) and verbal instructions (such as "freely") do not present significant technical difficulty for the performer, but the alignment of temporal events is approximate and potentially unreliable. Two other techniques that are more accurate are nested tuplets (such as a triplet embedded in a quintuplet) and what might be called non-quantized notation (a combination of dots, ties, and extremely small note values). These techniques can convey very precise rhythmic information at the cost of imposing high levels of performance difficulty, complicating practical issues such as ensemble synchronization and rehearsal time. The opposite is true of TM, which can achieve significant rhythmic complexity by means of a limited set of note values, combining high precision with low difficulty. How?

A TM causes a change in the hierarchical relationship between the perceived beat subdivision and all potential subdivisions belonging to the new tempo. For instance, in Figure 1-a, the relationship between the sixteenth-note and the eighth-note is at first 2:1. After the modulation, this relationship becomes 7:2. Even though the old sixteenth-note is now notated as a septuplet, both values proceed at the same speed. So, all else being equal, they are perceived as identical.

A tempo marking denotes not only the frequency of a reference tactus but also the durations of its subdivisions, a set of *discrete* values. For example, at 100 bpm, eighth-note IOIs = 300 ms, eighth-note triplets = 200 ms, sixteenth-notes = 150 ms, etc. In other words, only a limited number of available durations is available, determined by the metronomic marking (though in theory an infinite number of subdivisions is conceivable). If one wanted to speed up the triplets slightly, the next available "notch up" yields a not-so-subtle difference of 50 ms (the sixteenth-note). Since a feature of (micro)rhythmic complexity is the unrestricted malleability of durations, the capacity to narrow the gaps between different subdivisions would seem desirable. Table 1 illustrates how *subdivision space* can be saturated by means of TM. As stated above, other notational approaches can be used to fill in the gaps between subdivision values. The advantage of TM is that it can do so by resorting to simple ratios and without increasing notational complexity.

80 bpm	**ms**	**ms**	***100 bpm***
quarter-note	750		
		600	*quarter-note*
dotted eighth-note	563		
quarter-note triplet	500		
		450	*dotted eighth-note*
		400	*quarter-note triplet*
eighth-note	375		
eighth-note quintuplet	300	300	*eighth-note*
eighth-note triplet	250		
		240	*eighth-note quintuplet*
		200	*eighth-note triplet*
16th-note	188		
16th-note quintuplet	150	150	*16th-note*
		120	*16th-note quintuplet*

Table 1: A 5:4 modulation (80 bpm → 100 bpm) yields new subdivision values formerly unavailable in the original tempo, filling the gaps between adjacent subdivisions.

As an additional example, consider a scenario where the only available subdivisions are the eighth-note triplet, the sixteenth-note, the sixteenth-note quintuplet, and the sixteenth-note septuplet (i.e., 3, 4, 5, 7). Beginning with any of these subdivision speeds, only three possible destinations are possible. For instance, 5 can transition to either 3, 4, or 7, equal to proportional relationships of 5:3, 5:4, and 5:7, respectively. TM offers nine *additional* destinations, significantly increasing the palette of rhythmic expressivity (Table 2).

From a listener's standpoint, the precise moment of modulation goes by unnoticed (unless it is made explicit through accentuation, etc.), because the subdivision speed of the old tempo equals that of the new tempo. It is not until the first subdivision change after the modulation that the listener registers an acceleration or deceleration in the frequency of onsets. Now, let us suppose that (1) the musical context is such that the point of modulation is preceded by a long section where the predominant subdivision is one other than that effecting the modulation, and (2) there is a change of subdivision immediately following the modulation, as shown:

$$a \;\text{....} \rightarrow \underbrace{b = c}_{\textit{brief}} \rightarrow d \;\text{....}$$

5 =	→	ratio
3	7	3:7 (.43)
4	7	4:7 (.57)
3	5	3:5 (.60)
3	4	3:4 (.75)
4	5	4:5 (.80)
4	3	4:3 (1.33)
7	5	7:5 (1.40)
7	4	7:4 (1.75)
7	3	7:3 (2.33)

Table 2: In ascending order of magnitude, nine possible transitions from a quintuplet, using TMs with a restricted set of subdivisions. Note that at certain tempos, some resulting subdivisions may be exceedingly fast, so "octave" adjustments (e.g., eighth-note septuplet instead of sixteenth-note septuplet) may become necessary.

a	b = c	→ d	ratio
4	3 = 4	3	16:9 (1.78)
		5	16:15 (1.07)
5		3	20:9 (2.22)
		5	4:3 (1.33)
4	3 = 5	3	20:9 (2.22)
		4	5:3 (1.67)
5		3	25:9 (2.78)
		4	25:12 (2.08)
3	4 = 5	3	5:4 (1.25)
		4	15:16 (.94)
5		3	25:12 (2.08)
		4	25:16 (1.56)

Table 3: Expanded set of modulation ratios. Only 3-, 4- and 5-type subdivisions (without reciprocals) are shown.

In such a situation, the arrival of **d** elicits not only a comparison with **b** or **c** (as illustrated in Table 2), but also with **a**, which is stored in memory. This gives rise to a new and expanded set of proportional relationships (Table 3).

It should be emphasized that the rhythmic flexibility afforded by TM is memory-based; that is, it operates horizontally. Since the vertical alignment of available subdivisions remains invariant regardless of tempo, modulating does nothing to increase or reduce the potential complexity of rhythmic polyphony.

As we have seen, the fine temporal gradation made possible by certain modulation ratios can affect rhythmic behavior on the microtemporal level. Additionally, a tempo change establishes a new hierarchy of durations whose different configurations influence our perception of the tactus, meter, and rhythmic groupings. These operate on the level of subjective rhythmization, with different metric options tapping into different sub-regions of temporal perception (London 2002). On yet another level, TMs can be used to control the organization of tempos as structural elements across longer spans of time, as shown next.

4. TEMPO NETWORKS

So far we have considered modulations between only two tempos. The concept of a *tempo network* allows us to connect several different tempos via modulations, expanding the scope of TM function to longer stretches of time. Let us define a tempo network as the set T containing a "tonic tempo" (**tt**) and other tempo values related to **tt** by ratio multiplication. How many tempos can or should a network contain? Unfettered ratio multiplication yields an infinite number of possible tempos -- hardly a desirable compositional paradigm. Therefore it will prove helpful to determine factors that can regulate the number of tempos in a network. A logical choice is to limit the quantity of available pivot ratios. For example, a modulation such as "dotted eighth-note quintuplet equals sixtyfourth-note" is conceivable but probably impractical. Another way to control the number of tempos in a network is to limit the number of multiplicative steps used to obtain new tempos. Let R be the set of usable ratios in T. Suppose R contains all six fraction combinations of 3, 4, and 5, and that **tt** = 80 bpm. A 1-step limit yields the following tempo network: T= {48, 60, 64, **80**, 100, 106.67, 133.33}, where each element in the set is obtained by multiplying 80 by a ratio containing either 3, 4, or 5 in the numerator or denominator. Using a 2-step limit expands the same network from 7 to 19 elements; now T = {28.8, 36, 38.4, 45, *48*, 51.2, *60*, *64*, 75, **80**, 85.3, *100*, *106.67*, 125, *133.33*, 142.22, 166.66667, 177.78, 222.22} (step-1 values are in italics). The number of tempos brought on with each additional step is given by the "unordered sample with replacement" formula, (n-1+r)!/[(n-1)!r!], where n = number of ratios in R, and r = number of steps from **tt**. Thus, a 4-step network employing six ratio types (such as 7:5, 7:4, 7:3, 5:4, 5:3, 4:3) yields a total of 210 tempos: 1 (**tt**) + 6 (r = 1) + 21 (r = 2) + 56 (r = 3) + 126 (r = 4). If the ratio reciprocals are added (musically this would make sense), then the total number of tempos in this network is 209 + 209 + 1 = 419. For most practical purposes, clearly this amounts to an unnecessarily large number of tempos. The example is presented because it illustrates how basic iterative operations can produce more than enough tempos to structure a composition and to yield useful microrhythmic combinations such as those listed in Table 1.

5. PERFORMANCE , PERCEPTION

How difficult are TMs to perform? The degree of temporal precision attainable by the performer will depend not only on his or her technical skill, but also on other factors such as tempo, dynamics, and instrument type. In a music-making setting, other elements such as pitch direction, melodic skips, and accent structure will affect not only the difficulty of the passage but also the listener's tempo judgment (Boltz 1998). At best, highly skilled musicians are able to reduce the variability of isochronous tapping tasks to just under 3% of the mean IOI (Madison 2000); the range of 3-6% has been widely reported. This is roughly equivalent to the range of tempo discrimination using monotonic, isochronous sequences (Drake & Botte 1993). For IOIs under 250 ms, listeners can discriminate changes as small as 6 ms (Friberg and Sundberg 1995). To put these figures in context, this means that a transition from sixteenth-notes at 80 bpm (188 ms) to eighth-note triplets at 100 bpm (200 ms) (refer to Table 1) would probably be detected by a listener. Of course this seems highly optimistic given a musical situation that involves human performers playing acoustic instruments, but the data serve as a point of departure for further study. Exactly how accurately different TMs can be performed and perceived can only be ascertained empirically, an avenue that is certainly advocated here.

As Table 3 shows, some modulation functions are nearly identical, which raises the question of why one might choose to modulate using (say) a 25:12 ratio rather than 20:9 or simply 2:1, since perceptually these will appear virtually equal. From a composer's perspective, there are three reasons for selecting a modulation ratio from among nearly-equal possibilities. First, the subdivision type offered by the modulation function may provide the most appropriate vehicle to develop the musical material at hand. Hence it is not so much the destination tempo that concerns us, but rather the process of transformation itself. (The same argument applies to harmonic modulations, especially for listeners who do not have absolute pitch and/or who are unable to track chromatic modulations: the intermediary steps are more easily followed (and enjoyed?) than the actual denomination of the destination key.) In fact, this reason holds even when two possible modulations *do* offer exactly equal ratios. Second, since the available subdivisions are dependent on the present tempo, a difference between two nearly equal destinations may be significant insofar as what subdivision possibilities the new tempos can offer. Third, if the work employs digital technology or a click track, nearly-equal ratios may not be readily distinguished but they can be compounded and rendered perceivable over a longer span of time. For example, a 1-minute section composed of a modulation loop 16:15 → 16:15 → 16:15 … constitutes a gradual accelerando.

6. CONCLUSION

TM is a powerful compositional technique because it works on three general areas of time perception: the microrhythmic level, the level of subjective rhythmization, and the larger structural level spanning anywhere from several seconds to hours. Moreover, TM can generate complex rhythms without significantly increasing performance difficulty. This claim is not proven empirically here, but it is supported by a discussion of relevant research findings. Therefore the ideas presented in this paper will no doubt expand in scope as they are tackled directly in ensuing experiments.

7. REFERENCES

1. Bernard, J. W. "The Evolution of Elliott Carter's Rhythmic Practice." Perspectives of New Music, Vol. 26, No. 2, 164 - 203, 1998.
2. Berry, W. Structural functions in music. New York: Dover, 1987.
3. Boltz, M. "Tempo discrimination of musical patterns: Effects due to pitch and rhythmic structure." Perception & Psychophysics, 60, 8, 1357 - 1373, 1998.
4. Drake, M., and Botte, M.C. "Tempo sensitivity in auditory sequences: Evidence for a multiple-look model." Perception & Psychophysics, 54, 3, 277 - 286, 1993.
5. Epstein, D. Shaping time: Music, the brain, and performance. New York: Schirmer Books, 1995.
6. Friberg, A., and Sundberg, J. "Time discrimination in a monotonic, isochronous sequence." Journal of the Acoustical Society of America, 98, 2524 - 2531, 1995.
7. London, J. "Cognitive constraints on metric systems: Some observations and hypotheses." Music Perception, Vol. 19, No, 4, 529 - 550, 2002.
8. Madison, G. "On the nature of variability in isochronous serial interval production." Rhythm: Perception and production. The Netherlannds: Swets & Zeitlinger, 95 - 113, 2000.
9. Nicolas, F. "Le feuillete du tempo: Essai sur les modulations métriques." Entretemps, 9, 51 - 77, 1990.
10. Perkins, J. M. "Note values." Perspectives of New Music, Vol. 3, No. 2, 47 - 57, 1965.
11. Schick, S. "Developing an interpretive context: Learning Brian Ferneyhough's Bone Alphabet." Perspectives of New Music, Vol. 32, No. 1, 132 - 153, 1994.
12. Schuller, G. Early jazz: Its roots and musical development. New York: Oxford University Press, 1968.