Actes du Congrès international de recherche en sciences humaines et sociales, Paris 24-28 juillet 2012
(G. Tchibozo, dir.) – Strasbourg (France) : Analytrics, 2012, p. 131–139

Proceedings of the International Congress of Humanities and Social Sciences Research, Paris 24-28 July 2012
(G. Tchibozo, ed.), Strasbourg (France): Analytrics, 2012, pp. 131–139

MUSICAL SCALE HIERARCHY

Pavel B. Ivanov

Abstract

L. Avdeev's theory of musical scale hierarchy formation is discussed. The model provides an adequate mathematical description of the already known scales and reveals some other new possibilities. A measure of the information difference between two probability distributions is employed to construct a numerical estimate of a contradiction between two compound tones. The discordance function thus obtained allows to determine a musical scale. A dissonance function is introduced, which reveals the scale as a set of zones. Stationarity under nonlinear transformations and maximum regularity provide the numerical criteria for selecting the preferable scales. The formant structure of the internal timbre pattern characterizes the stability of local hierarchical structures. Harmonic, modal, and chromatic types of scale lability are described. Musical scale as a movable hierarchy of zone structures unfolds itself in various ways, forming the musical context.

Introduction

The traditional application of mathematics to music, from Pythagoras to now, is reduced to mere combinatorics and the search for statistical regularities. The insufficiency of such an approach has become evident in developing the tools for computer aided music composition. It has become clear that a human composer is never stochastic, and we need to comprehend the psychology of music perception, to model some aspects of composition. And, primarily, we need to explain the origin of the 12-tone chromatic scale as its principal material. In 1980s, Leonid Avdeev has suggested a consistent theory of scale formation (Avdeev & Ivanov 1993; Avdeev *et al.* 2006) based on the following principles.

1. There is no direct correlation between physical properties of sound and the perceived intonation (Gelfand 1981; Pozin 1978). We rather construct an internal model of the sound and then try to fit all what we hear into this pattern. That is why we speak about sound *perception*, which assumes sound sensation shaped by representation (Ivanov 1994).

2. Both pitch perception and intonation are of a *zone* character (Garbuzov 1948; Johnson 1966).

3. Historically, mere pitch discrimination precedes the notion of a musical interval (Kvitka 1971). The fifth and the octave take their place in music rather late; hence one cannot seriously speak about combinatorial origin of musical modes.

4. Generally, musical hearing evolved in the direction of distinguishing still higher overtones and a more detailed perception of timbre (Schönberg 1973).

5. Any human activity is *hierarchically* organized (Moles 1958; Pribram 1971; Leontiev 1975), including the perceptual activity.

6. Such musical phenomena as consonance and dissonance, tension, instability and steadiness arise

only in a specific *context* (Schönberg 1973), and the theory must somehow describe it.

7. A detailed investigation of the mechanisms of perception gives little for the comprehension of perception itself. Rather, construction of higher-level models permits to bind physiological data together.

8. Human perception is a kind of information processing, and its models should be compatible with *information* theory (Moles 1958; Pribram 1971; Golitsyn 1980).

9. The available physiological and psychological data reveal the complex organisation of hearing (Gelfand 1981; Pozin 1978; Garbuzov 1948, 1956), and the theory must comply with those data.

Hierarchical Scaling

On the lowest level, the perception of a pure tone if pitch (logarithm of sound frequency) h is modeled with an inner pitch distribution

$$f(x) = (2\pi)^{-1/2} \sigma^{-1} \exp\left[-\frac{1}{2} (x-h)^2 \big/ \sigma^2 \right].$$

The dispersion σ is a historically formed parameter of perceptual tuning related to the overall sharpness of hearing (intonation distinction).

A musical tone (abstracted from noises, phase effects, etc, which are unessential for the theory of scaling) is represented by its harmonic series, that is, the set of partials h_j with the amplitudes t_n $(n = 1 \ldots N)$. This set of amplitudes will be called *internal timbre*. This is a perceptual set only related to pitch hearing; it does not depend on the physical or perceived timbres of the musical instruments, since timbre hearing is a different activity.

In the process of hearing, people establish relations between tones. Subjective interference between two partials can be estimated as a quantity of information in one distribution relative to another (Golitsyn; Avdeev & Ivanov 1993):

$$d(R) = (R^2 / \sigma^2) \exp\left(-R^2 / \sigma^2 \right),$$

where R is the pitch difference. The quantity $d(R)$ is called the *discordance value* for the two partials. For two complex tones represented by the internal timbre, the discordance is calculated as

$$\Delta(h) = \sum_{m,n=1}^{N} t_m t_n d[h + \log_2 (m/n)].$$

This value characterises the contradiction of the two sounds, the difficulty of including them in a common system of tones. A few samples of discordance function given in Fig. 1 manifest a number of minima, which might be related to the possible degrees of a musical scale.

On the next level, various tone structures formed of the minimally discordant tones are perceived. Not all timbres are of equal worth for articulated structure perception. Thus, if the timbre is not preserved during various inner operations, it will hardly become culturally fixed; the criterion of timbre stationarity $0 \le \tau < 1$ is introduced as a measure of the difference between the original timbre and its quadratic transform. Timbres with high τ are called *stationary*.

However, mere stationarity is not enough for musical scale formation. The tuning of perception to a leading "rhythm" in the discordance function leads to a *regularity* requirements; stationary timbres producing regular discordance functions are called *optimal* timbres; they are the most suitable for scale formation.

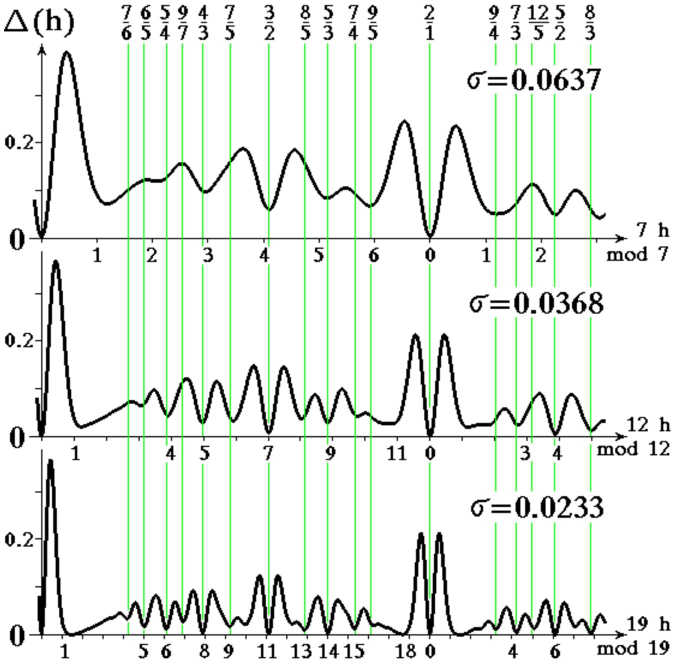

Figure 1. Discordance functions.

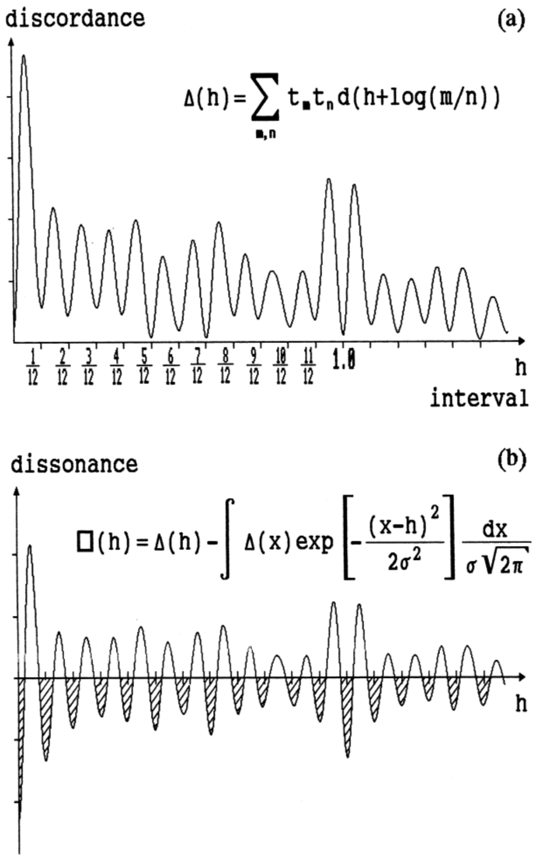

Figure 2. Discordance and dissonance.

The degrees of a musical scale can be introduced as some zones around the minima of discordance function. The belonging of a tone *h* to a zone is defined through the *dissonance function* derived from the primary discordance:

$$\square(h) = \Delta(h) - \int_{-\infty}^{+\infty} dx \Delta(x) f_h(x) .$$

The integral represents a local average with the Gaussian weight. The discordance and dissonance functions are compared in Fig. 2. Now, one can define a musical scale as a collection of pitch zones (Garbuzov 1948).

Computer-aided calculations permit us to find a series of optimal stationary timbres. The properties of some corresponding scales (for example, with 2, 3, 5, 7, and 12 tones to the octave) are well known in musicology. We can therefore understand how a timbre pattern correlates with the scale properties, and with a good certainty describe some new scales, hardly yet established in musical practice.

It has been discovered that each optimal timbre has a well pronounced *formant structure*. The musical quality of a scale depends on the formant pattern of its optimal timbre. The partials of the first formant indicate which intervals in the scale can sound in accord simultaneously, so we refer to the first formant as *harmonic*. The length of the harmonic formant correlates with the richness of harmony in the scale, and musical hearing evolves in the direction of ever more complex chords. There are *anomalous* scales, with only the octave possible as a harmony. Four partials in the harmonic formant give *quintal* scales which admit also the fifth (and the fourth) as a harmonic interval; they are still rather poor for the modern musical thought. *Tertial* scales, with six partials in the first formant, introduce the third in harmony, which makes it rather good in most cases. The most interesting harmony can be achieved in *harmonic* scales, with more than six partials in the first formant. Only in such scales, harmony can freely use the seventh and the second.

We refer to an interval that is generated *inside* the second formant as a *melodic move*, while an interval between a partial of the second formant and a partial of the first formant is called a *melodic jump*. Both moves and jumps define characteristic intervals of a *mode*, so we call the second formant *modal*. When harmonic intervals prevail, the scale is *harmonically labile*, so that any chord may play a centralizing role, and music cannot be tonally organized in a wider range. The opposite case, when modal intervals prevail, leads to *modal lability*, with weak tonal definiteness in a melodic sequence. In both cases, higher levels of musical hierarchy can hardly be built. This property is usually exploited to produce some coloristic or stylistic effects, but generally it restricts the practical use of the scale. The criterion of scale lability is introduced as

$$L = \log \frac{M(M-1) + 2MH}{H(H-1)} ,$$

where *H* and *M* are the numbers of partials in the harmonic and modal formants respectively. Now, $L < 0$ indicates harmonic, $L > 0$ modal lability. Well balanced timbres with $|L| < 1$ lead to *stable* scales, which can be universally used. The well-known example is the common 12-tone scale. Scales with lower number of zones cannot be stable.

Higher partials usually concentrate in a vast *chromatic* formant which is responsible for the distinction of narrow intervals. The basic scale becomes thus more pronounced, quite like the higher vocalist's formant makes vocal performance more articulate. Too long a chromatic formant may, however, lead to *chromatic lability* of the scale, that is, an easy drift from any musical structure by a narrow interval. On the other hand, if the maximum chromatic move is too short, then some modal moves will not have chromatic analogues. Also, for chromatic stability, the minimum chromatic move should be narrow enough, so that a chromatic alteration of a degree would not change its function.

Chromatically stable scales are those with $\nu = 10, 12, 17, 19, 22, 31$.

With a pre-selected optimal timbre, one can tune hearing to the structures different from the basic scale as defined by the dissonance function. Formally, this corresponds to dispersions $\zeta < \nu$ and the timbre truncated to the number of partials $Z < N$. Such a reduction is not arbitrary; the possible *embedded* scales are determined by the Fourier spectrum of the dissonance function. Each scale has a number of possible subscales. An example of the hierarchy of possible embeddings is presented in Fig. 3.

This hierarchy can be unfolded in different ways. In particular, the systems of embeddings $12 \supset 7 \supset 3$ and $12 \supset 7 \supset 5$ can form. If an embedded scale is harmonically labile, it can play the role of harmony in the basic scale (as the imbedding $\zeta^Z = 3^6$ in the $\nu = 12$ scale). A modally labile or stable embedded scale defines a possible mode or tonality ($\zeta^Z = 5^9, 7^{10}$). Naturally, the reference tone of any embedding may differ from the key tone of the basic scale, which just specifies the pitch of a layer in a musical piece.

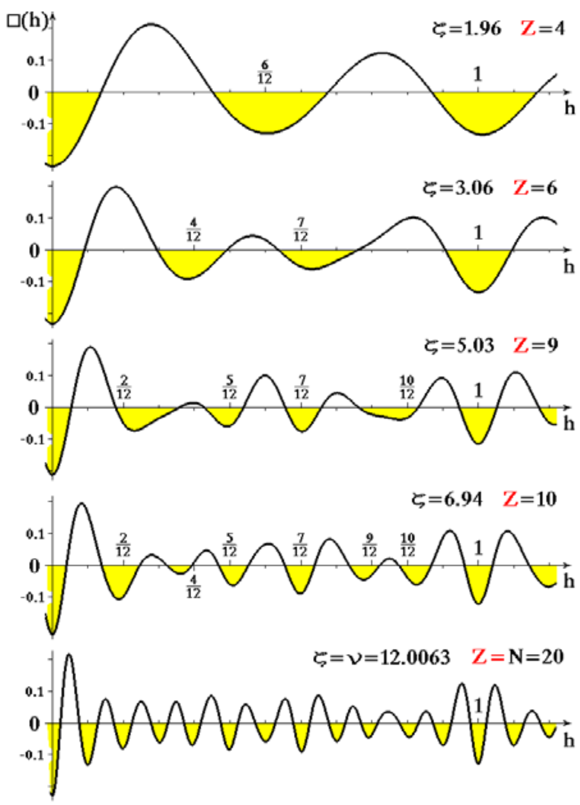

Figure 3. Scale hierarchy for the 12-tone scale.

The movable hierarchical system of embeddings represents the current musical *context*, defining dissonances and tensions on various levels. In general, a dissonance on a certain level occurs whenever an actually sounding tone does not fit into the current imbedded scale of that level. The tensions of a sound according to the imbedding with the discordance function $\Delta(h)$ can be estimated by the integral

$$\rho(h_1, h_2) = \int_{h_1}^{h_2} \Delta(h) dh .$$

For example, in the $12 \supset 7 \supset 3$ system of embeddings, describing the usual tonal music, there exist harmonic dissonances ($\zeta = 3$ level) and modal dissonances ($\zeta = 7$ level), which can be resolved inside the 12-tone scale. There are two ways of releasing the tension: a move to the nearest consonance, and a change of the current embedding system; both take place in real music.

The 19-tone Scale

The $\nu = 19$ scale is a stable harmonic scale with 10 partials in the harmonic format and 5 partials in the modal formant; there is also a well balanced chromatic formant. The hierarchy of scale embeddings for the 19-tone system is displayed in Fig. 4. As one can see, the scale allows modelling the usual tonal music within the $19 \supset 7 \supset 3$ system of embeddings. Additionally, there are 12-tone embeddings, a sort of hypermodes, or "hypertonalities" (Schönberg 1973). Some of them reproduce the usual ascending and descending chromatic scales; however, in 19-tone music, the tones of the hypermode are not evenly spaced, manifesting a specific structure of tensions. Also there are hypermodes that involve no diatonic, demanding a qualitatively different system of embeddings: $19 \supset 12 \supset 5$. Here the $\zeta = 5$ level plays the role of harmony in the 12-tone hypermodes. Note that the 5^9 imbedding in the 19^{31} scale is harmonically labile, though the pentatonic $\zeta = 5$ is modally labile as an independent scale.

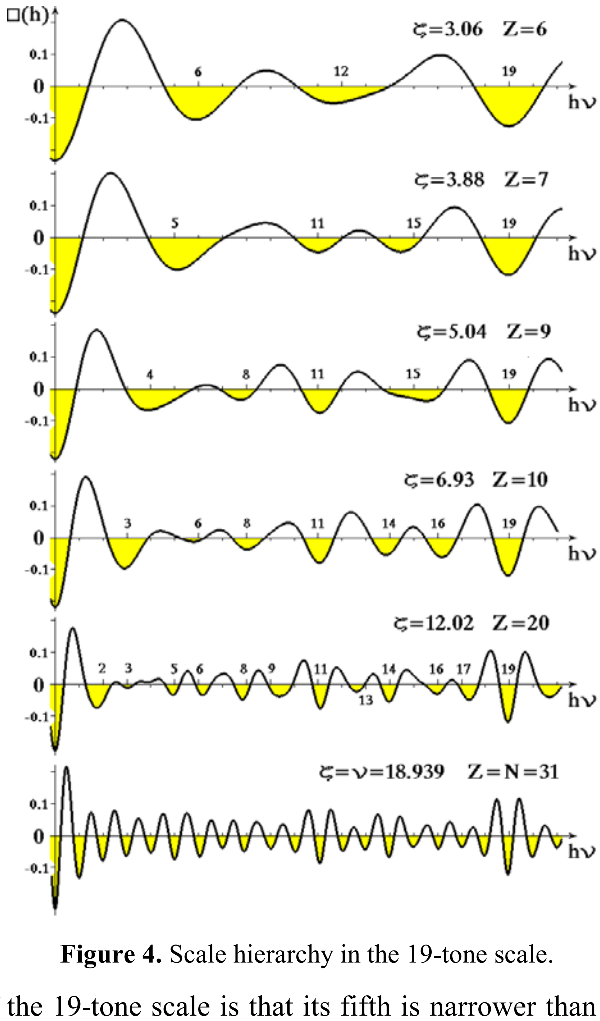

Figure 4. Scale hierarchy in the 19-tone scale.

The important feature of the 19-tone scale is that its fifth is narrower than the fifth of the traditional 12-tone system (which is close to the natural fifth). For hearing tuned to the 12-tone scale, such a "tempered" fifth sounds rather unusual; it is softer and richer. In vocal intonation, it does not require a register shift and can be even reproduced by untrained people. In the context of 19-tone scale, this fifth is perceived as a full-fledged harmony, so that the traditional fifth avoidance is no longer necessary.

Playing 19-tone music requires specially designed instruments. Of course, one can tune a violin or a cello by the tempered fifths; however, the acoustic properties of the traditional instruments will hardly reveal in full the richness of new harmony. One could build a piano suited for 19-tone music adding one more row of keys (for example, painted in red), as shown in Fig. 5. Thus, instead of every black key, between the adjacent white keys there are now *two* intermediate degrees in case of a diatonic whole tone, while a diatonic semitone (*B–C*, *E–F*) is *divided* by a red key in half. With the equal temperament, all the intervals between the nearest keys are the same (1/19 of the tempered octave which should be a little stretched to the frequency ratio $2^{19/v} \approx 2.00447$). The white and red keys separately form two diatonic 7-degree scales, while the black keys, as previously, assemble into a pentatonic.

The notation of the 19-tone music extends the usual musical notation introducing the signs for half-flat and half-sharp along with the usual flat and sharp. Such signs can already been found in the musical literature; but they traditionally refer to the half of the semitone in the 12-tone scale, while, in the 19-tone system, they have significance on their own, as the format structure of the corresponding optimal timbre predicts. A “whole tone” contains three steps of the scale; a “semitone” (the distance between *mi* and *fa*) contains two steps of the scale. Additionally, an “introductory semitone” corresponds to the interval of one step of the scale. Introductory intervals are widely used in music, though they are poorly reproduced on 12-tone tempered instruments.

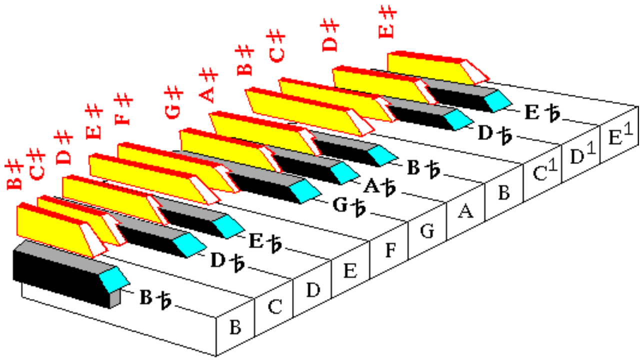

Figure 5. The 19-tone keyboard.

For usual tonal music performed within the $19 \supset 7 \supset 3$ system of embeddings, the full sharp and flat are mainly used in melodic alteration, while the half-sharp and half-flat are related to harmonic modulations. In particular, the traditional quintal circle will be constructed using half-sharp and half-flat (which will correspondingly appear on the score as the key signs).

For example, in Fig. 6 (upper), the ascending chromatic scale of C major is given in the 19-tone notation; it is obviously different from the descending major chromatic scale, as well as from the minor chromatic scales and the major-minor scale.
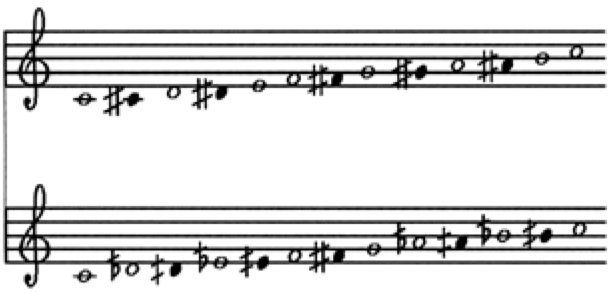
Figure 6. Two hypermodes.

An example of a hypermode with 12 tones that cannot be reproduced in the 12-tone scale is given in Fig. 6 (lower).

Conclusion

L. Avdeev's theory of pitch scales in music is capable of treating the scaling phenomena on any level of musical perception. In its static aspect, it describes all existing and possible in the future scales with regard to their musical value and expressive abilities. Also, it shows how ever more complex scales have been forming themselves with the growth in the number of partials comprehended in musical tones and the tightening of the perceptional tuning associated with the dispersion □ The stages of scale development are embodied in the hierarchy of possible scale embeddings, which is linked to modal and harmonic levels of music. The model has been elaborated in the framework of a general theory of hierarchies (Ivanov 2009).

Though the theory is based on the history of the European music, it can be applied, with minor changes, to musical systems that do not lean so much on scaling, such as oriental modes, ragas, or modern serial and aleatoric investigations. Any non-harmonic overtones can as easily be taken into account without complicating the mathematical formalism. We have thus presented a novel outlook at the acoustic foundations of music, ascending from the sensory level of the Helmholtzian physiological acoustics to the psychophysics of perception. Similar phenomena can be observed in the perception of the other aspects of music and in visual arts (Ivanov 1995). This makes hierarchical scaling one of the fundamental mechanism of aesthetic perception.

References

Avdeev L. V. & Ivanov P. B. (1993), A Mathematical Model of Scale Perception, *Journal of Moscow Physical Society*, **3**, 331-353

Avdeev L. V., Varivoda Y. I., Dubovik V. M. & Ivanov P. B. (2006), *The origin of the musical scale*, St. Petersburg: BODLib

Garbuzov N. A. (1948) *The Zone Nature of Pitch Hearing*, Moscow, Leningrad: USSR Academy of Sciences

Garbuzov N. A. (1956) *The Zone Nature of Timbre Hearing*, Moscow: Muzgiz

Gelfand S. A. (1981) *HEARING: An Introduction to Psychological and Physiological Acoustics*, New York, Basel: Marcel Dekker

Golitsyn G. A. (1980) *Information and the Laws of Aesthetic Perception*, Moscow: Znanie

Ivanov P. B. (1994), A Hierarchical Theory of Aesthetic Perception: Musical Scales, *Leonardo*, 27(5), 417-421

Ivanov P. B. (1995), A Hierarchical Theory of Aesthetic Perception: Scales in the Visual Arts, *Leonardo Music Journal*, 5, 49-55

Ivanov P. B. (2009), *Philosophy of Consciousness*, Victoria, BV, Canada: Trafford

Johnston B. (1966) *Persp. of New Music*, **5**, 112

Kvitka K. (1971) *Primitive Scales*, Moscow: Soviet Composer

Leontiev A. N. (1975) *Activity, Consciouness, Personality*, Moscow: Politizdat

Moles A. (1958) *Théorie de l'Information et Perception Esthétique*, Paris: Flammarion

Pozin N. V. (ed.) (1978) *Elements of the Theory of Biological Analyzers*, Moscow: Nauka

Pribram K. H. (1971) *Languages of the Brain*, Englewood Cliffs, NJ: Prentice-Hall

Schönberg A. (1973) *Persp. of New Music*, **11**, 3

---

http://unism.pjwb.net/arc
http://unism.pjwb.org/arc
http://unism.narod.ru/arc