24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

SEPARATION OF STRUCTURAL TIMBRE AS NARROW-BAND DIGITAL PROXIMAL SINUSOIDAL SPIKES

Ken ITO

*Div. Composition and Conducting, Interfaculty Initiative in Information Studies,
University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, 113-0032 Tokyo, JAPAN, e-mail: itosec@iii.u-tokyo.ac.jp*

Akinori TAKEI

NTT DATA IOMC .2-7-9 Hirakawa-cho Chiyoda-ku, 102-0093 Tokyo JAPAN

ABSTRACT

Human listening to band noise complex has extreme nonlinearity. A typical case is string ensemble or string sections within an orchestra. Within first violin section are 12-16 players and they play same voicing part in "unison". Looking at this event carefully, a unison part with numbers of violin players containing various vibrati consists not of pure tone at all, but of narrow band noise complex. However our human ear would catch clear sense of "pitch" tone height from such band noise.

String ensemble in "good" condition has kinds of "gloss" as quality.

With combining sinusoidal waves in proximal narrow band area we could separate such quality as "structural timbre" and continue investigating its brain cognitive nature from stand point of strict science of music. The distribution of hair cells in human cochlea has its spectral density. If we set 2 sinusoidal waves in 1 Hz difference --- e.g. 440Hz and 441Hz --- we would hear "beat" with a periodicity of 1 Hz. If 440Hz and 540Hz we would listen heterodyne differential resonance with 100Hz; this is audible frequency. Between 16-60 Hz is called Binding frequency band which has much to do with combining different human sensations.

Now we can design an inversion of such process within 1 Hz. When we put 3 sinusoid as 440Hz, 440.5Hz and 441Hz, these proximal sinusoids would interfere and we observe extreme values in wave form. When this frequency is in binding band area, this is neither heard as beats nor audible tone height but give special kind of sound quality. We determined this as structural timbre. Structural timbre could have much to do with human sense of space cognition from brain cognitive information processing.

1. Bright Line spectrum and Band Noise

Conventionally, it was a major problem in the sound study, how much ratio of the harmonic component obtained as a line spectrum should be. Analog synthesizers are using these setting parameters to make certain timbres, from this point of view. With the digitalization of audio processing techniques, timbre handling made more freely. For example, morphing to connect between the different timbres "naturally", e.g. from violin to clarinet, became the main topic. It is obvious that the emphasized point is on how to control the coefficients of harmonic spectrum, since those instruments with a clear harmonic spectrum are selected, e.g. from violin as an initial state to clarinet as a final state. Vocoding technology have played an exceptional role from this background. While analog vocoder is mimicking the voice by band noise processed by a filter varies largely in time, digital vocoder became possible to use a more diverse ways.

For organisms, at basement membrane on the cochlea tract within inner ear, hair cells are arranged densely as detector corresponding to sound waves of each frequency of the audible range. Then, when the hair cell reacts to the sound wave for a certain frequency, the impulse of neuronal firing is sent to the brain through the inner ear nerve. Thus, it can be examined individually which range of hair cells can react to the test signal by the band noise of a specific frequency range. Decomposition and recomposition with sinusoidal model of spoken language [1] are established on the basis of this principle, our group has also obtained the results using the same [2]. However, fundamental considerations connecting the voice and timbres, in terms of music, have been left consistently since the 1950s. The music framework dealing with them systematically have been sought.

2. From Mesoscopic system of solid state physics …Bridge between Discrete spectra and Continuum

One of the authors was inspired from solid-state physics, e.g. for semi-conductor LSI. Here we describe in short. For Microscopic level, a single atom has discrete energy level. On the contrary, for Macroscopic level, cristale has several energy bands and gaps. However for Mesoscopic level, few-body system has in-between characteristics from discrete to continuum, which are proximate and dense discrete energy levels.

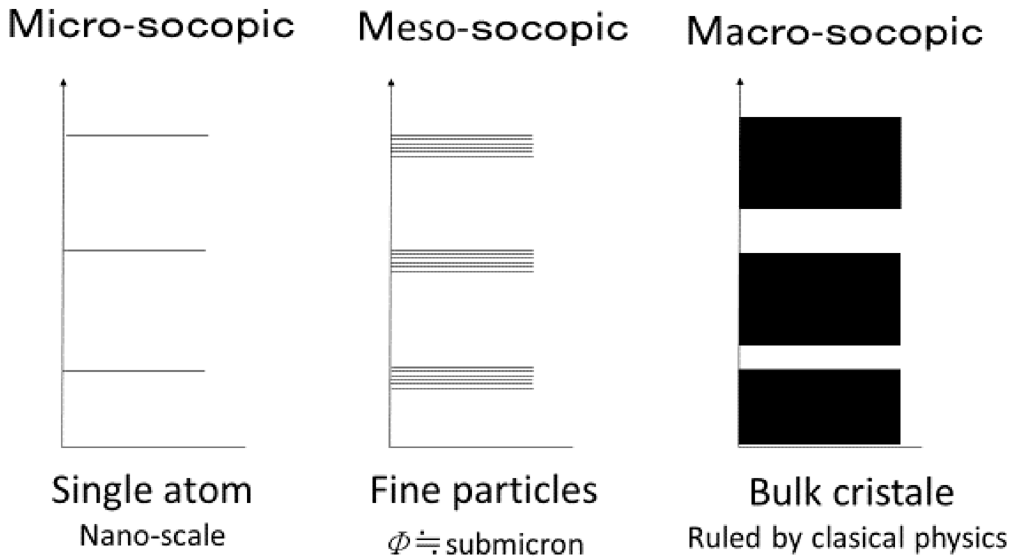

Figure 1. Schematic views of electron energy spectra in single atom (left) bulk cristale (right) and fine particles (middles)

3. Perception threshold and signal density

Human hearing is conducted frequency degradation by parallel processing in the cochlea duct, cochlear nerve is sending nerve firing impulse of upward to the brain as a signal of each frequency band. There appears difference depending on bands, depending on individuals also, assume that this resolution is Δ as model. If step-wise series of signals with shorter gaps than human cognitive resolution limit is less than Δ, human auditory cognition cannot recognize its discrete, and will listen as continuum (Fig. 2).

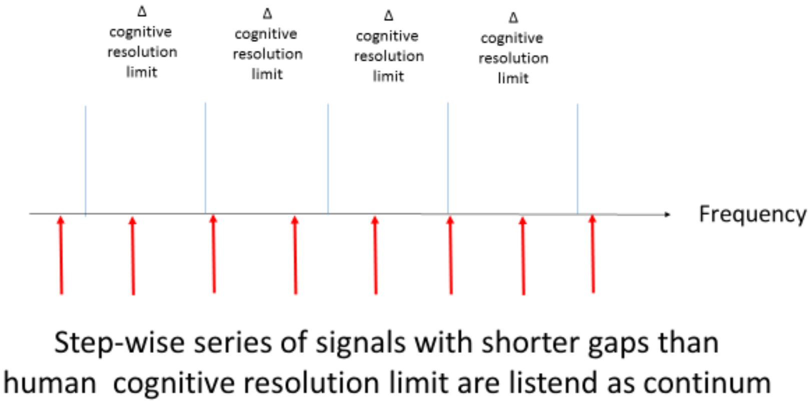

Figure. 2

Thus, we are able to make a series of appropriate test signals fundamentally, by selecting appropriate parameters, to make discrete digital spikes of signal but can be heard continuous enough.

4. Sinusoidal spike lattice and pitch-rhythm conversion

Let us think over a sine wave with frequency of 440 Hz.

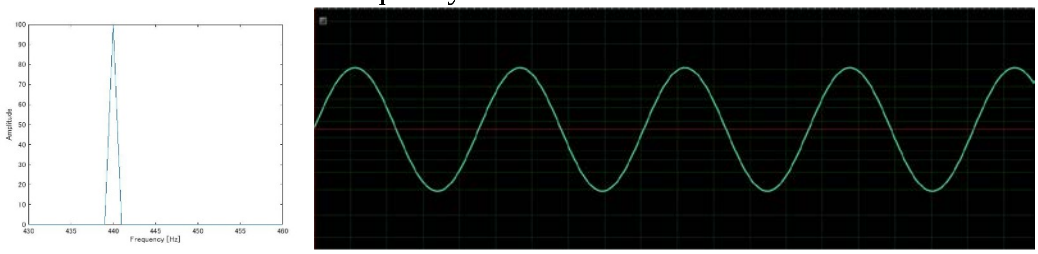

Figure 3. Sinusoidal spike with single bright line and its wave form(440Hz)

Then spectrum of this sound is shown above on left, and the wave form above on right. We hear a single tone with a pitch of La.

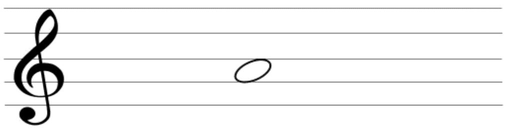

Figure 4.

The let us think about the second tone, a half tone lower in pitch from the former, namely sol #. This pitch would be given with the 15/16 of the first, i.e. 440×15/16 ＝ 412.5（Hz）.

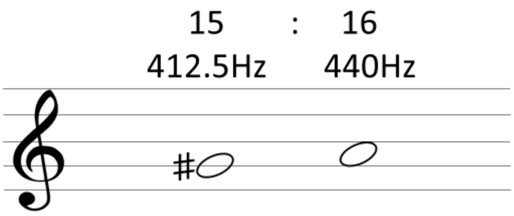

Figure 5-a.

With sound generation by computers, we certainly hear two pitches, but more prominent is BEAT caused by the difference of two frequencies.

Figure 5-b.

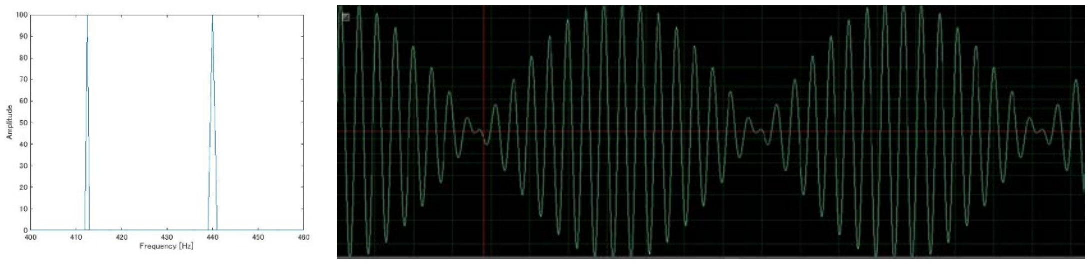

Figure 6. 2 Bright lines' spectral spikes and superposed sinusoidal wave form(440Hz、412.5Hz)

Similarly we obtain a pitch with major second lower than the original 440 Hz. It would be calculated with multiplying 7/8, i.e. 440×7/8 ＝ 385（Hz）, thus the beat would be heard with a frequency of 55Hz.

Figure 7.

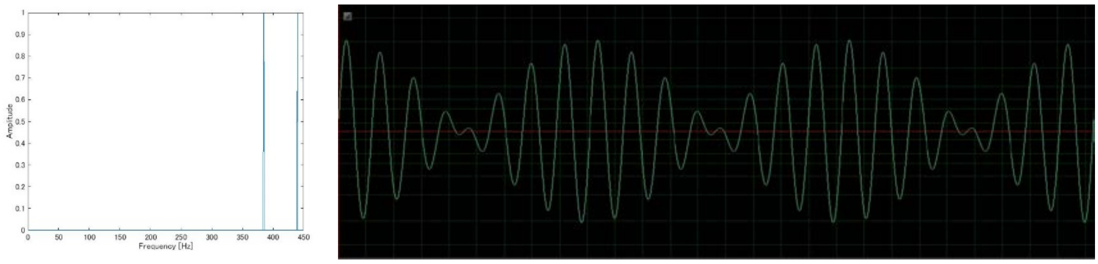

Figure 8. 2 Bright lines' spectral spikes and superposed sinusoidal wave form(440Hz、385Hz)

Now let us consider possibilities to fill the gap between the intervals above with sinusoid(s), the same amplitude and phase in intermediate frequenc-y(-ies).

If we divide the interval with the mean value, and let three sinusoids in 440, 412.5, and 385Hz ring altogether, we would listen an interesting result.

The intermediate frequency 412.5Hz would be hardly heard as pitch, but as beats interfering with the other two. Also the subjective sense of pitch of the upper tone declines a little; though very subtle, but for those who have an absolute pitch, this phenomenon is quite apparent.

We would hear the superposed synthetic sound as Major seconds with a little narrower interval and with much beat – flatter in --- round --- 27.5Hz.

Figure 9. 3 Bright lines' spectral spikes and superposed sinusoidal wave form (440Hz、412.5Hz and 385Hz). The subjective sound image is not at all linear because of the interference and it is difficult or impossible to segregate original three sinusoids.

Similarly, dividing the same interval in 4 and by adding five sinusoids with friequencies of 440, 426.25, 412.5, 398.75, and 385Hz, The upper pitch declines more deeper and we hear a multiple sound with rarer beat frequencies.

Figure 10. 5 Bright lines' spectral spikes and superposed sinusoidal wave form (440Hz、412.5Hz and 385Hz). The non-linearity of subjective sound image increases more.

If we promote this in such courses, with 8th division and 16th , we would hear synthesized multiple sound with beat attack of a frequency of common tolerance.

Figure 11. above 9 (above)/ 17(beneath) Bright lines' spectral spikes and superposed sinusoidal wave form (evenly divided between 440Hz and 385Hz). As the result of phase matching we hear interfering attacks with the frequency of common tolerance.

Figure 11. beneath

The 16^{th} even division between 440 and 385 Hz gives a common tolerance of 3.4375Hz.

If we fill the frequency gap with mostly evenly divided sinusoid with a gap of 3Hz, we would hear multiple sound with 3Hz. If common tolerance is 2Hz, beat would be also 2Hz and 1Hz, 1Hz respectively.

Thus, if we superpose many sinusoids in-phase with common gap of 0.1 Hz, we would hear interfering accents every 10 seconds.

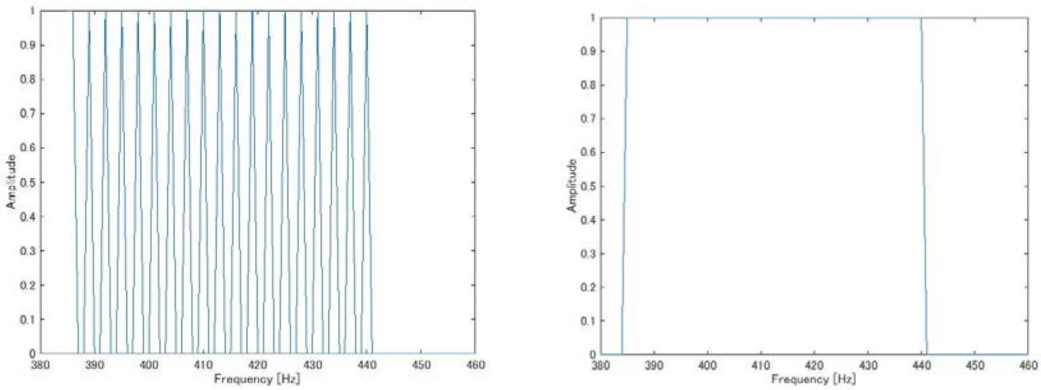

Figure 12.

In-phase sinusoids with the common tolerance frequency of 2Hz / 1Hz, n Hz would give regular rhythmic accents with 2Hz / 1Hz / n Hz. Note that here frequency gap gives not the pitches but rhythmic structure.

Now we would like to change our way of thinking from "division" to structure. If we form a lattice structure of sinusoids with common sequence tolerance of 0.1 Hz and downwards from 440Hz, what would our ears listen or find?

If we take a sample shorter than 10 seconds, we would get a synthetic sound with no periodical attacks and can observe the details of the sonority.

When we put sinusoids from 440 to 426.25 Hz, the interval for a quarter tone, with sequence tolerance of 0.1 Hz, we strangely hear a single pitch near round 440 Hz. This fact is widely acknowledged in auditory psychology but little known in musicology, and only we practical musicians of bowed string instruments know the minute details of such nature of tones as part of kraftmanship.

The superposition of narrow band sinusoidal spikes would give, even in-phase situation, kind of VIBRATI with the result of interference and gives rather steadfast sense of SINGLE PITCH. This comes from the cognitive structure of human inner ear, especially the distributions of hair cells on basilar membrane of cochlea. On this physiological principle we can device number of methods in vibrati for various instruments.

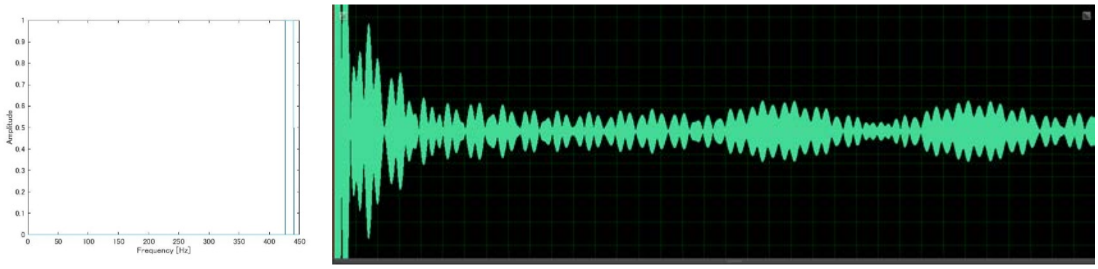

Figure 13. In-phase sinusoidal spikes with the common tolerance frequency of 0.1Hz between frequency area of 440 and 426.25 Hz. Note that the distribution of frequencies here gives not pitch divergence but rhythmic ones.

When we broaden the area of frequency further, from 440 to 412.5, an interval of a semitone, we would hear complex sound with pitches fluctuating mostly in two pitch levels --- upper end ant the lower --- and the quality of sound, if we may call this “timbre” is almost different from that of single sinusoidal wave.

Figure 14. In-phase sinusoidal spikes with the common tolerance frequency of 0.1Hz between frequency area of 440 and 412.5 Hz. The wound quality begins differ from that of single sinusoid; a congnitive phenomenon caused by human inner ears and central cognitive system.

And if we cover a frequency area wider than a half tone, e.g. from 440 Hz to 398.5, this gives an interval of 3/4 tone, the quality of sound is already completely different from that of simple sine waves.

This “timbre” is yielded not with harmonic superposition of overtones, but systematically added sinusoids and causes special sense of audible quality. Some says “depth”, some “clarity” and some “luster” or “sparkle”, but we can safely say sound qualities like this is quite often get by electric quitter amplifiers or analogue / digital effecter in popular music, and in the best traditions of classical and folk music all over the world.

5. ”Structural Timbre” and the structural colours

We call this modeled sound quality “structural timbre” and would investigate the physical and physiological nature of this in the context of music. We can introduce various common tolerance and uneven ones, with / without distribution of amplitudes and phases, in various pitch register and with various “band width” and “band gaps”. This naming is derived the case of super poly-chrome structure of colour cognition (structural colours)

Distinctive shine can be seen on the body of morphoes and jewel beetles. It is known that the "Structural Colour" can show the scatter response characteristic for not only the sunlight but also various incident lights.

Figure. 15 "Structural Colors" from a) Morphoes and b) Jewel beetles

The idea of semiconductor super lattice by Leo Ezaki. In comparison to this structural timbre we can call ordinal superposition of spectral harmonics as "pigment timbre" and differentiate the former from the latter.

Although little noticed ever, structural timbre components are universal and we can extract them from almost every fragment of music, regardless of genre, place and era of any civilization on the earth.

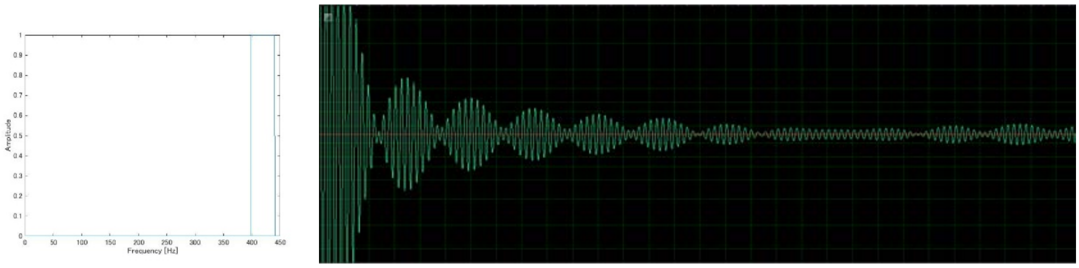

Figure 16. In-phase sinusoidal spikes with the common tolerance frequency of 0.1Hz between frequency area of 440 and 398.5 Hz i.e. 3/4-tone interval. Sinusoidal lattice structure with the simplest, even deviation, and the synthesized wave form of this sample structural timbre.

6. "Structural Timbre" as models for Tutti-vibrati of String instruments and or Chorus.

Structural timbre could be regarded as a model of vibrati in string instruments' performance or chorus. With this scope we are continuing changing various parameters of the spike lattice and evaluated the human hearing.

Also we are continuing brain blood flux measurement during the listening of structural timbre. Checking the activation area and its differences between pigment timbres and structural timbres, we would like to approach to investigate brain cognitive mechanism of such proximally condensed frequency deviation and its human evaluation within the whole scope of human perception for space and time.

REFERENCES

1 Takeyuki, HIDA, "White Noise: An Infinite Dimensional Calculus", Springer Science & Business Media, 1993