Measuring Dissonance: considerations for the microtonalist

Richard Duckworth
Department of Music
NUI Maynooth
Maynooth, Co Kildare
richard.duckworth@nuim.ie

Abstract
One area of microtonal research is concerned with digitally re-defining musical spectra so that bespoke microtonal scales can be paired so that new instruments capable of coherently performing exotic scales can be realised. In this field of timbre-matched microtonal scales, sensory dissonance perception models can provide invaluable information about the way that the human ear reacts to the spectra of musical instruments. This paper sets out to: survey the sensory dissonance models (Plomp and Levelt, Vassilakis, Hutchinson and Knopoff, Parncutt et al.), to assess the relative merits of these, and to demonstrate some tuning and tuning/timbre musical examples which have been implemented by the author in the Pure Data audio development package. The proof-of-concept examples of microtonal scale-timbre pairs that have been created in the software utilise the data from sensory dissonance models to permit the use of scales that would be unusable on conventional instruments.

1. Introduction

Considerable confusion is created by the fact that the terms used to describe phenomena and effects in the field of dissonance perception are also used to describe properties of intervals and chords in traditional harmony. This terminological woolliness is not aided by the plethora of models that the newcomer to this area must contend with. So initially, the nomenclature and the philosophies must be assessed. Notable commentator James Tenney examines the ways in which the terms consonance and dissonance have been used, and he describes the following types of dissonance: Melodic dissonance, Polyphonic dissonance, Contrapuntal dissonance, Functional dissonance and Psychoacoustic dissonance. [1] For the purposes of this paper, only the last category, psychoacoustic dissonance, will be considered, specifically sensory dissonance or roughness, and also distonalness. Vassilakis maintains that roughness is a dimension of the sensation of dissonance.[2]

2. Aspects of Sensory Dissonance

2.1 Roughness: the path from Helmholtz

Most listeners will be familiar with the sensorial effect of what Kameoka and Kuriyagawa call ‘turbidity’ - the rough, disruptive sensation caused by two tones close in pitch sounding simultaneously.[3] The fact that this sensation magically disappears and is replaced by a sensation of ‘clearness’ when the unison or the octave is achieved is the genesis of both the Periodicity and the Frequency Ratios theories. Theories based on frequency ratios state that intervals with small frequency ratios sound more consonant than those with large frequency ratios. Theories based on periodicity agree with this since they observe that intervals with small repeat periods sound more consonant. Since the repeat period for a small ratio is short, the two theories agree. The mechanism for the perception of complex tones is seen as an extension of the one used for simple tones and we will examine the implication of this below.

The Helmholtzian assumption is that the ear acts as a type of organic frequency or spectrum analyser - decomposing complex timbres into simpler components as they transmit through the cochlear windows and agitate the elements of the cochlear partition. One consequence of such a decomposition is creation of spectral data for processing by the neural system. In order to accomplish this task it needs to separate out the spectrum into divisions by using audio bandpass filters. These were identified by Fletcher in subjects by playing tones and then masking the tones with noise, sharply defined narrow-band noise - and given the name ‘critical bands’.[4] The width of the masking noise revealed the width of the critical bands. Each has a bandwidth, or a range of frequencies that it is responsible for processing, and a clearly defined boundary or a critical point, the traversal of which results in excitation of the adjacent critical band. So the term *critical band* refers both to a conceptual frequency range, and to a physical locus on the basilar membrane.

Experimental data reveals the frequency coordinates on the basilar membrane - high frequencies are processed near the basal portion of the structure (the end adjacent to the cochlear windows) and low frequencies are processed by the part near the apex of the cochlea. A frequency map of the basilar membrane been plotted - and Zwicker has assigned a 24-unit critical band scale covering the canonical 20Hz - 20kHz range of human hearing - the unit is named Bark and there are 24 Barks in the audio pitch range.[5] Tonotopic or ‘place’ theories of roughness perception rely upon this research.

The important finding for dissonance calculation is Plomp and Levelt's assertion that if two tones lie within the same critical band they are perceived as being dissonant. The term roughness is sometimes used so that confusion with the term dissonance as it is used in harmony is avoided. It might be more accurate to describe the sensation created by a dyad of pure tones or sine waves falling on the same critical band as 'roughness', and the feeling experienced when a dyad of complex tones falls on a critical band as 'dissonant'.

To return now to the perception of complex tones - in this model, a dyad of complex tones will have the respective timbres decomposed and then the ear will process the spectral data. If many of the partials of the two tones settle upon common critical bands, and they do not cohere, then dissonance will be experienced.

2.2 Resonator...or Travelling Wave?

In marked contrast to the critical band 'resonator' model of Helmholtz *et al*, albeit still a tonotopic theory, is the Travelling Wave model first proposed by Georg Békésy.[6] His precise experiments on animal and human cochleas from recently deceased subjects led to the creation of this model. He found that pitches channeled into the cochlea via the stapes created a wavelike crest that would travel down the length of the basilar membrane. Most importantly, this wave would peak at different loci for different frequencies - and Békésy postulated that it was at this peak point that the frequency was registered. Each position has a characteristic frequency and this was later encapsulated in a formula by Donald Greenwood.

Greenwood's Function:

$$F = A(10^{ax} - 1) \qquad (Eq. 1)$$

describes mathematically the 'place-frequency map' of the basilar membrane.[7]

The main criticism of this theory is that it does not explain the fine frequency discrimination of the ear - the fact that trained listeners can perceive pitch accurately to within a few cents. The physically broad profile of Békésy's wave shape is physically incapable of a narrow enough profile to permit such accurate pitch detection, though it has been proposed that the process may work in conjunction with a type of neuronal data filtering at the ganglial level.
A contemporary survey reveals that the cochlear partition is viewed as an active, bio-mechanical structure - the model of the cochlear partition as a series of tuned active resonators, as proposed by Gold and Pumphrey in 1948, would seem to prevail.[8] Recent research (Nilsen and Russell) reveals complex shearing forces at work in the basilar membrane and these, coupled with the interplay between the outer hair cells and the inner hair cells could explain the mechanics of the active amplification process at play on and below the surface of the Basilar membrane.[9]

2.3 Tonalness

The combination of data from research into the physiology of the cochlea with research of a psychological nature has resulted in a body of work which sees the neural apprehension, assessment, and processing of spectral data as only part of a larger picture which involves higher-level brain functions such as learning. This model, proposed by Terhardt, views the cochlear frequency-mapping device as a component of a larger perceptual apparatus.[10] Diagrammed in his landmark 1973 paper, 'Pitch, consonance, and harmony', as a neural-net type structure - the model includes provisions for cognitive feedback so that 'learned' patterns can affect the system.[11] This theory accounts for the influence of culture, and ties in *Gestalt* principles with the introduction of the concept of virtual pitch. The analysis and synthesis of tones is differentiated - Terhardt considers that the pitch assigned to a complex tone by a listener is a virtual one and not only is this pitch a construct of the perceptual system, it is also a product of learning.
This model and its extensions (Parncutt, MacCallum) have the potential to yield the most useful algorithms for the microtonalist due to the inclusion of provisions for inharmonic scales and compensation for masking.

3. Measuring Roughness

3.1 Models and metrics

Models for dissonance curve calculation begin with Plomp and Levelt's contour for two pure tones. This algorithm has parameters for the frequencies of the two waves and the respective amplitudes and from these inputs a 'dissonance' figure can be returned for any given interval. Here is the original Plomp and Levelt dissonance curve:[12]

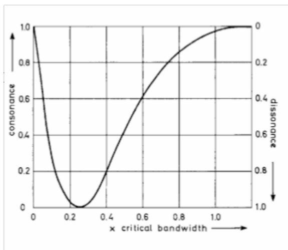

Fig 1. Roughness Curve for two pure sine tones

Essentially they tie the dissonance maximum to the critical bandwidth of the ear. This basic dissonance calculator has been extended to include complex timbres. Spectral decomposition is necessary as each individual sinusoid in the complex timbre must have a separate dissonance calculation carried out on it w.r.t. the other components. The sum of all of these operations yields the total dissonance curve for the two timbres. When implementing routines for describing these curves, the critical band factor, derived from experimental data, is most important. Also, the descendants of Plomp and Levelt's curve calculators are more appropriate in that they take into account partials of different amplitudes. Here is Sethares' parameterisation of Plomp and Levelt's curve:[13]

$$d(x) = e^{-3.5x} - e^{-5.75x} \qquad (Eq. 2)$$

3.2 Additive nature of dissonance

Terhardt attributes the sensation of roughness directly to temporal fluctuations in amplitude when they appear within the 'spectral regions' known as critical bands (CBs). His theory takes account of the additive nature of roughness perception: the roughness created across all of the CBs is summed to give an overall roughness effect. This confirms the assumption that, for a pair of timbres, adjacent partials of each timbre that lie within the same CB will contribute to the overall sensation of roughness. He also posits that roughness is greatest for spectrally coherent sounds lying within the same CB, and that the roughness effect can be ameliorated by randomising the amplitudes and phases of the components (he points out that reverberation can provide a randomising effect).

Terhardt's ideas were extended by Parncutt, who applied them to common practice music and to 12-tone Equal Tempered music. He also implemented an extended version of Hutchinson and Knopoff's roughness algorithm, represented by this equation:

$$d(x) = \left(Axe^{\frac{x}{B}}\right)^2 \qquad (Eq. 3)$$

3.3 Inclusion of roughness

It could be said that the emancipation of roughness became complete with the publication of Vasilakis' pan-musical interdisciplinary dissertation in 2001.[14] He makes the case that roughness, often suppressed by the builders of instruments designed to play western 12-tone Equal Tempered musical material, is an integral sonic attribute of the sound of many non-western musical instruments. The confusion surrounding amplitude fluctuation and amplitude modulation depth is clarified, and as well as this a compensating equation is proposed which corrects Eq. 2 by accounting for amplitude fluctuations:

$$Amp(A_1 A_2) = (A_1 A_2)^{k1} K_2 \left(\frac{2A_1}{(A_1 + A_2)}\right)^{k3} \qquad (Eq. 4)$$

4. Implementations in pure data

4.1 Bohlen-Pierce scale in Pd

The Bohlen-Pierce (B-P) scale is a non-octave scale with 13 steps and it repeats every 12th (an octave + a fifth). This 3:1 repetition interval is called the tritave. The scale was independently proposed by Heinz Bohlen in 1978, and by John R. Pierce in 1984.[15 16] Bohlen's original B-P scale was a Just Intonation (JI) scale with the following ratios:

1/1, 27/25, 25/21, 9/7, 7/5, 75/49, 5/3, 9/5, 49/25, 15/7, 7/3, 63/25, 25/9, 3/1.

J. R. Pierce's scale is a 13-tone Equal Tempered scale where:

$$one\ step \ = \ \sqrt[3]{13}$$

The scale is based on the 3:5:7:9 tetrachord and consists of 13 steps in the chromatic version.

4.2 Implementation of the B-P scale in Pd

The B-P scale was implemented in Pd and controlled by a repurposed MIDI keyboard. The keyboard provides a 'front end' for an additive synthesis instrument implemented in Pure Data. The current version of the instrument contains a polyphonic voice engine and additive synthesis voices consisting of twenty partials. The partials have high resolution frequency and amplitude controls, and there are individual amplitude envelopes on each partial.

The B-P scale does not sit well on the physical key arrangement of the 12-TET keyboard, so, following from Elaine Walker (and from Heinz Bohlen), a pair of standard Roland PC-70 MIDI keyboard controllers were fashioned into a single controller with a key layout that complements the popular 'Lambda' B-P mode. In this implementation the base pitch is situated on the key formerly called low C, and the white keys are assigned the diatonic B-P scale ascending:

1/1, 25/21, 9/7, 7/5, 5/3, 9/5, 15/7, 7/3, 25/9, 3/1

The ratios corresponding to the chromatic notes are mapped to the reconfigured black keys. The short white keys are not used for note values but may be used for controller messages and used to switch timbres or actuate other processes. The messages are accepted into Pd using a standard MIDI port, however, only parts of the MIDI protocol are used as intended by the creators. The Velocity message is used to control dynamic, but the frequency values ascribed to the MIDI note numbers are not used. Instead, each unique MIDI note ID is re-assigned to trigger scale maps created within the Pd patch itself.

Fig 2. Repurposed MIDI keyboard suitable for B-P Lambda mode

The Pd patch allows the user to select either 13-TET B-P or JI B-P scales. The 13 TET scale is derived by using the formula:

$$\text{scale step } n \;=\; 3^{\frac{n}{13}}$$

and this is coded into an expression in Pd. The B-P JI scales are calculated by sequentially routing the B-P ratios through an expression, and storing the results of the calculations for each ratio in an array - this then is used to create the scale map.

4.3 Spectra related to the B-P scale

Since all of the B-P scale intervals are based on frequency ratios derived from odd integers, it follows that the B-P scale has maximum relatedness to a timbre possessing only the odd-numbered partials. It is true that a B-P scale will perform admirably with a square wave or a triangle wave timbre. There has, however, been a substantial amount of success with timbres possessing both even and odd partials - and so an opportunity for B-P spectrum-scale matching became apparent. The principle of coinciding partials is manipulated via a method described by Sethares: the symbolic computation of spectra.[17] When considering roughness curves for spectra, it is found that many of the points of minimum roughness can be found at points where partials coincide, and this effect is called the principle of coinciding partials. This is encapsulated by the formula[18]:

$$\text{ratio } r = \frac{f(i)}{f(j)} \;\; \text{where } f(i) \text{ and } f(j) \text{ are partials of } F$$

By using this principle, a spectrum related to the B-P scale can be created in which the ratios between the spectral components are equal to intervals between the scale steps. To simplify this fairly unwieldy process, Sethares uses an organisational device called an O-plus table to codify all of the permissible intervals, i.e., those that result in legal scale steps. Fig. 2 shows how Sethares method was adapted for use with the non-octave B-P scale:

	1/1	27/25	25/21	9/7	7/5	75/49	5/3	9/5	49/25	15/7	7/3	63/25	25/9	3/1
⊕	**(0,0,0,0)**	**(1,0,0,0)**	**(1,1,0,0)**	**(2,1,0,0)**	**(2,1,1,0)**	**(2,1,1,1)**	**(2,1,2,1)**	**(3,1,2,1)**	**(3,1,3,1)**	**(3,1,3,2)**	**(3,1,4,2)**	**(4,1,4,2)**	**(4,2,4,2)**	**(5,2,4,2)**
(0,0,0,0)	(0,0,0,0)	(1,0,0,0)	(1,1,0,0)	(2,1, 0,0)	(2,1,1,0)	(2,1,1,1)	(2,1,2,1)	(3,1,2,1)	(3,1,3,1)	(3,1,3,2)	(3,1,4,2)	(4,1,4,2)	(4,2,4,2)	(5,2,4,2)
(1,0,0,0)	(1,0,0,0)	*	(2,1, 0,0)	*	*	*	(3,1,2,1)	*	*	*	(4,1,4,2)	*	(5,2,4,2)	(1,0,0,0)
(1,1,0,0)	(1,1,0,0)	(2,1,0,0)	*	*	*	*	*	*	*	*	(4,2,4,2)	(5,2,4,2)	*	(1,1,0,0)
(2,1,0,0)	(2,1,0,0)	*	*	*	*	*	*	*	*	*	(5,2,4,2)	(1,0,0,0)	(1,1,0,0)	(2,1,0,0)
(2,1,1,0)	(2,1,1,0)	*	*	*	*	*	*	*	*	(5,2,4,2)	*	*	*	(2,1,1,0)
(2,1,1,1)	(2,1,1,1)	*	*	*	*	*	*	*	(5,2,4,2)	*	*	*	*	(2,1,1,1)
(2,1,2,1)	(2,1,2,1)	(3,1,2,1)	*	*	*	*	(4,2,4,2)	(5,2,4,2)	*	*	*	*	*	(2,1,2,1)
(3,1,2,1)	(3,1,2,1)	*	*	*	*	*	(5,2,4,2)	(1,0,0,0)	*	*	*	*	(2,1,2,1)	(3,1,2,1)
(3,1,3,1)	(3,1,3,1)	*	*	*	*	(5,2,4,2)	*	*	*	*	*	*	*	(3,1,3,1)
(3,1,3,2)	(3,1,3,2)	*	*	*	(5,2,4,2)	*	*	*	*	*	*	*	*	(3,1,3,2)
(3,1,4,2)	(3,1,4,2)	(4,1,4,2)	(4,2,4,2)	(5,2,4,2)	*	*	*	*	*	*	*	*	*	(3,1,4,2)
(4,1,4,2)	(4,1,4,2)	*	(5,2,4,2)	(1,0,0,0)	*	*	*	*	*	*	*	*	(3,1,4,2)	(4,1,4,2)
(4,2,4,2)	(4,2,4,2)	(5,2,4,2)	*	(1,1,0,0)	*	*	*	(2,1,2,1)	*	*	*	(3,1,4,2)	*	(4,2,4,2)
(5,2,4,2)	(5,2,4,2)	(1,0,0,0)	(1,1,0,0)	(2,1, 0,0)	(2,1,1,0)	(2,1,1,1)	(2,1,2,1)	(3,1,2,1)	(3,1,3,1)	(3,1,3,2)	(3,1,4,2)	(4,1,4,2)	(4,2,4,2)	(5,2,4,2)

Fig 2. O-plus table showing legal intervals for B-P scale. Symbol * denotes result that is not a scale interval.

The presence of a large number of * symbols foretells that the choice of partials will be quite restricted. However, there is always the option to space conflicting partials widely, thereby essentially nullifying the roughness effect. Now this table is used to construct the symbolic spectrum table:

i	1	2	3	4	5	6	7	8	k
t(i)	(0,0,0,0)	(5,2,4,2)	(9,4,8,4)	(12,5,8,4)	(13,5,10,5)	(13,5,12,6)	(14,5,12,6)		
s(i)	(0,0,0,0)	(0,0,0,0)	(4,2,4,2)	(2,1,0,0)	(3,1,2,1)	(3,1,4,2)	(4.1.4.2)		
r (i,k)		(5,2,4,2)	(4,2,4,2)	*	*	*	(1,0,0,0)		**1**
			(4,2,4,2)	(2,1,0,0)	*	*	*		**2**
				(2,1,0,0)	(3,1,2,1)	(4,1,4,2)	*		**3**
					(3,1,2,1)	(3,1,4,2)	*		**4**
						(3,1,4,2)	(4.1.4.2)		**5**
							(4.1.4.2)		**6**

Fig 3. Symbolic spectrum table derived from O-table in Fig. 2.

This table gives us our partial frequency values expressed as exponents of JI ratios. The four intervals present in the B-P scale are:

27/25, 625/527, 49/45, 375/343,

so, by simply carrying out the JI math, a set of partials suitable for the B-P scale can be found. This particular spectrum is not perfect, note the presence of * symbols. The effect of the scale is intriguing: bittersweet yet with enough coherence to realise B-P tonalities. The partial frequencies derived from this process were imported into Pd as octave-adjusted JI ratios by applying the corresponding exponential values, then they were converted to floating point values and sent to the precision partial frequency controls. The amplitudes were initially set to amplitudes 1/n where n = partial number, and then adjusted by ear.

References

[1] Tenney, James. *A History of Consonance and Dissonance*. (New York: Excelsior, 1988).

[2] Pantelis Vassilakis, 'Perceptual and Physical Properties of Amplitude Fluctuation and their Musical Significance', Doctoral dissertation. Los Angeles: University of California, Los Angeles. (2001)

[3] Akio Kameoka and Mamoru Kuriyagawa, 'Consonance Theory Part I: Consonance of Dyads', J. Acoust. Soc. Am. 45, (1969): 1452

[4] Fletcher, Harvey, 'Auditory Patterns', Rev. Mod. Phys., 12, (Jan, 1940), 47-65

[5] Eberhard Zwicker, 'Subdivision of the Audible Frequency Range into Critical Bands', J. Acoust. Soc. Am. 33, (February 1961), 248

[6] Georg Békésy, 'The Variation of Phase Along the Basilar Membrane with Sinusoidal Vibrations', J. Acoust. Soc. Am. 19, (May, 1947), 452

[7] Donald Greenwood, 'Auditory Masking and the Critical Band', J. Acoust. Soc. Am. 33, (April 1961), 484

[8] T. Gold and R.J. Pumphrey, 'Hearing. I. The Cochlea as a Frequency Analyzer', Proceedings of the Royal Society of London, 135, No. 881, (Dec, 1948), 462-491

[9] Nilsen, K.E., Russell, I.J., 'The timing of cochlear feedback: Spatial and temporal representation of a tone across the width of the guinea pig basilar membrane', (1999), Nature Neuroscience 2, 642 - 648.

[10] Ernst Terhardt, 'Pitch, consonance, and harmony', J. Acoust. Soc. Am. 55, (May 1974), 1061-1069

[11] Ernst Terhardt, 'Pitch, consonance, and harmony', J. Acoust. Soc. Am. 55, (May 1974)

[12] R. Plomp and W. Levelt, 'Tonal Consonance and Critical Bandwidth', J. Acoust. Soc. Am. 38, April 1965), 548-560

[13] William A. Sethares, Tuning, Timbre, Spectrum, Scale, 2nd edn (London: Springer-Verlag 2005), page 345

[14] Pantelis Vassilakis, 'Perceptual and Physical Properties of Amplitude Fluctuation and their Musical Significance', Doctoral dissertation. Los Angeles: University of California, Los Angeles. (2001)

[15] H. Bohlen, '13 Tonstufen in der Duodezeme', **39**, 67-86 (1978).

[16] M.V. Matthews, L.A. Roberts, and J.R. Pierce, 'Four new scales based on nonsucessive-integer-ratio chords', J. Acoust. Soc. Am. Suppl. 1 75, S10 (1984)

[17] William A. Sethares, Tuning, Timbre, Spectrum, Scale, 2nd edn (London: Springer-Verlag 2005), page 257

[18] William A. Sethares, Tuning, Timbre, Spectrum, Scale, 2nd edn (London: Springer-Verlag 2005), page 123