Low Energy and Equal Spacing
The multifactorial evolution of tuning systems

Georg Hajdu

Center of New Music and Audio Technologies (CNMAT)
University of California
1750 Arch Street, Berkeley 94709
georg@cnmat.berkeley.edu

Abstract

Studies of tuning systems established in different historical and ethnic cultures suggest that their evolution is mainly dependent on two different, mathematically incompatible, principles: **equidistance**, based on a mostly logarithmic perception of the audible frequency range, and **harmonic consonance** generated by simple integer pitch ratios (SIPR). Since for tuning systems the number of relevant SIPRs was overestimated in older theories (often mixing a scientific approach with irrational numerology), a new perceptual model based on both, harmonic consonance and pitch strength, the latter is closely related to **timbre**, is presented here. A method for numeric evaluation of the harmonic consonance of any interval in pitch space allows a mathematical and graphical representation of **harmonic energy**. The quality and stability of tuning systems thus can be calculated by combining the energy values for each element of their interval set. The combinatorial energy for tempered systems is plotted in order to determine the system with the lowest value, yielding different systems for different pitch strengths. The high pitch-strength plots are dominated by the most common system, 12-tone equal temperament, whereas the low pitch-strength plots overwhelmingly give priority to the equal tempered pentatonic system - a finding which suggests a strong relationship between the historical evolution of tuning systems and a culturally determined focus on different musical parameter like melody or rhythm.

I. **Introduction**

In order to establish a model for the evolution of tuning systems it is necessary to analyze the main physical and psychophysical factors which may have contributed, and still contribute, to their formation[1]. Different factors such as the following have been discussed in music theory:

1. **Simple integer pitch ratios** of complex tones (intervals contained in harmonic spectra).

This factor is known since the classical Greek period (Pythagoras) and has often been reconsidered by mathematicians and theorists/composers. The basic psychological term connected with simple integer pitch ratios (SIPRs) is *pleasantness* (or more recently *harmonic consonance*)[2]. Numerous mathematical models have tried to correlate the degree of pleasantness to the nature of a given SIPR.

2. **Logarithmic perception of pitch space** (pitch detection on the basilar membrane).

Weber and Fechner have described a logarithmic relationship between sensory stimuli and their perception by the sensory apparatus. Von Békésy (1960) has been able to match resonance maxima on the basilar membrane with perceived pitches. The monotonous curve he obtained approximates a logarithmic curve.

3. **Octave similarity/identity** (helical structure of pitch space).

This aspect of pitch perception (especially of complex tones) is captured by R. Shepard's helical model of pitch perception/mental representation (1964). Pitches separated by one or more octaves are judged as most similar and therefore located on the same side of the helix.

4. **Psychophysical consonance and dissonance** (coinciding partials).

Sinosoids slightly out of tune produce beating patterns which - in a certain range - are perceived as roughness. Thus, simultaneous compound tones display more or less complex beating patterns according to the amount of coinciding partials. As the computational model by Plomp and Levelt (1965) suggests, SIPRs of compound tones are characterized by reduced roughness and low unpleasantness values.

5. **Difference tones.**

Hindemith (1940) based his theory of tonality/musical intervals on the fact that when two pitches are played a third one can be heard. In case of SIPRs the difference tone often forms the fundamental pitch of the perceived three tone complex.

Studies of tuning systems in different historical and ethnic cultures propose that mainly the first two factors (the third being a special case of the first) are essential for their evolution (von Hornbostel, 1905-06)[3]. The two notions derived from these two factors and used throughout this paper are **harmonic consonance** and **equidistance**. Both principles are mathematically incompatible since a set of intervals derived through equal geometrical division of an SIPR (mostly the octave 2/1) never matches exactly the frequency ratio of any large integers[4].

Fig. 1 shows a scheme for the coevolution of both principles which brought forth the standard well-tempered system. If one had to represent different historical or ethnic systems as areas within the triangle one could locate Indonesian slendro on the right, and the drone-based Indian rags on the left. The transition from an equidistant system to a harmonically pure one is called **rationalization**, whereas the opposite transition is defined as **tempering** (Barlow, 1987).

II. **Simple Integer Pitch Ratios**

It is a fact of common sense that certain intervals have greater stability than other ones. As Hall & Hess (1984; see also Rasch, 1985 and Vos, 1988) were able to demonstrate in a series of tuning experiments, subjects with musical training are quite aware of slight deviations from SIPRs. Inquiring about the degree of pleasantness, they found V and U shaped curves in the vicinity of different SIPRs and assigned these differences in shape to different degrees of stability.

Euler, the famous Swiss mathematician of the 18th century, (Busch, 1970) tried to quantize an SIPR's degree of stability *(gradus suavitatis)* and found a correlation between stability and two mathematical properties of the simplified pitch ratios:

1. Size of involved prime numbers.
The smaller the prime factors of the numbers in the given ratio the more harmonic the ratio.
2. Number of exponents.
The more divisible the numbers in the given ratio the more harmonic the ratio.

Euler's Gradus function yields values which corresponds roughly to the experience of trained Western listeners. Nevertheless, it is questionable whether or not the ratios of 6/5, 7/6 or 9/8 have the same stability or pleasantness.

Barlow (1981,1987) took a similar approach. Based on the concept of mental representation of numbers ("indigestibilty"; developed independently by Shepard et al (1975)), he found a formula yielding harmonic-consonance values ("harmonicity") for given ratios of integers. In addition to Euler the polarity (root position) of the interval was also taken into consideration [5]:

Indigestibility:

$$\xi(N) = 2\sum_{r=1}^{\infty}\left\{\frac{n_r(p_r - 1)^2}{p_r}\right\} \quad \text{where} \quad N = \prod_{r=1}^{\infty} p_r^{nr},\ p \text{ is a prime, and } n \text{ is a natural number.} \qquad (1)$$

Harmonic Consonance ("Harmonicity"):

$$h(P,Q) = \frac{sgn\left[\xi(P) - \xi(Q)\right]}{\xi(P) + \xi(Q) - 2\xi\left(hcf_{P,Q}\right)} \qquad (2)$$

where $sgn(x) = -1$ when x is negative, otherwise $sgn(x) = +1$,

$hcf_{a,b}$ is the highest common factor of a and b, and $\xi(x)$ is indigestibility of x.

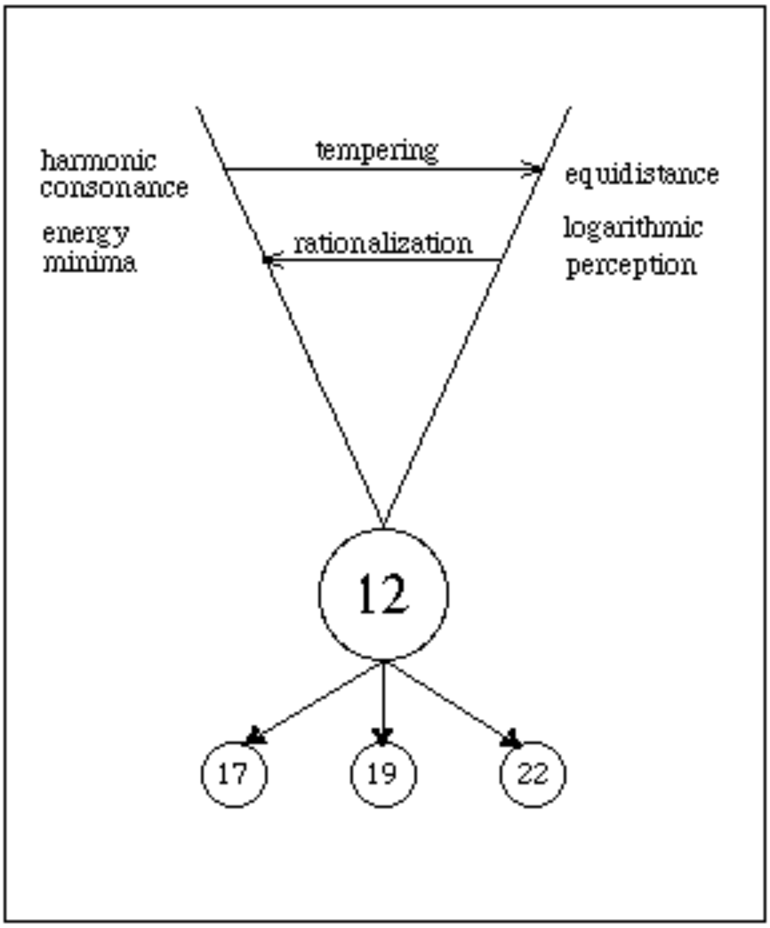


Fig. 1. Evolutionary triangle for the development of Western standard tuning. The systems with 17, 19 or 22 steps are mentioned as examples for a further, pluralistic development of tuning systems. For the future composer, choosing a tuning system will be as much a precompositional decision as selecting a set of instruments (Hajdu, 1990).

Both the plausibility of his theory and its successful application in his algorithmic compositions suggested the use of his formulae in the further quantitative study of musical intervals and systems. For this, the harmonic consonance values for approx. 2000 intervals (between 1/1 and 256/1, corresponding to 0-9600ct) have been calculated and plotted in Fig. 3. (The author assumes that—although *cents* is a relative measure—0ct can only be anchored within a limited frequency range in bass register.)

As one can easily see, the differences of the harmonic consonance values for the salient intervals decrease in the upper register - a finding which is consistent with predictions of music theory. Of greater interest is the fact that certain intervals belonging to the same interval class (e.g. 5/4, 5/2, 5/1, 10/1) reach their highest values in different octave registers[6].

Investigations of harmonic consonance values within ranges other than the first octave (Fig. 4) unveil a striking correspondence between the envelope formed by the strongest intervals of the third and fourth octave (the range known to contain the melodically important pitches of the overtone series) and the graph displaying tonal hierarchies in the context of the major scale (Krumhansl & Shepard, 1979). This finding suggests the existence of a virtual fundamental phenomenon also for horizontal pitch perception (melody) analogous to the virtual pitch phenomenon described by Terhardt et al. (1982) for simultaneous pitches (harmony). It seems likely that the listener's mind parses melodic events by referring to both a virtual fundamental located in the low register, and to the linear intervallic progressions.

Note that Barlow's formulae are also based on personal judgments about the "indigestibility" of numbers. It is questionable whether number 9 is in fact more digestible than number 5. In the case of the the major second, Barlow's assumption may cause the noticeable difference between the predictions by his model and the findings by the probe tone method. As far as the computational model is concerned, the harmonic consonance values only represent one possible quantitative input, i.e. any kind of quantitative estimate about the pleasantness/harmonic consonance/relatedness-of-musical-intervals could be used and tested in the model.

N	ξ(N)
1	0.000000
2	1.000000
3	2.666667
4	2.000000
5	6.400000
6	3.666667
7	10.285714
8	3.000000
9	5.333333
10	7.400000
11	18.181818
12	4.666667
13	22.153846
14	11.285714
15	9.066667
16	4.000000

Fig. 2. Indigestibilty values for the integers 1 to 16.

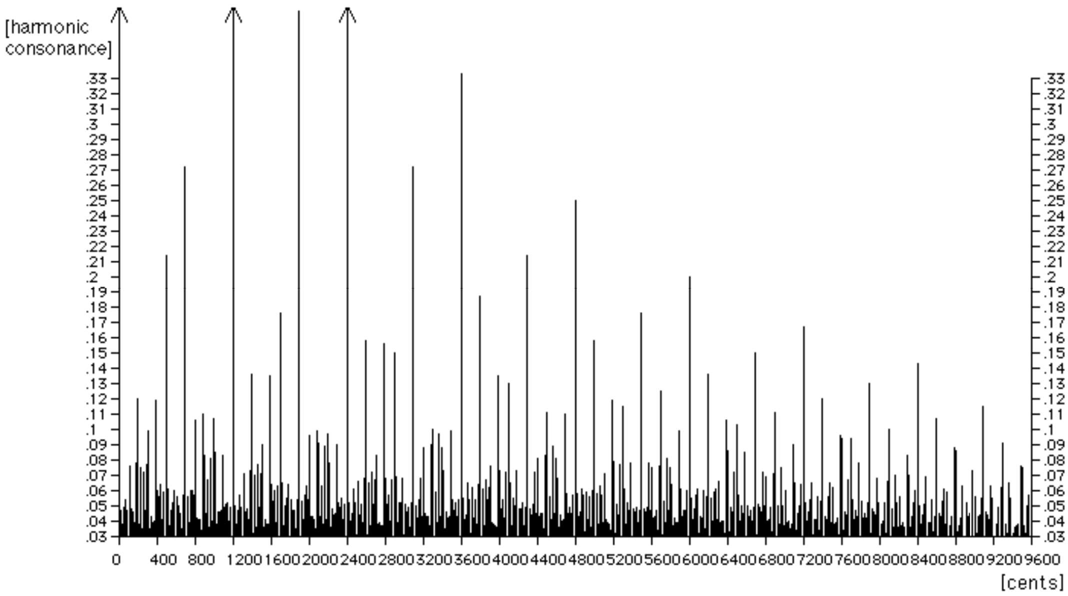


Fig. 3. Harmonic consonance for selected intervals between 0 - 9600ct (absolute values).

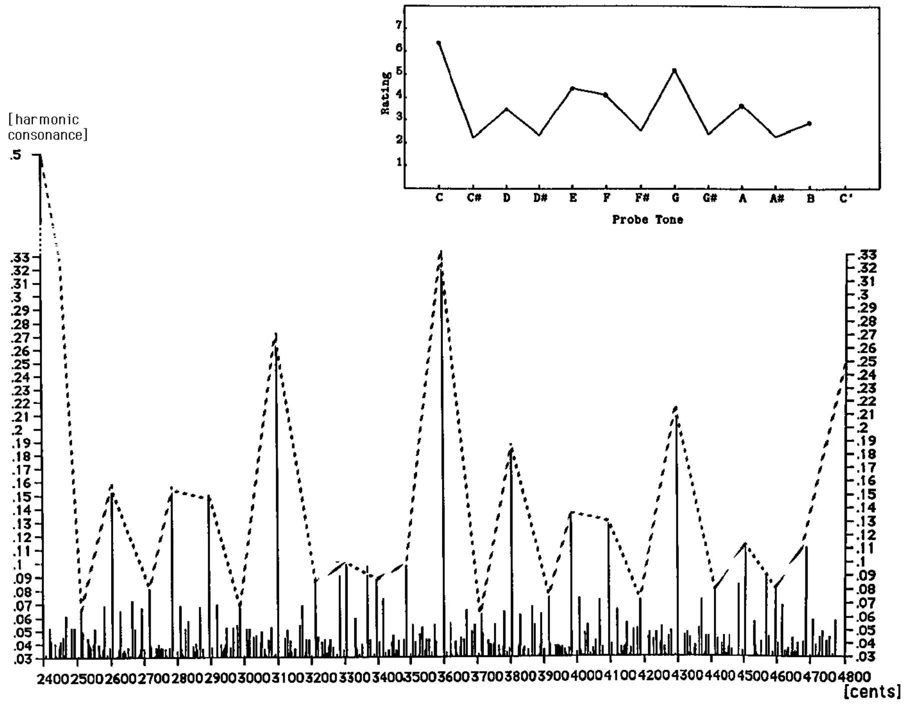


Fig. 4. Harmonic consonance values for important intervals between 2400 - 4800ct compared with probe tone ratings in major context. The salient intervals are connected by dotted lines (exception: 7/1, 14/1 are supposedly too far from tempered halfsteps or too close to intervallic categorical boundaries, and therefore rejected in the course of music history. See also Hall & Hess (1984)).

III. **Harmonic energy**

Detuning experiments show that SIPRs are not perceived as single spots in pitch space. Like resonances they tend to be extended over a certain range. Since both extension and stability are directly dependent upon the pitch ratio itself, an analogy to physical processes can be drawn by introducing the term *harmonic energy* as a measure for intervallic stability (Hajdu, 1989)[7]. For this, bell shaped Gaussian distribution curves are superimposed on the central harmonic consonance values of SIPRs, yielding decreasing harmonic consonance values for adjacent intervals. Since there are, theoretically, infinite values for a given point in pitch space (the extension of every single Gaussian curve is infinite) only the maximum values for all selected intervals are collected.

According to Fricke (1973), the extension of an interval depends on the musical context in which this interval is presented. But also the listener is far from being an ideal black box which receives and processes musical stimuli in a consistent and controllable way. Pitch perception in general is influenced by the internal (physiological and psychological) and external (physical) factors listed below:

Internal factors:	External factors:
- Cultural background	- Timbre[8]
- Training	- Register
- Attention	- Loudness
- Ability	- Duration

These factors can be subsumed under the notion of *pitch strength*, which thus serves as a measure for how strong a pitch sensation is evoked in the listener. A general relationship between pitch strength and harmonic consonance can be formulated:

The clearer the pitch percept the narrower the spatial extension (pitch width) of an SIPR and the lower its energy.

The hypothetical correlation between harmonic energy, harmonic consonance, pitch strength, and pitch width is given by (3)[9].

$$E(I) = - S \log(|H(I_0)|\, e^{-(I - I_0)^2 / W^2}) \qquad \textbf{(3)}$$

$E(I)$	= Harmonic Energy of an interval
$\lvert H(I_0)\rvert$	= Harmonic consonance of interval I_0 (absolute value)
I_0	= $1200 \log(F_1/F_2)/\log(2)$
W	= Pitch Width
S	= Pitch Strength
	= $21.5ct\ (1 / \sqrt{W})$ (21.5ct = syntonic comma)

The harmonic energy distribution curve has been calculated for a mean pitch strength value (W=28ct) and plotted for a selected range in Fig. 5. The graph suggests that only a limited amount of SIPRs actually have a psychoacoustical significance. Most of the SIPRs entered in the calculation are covered by more salient intervals. Thus, the consideration of complex ratios in numerous treatises (both older or more recent ones, e.g. Partch 1949) is questionable, assuming that the model presented here better captures the realities of pitch perception. The arrows correspond to local energy maxima. Their coincidence with categorical boundaries is striking, thereby creating an open field for further speculation.

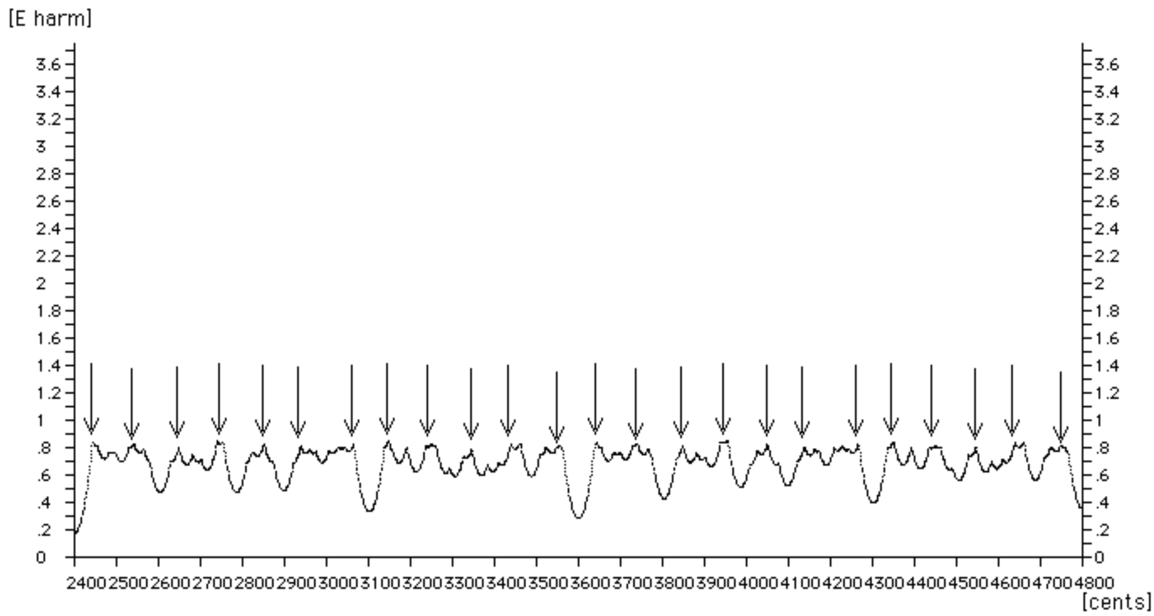
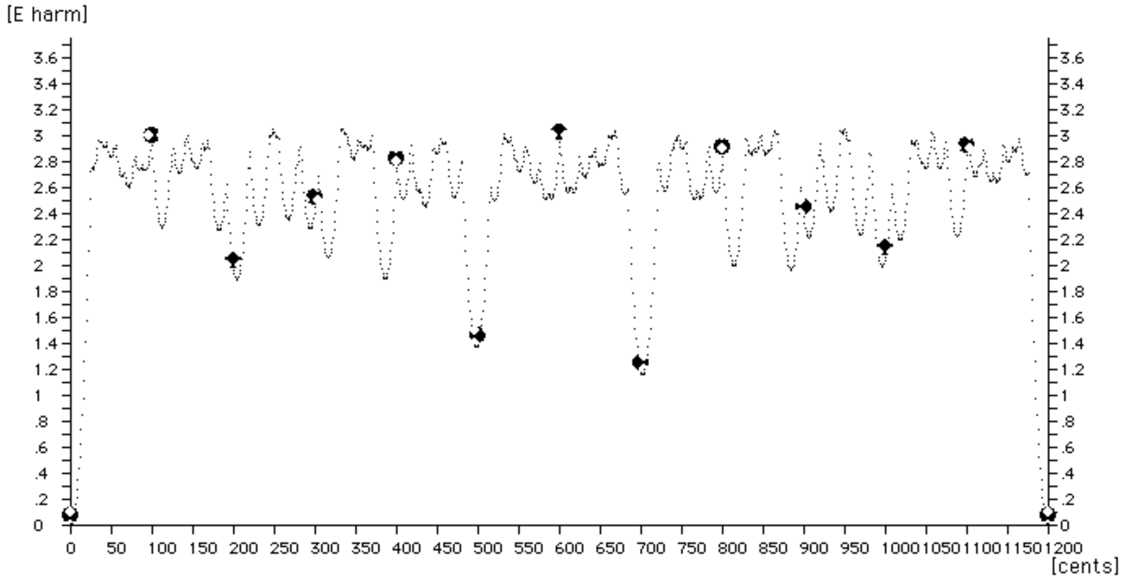


Fig. 5. Harmonic energy distribution for the third and fourth octave apart from a given fundamental (W=28ct). The arrows correspond to absolute energy maxima within windows of 70-120ct in size. They coincide with boundaries of categorical pitch perception.

IV. **Tuning systems**

Intervals generated by equal geometric subdivision of an SIPR are considered to be irrational in the sense that they don't match integer ratios. As mentioned earlier, the identification of an irrational interval with its ideal harmonic counterpart (like 1210ct being identified with 2/1) is called rationalization. Fig. 6 shows two examples of such rationalization in (extremely) different pitch-strength contexts: 12-tone temperament in a well-trained listener/clear pitch percept context (timbre: e.g. violin ord.: Fig. 6.1) and equidistant pentatonic scale in an inexperienced listener/low pitch-strength context (timbre: e.g. marimba: Fig. 6.2). Irrational intervals (a,c) - partly unstable - are *bent into place* at local energy minima (b,d). (Intervals are represented by ball-like shapes.)

a.)



b.)

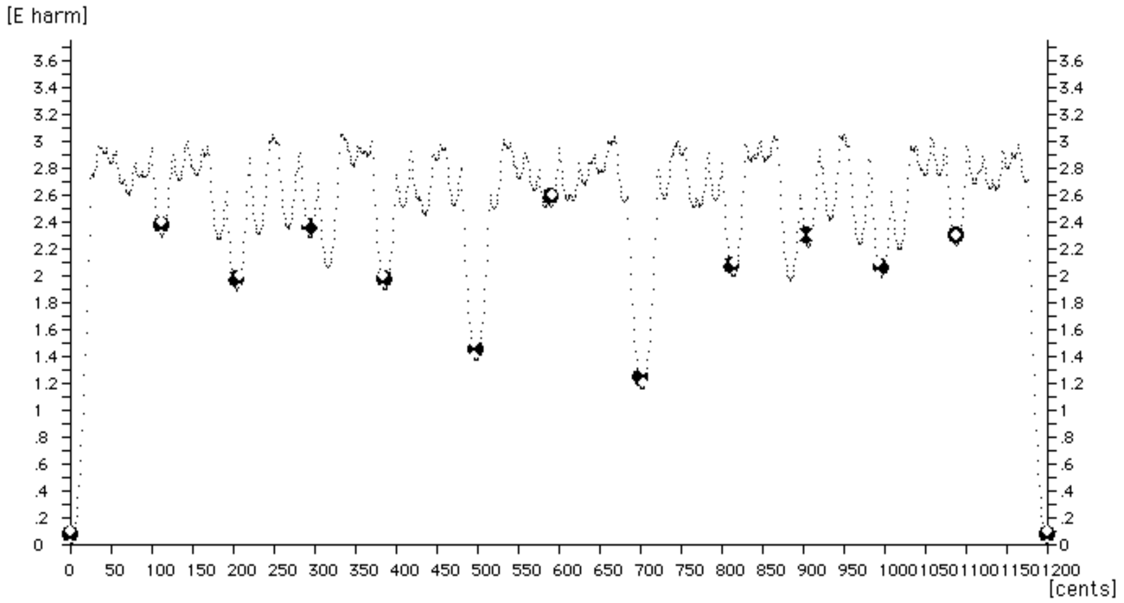
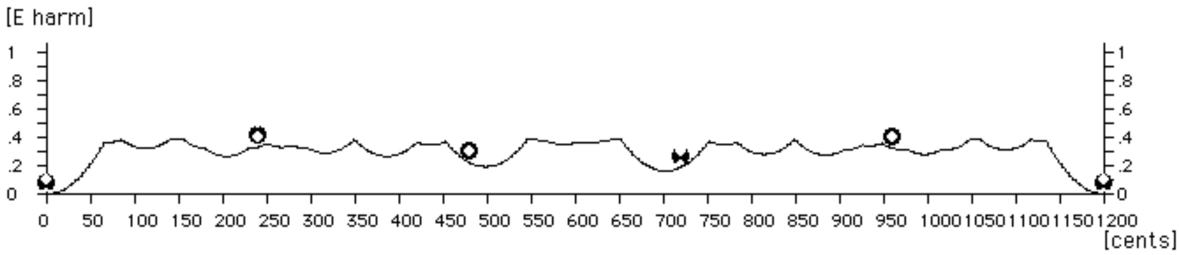


Fig. 6.1. Rationalization of a set of harmonically 'irrational' pitches in high pitch strength context (W=15ct).

c.)



d.)

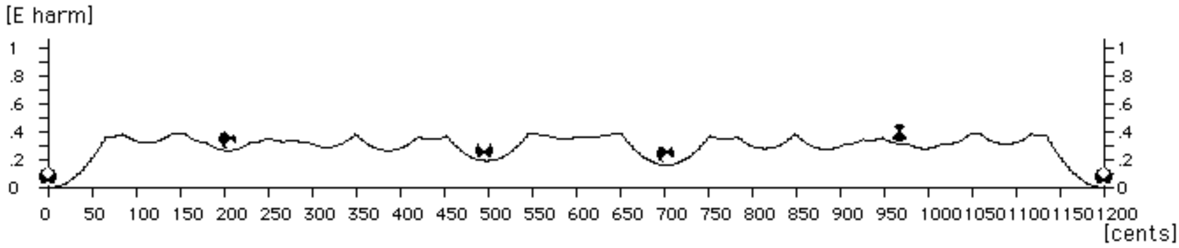


Fig. 6.2. Rationalization of a set of harmonically 'irrational' pitches in low pitch strength context (W=40ct).

V. **Combinatorial energy/relevance/quality**

Since systems in general achieve greater stability by minimizing their internal energy, it could be an interesting task to determine the tempered tuning system with the lowest combinatorial harmonic energy, and to compare the obtained data with the combinatorial energy of its rationalized counterpart. The energy ratios themselves serve as a measurement of how well the tempered system with its set of harmonically irrational intervals is able to capture the underlying harmonic meaning.

First, the combinatorial energy is calculated by applying formula 4. Given the present stage of the theory, only pitches within the first octave were considered. Energy values (constant W values) obtained by creating all possible pairs of intervals were added, and finally divided by the number of combinations[10].

$$\sum E = \frac{\sum_{m=1}^{s-1} \sum_{n=m+1}^{s} E(I_o)}{s\ (s-1)\ /\ s} \qquad \textbf{(4)}$$

s = Steps per frame interval
m,n = Indices for given pair of steps

The different combinatorial energy values for systems with 1200ct/W steps (systems with more steps per octave are not considered because of decreasing pitch discrimination; different scale degrees being identified) are compared by applying formula 5. In order to compare different systems, and to thereby determine the their **relevance,** the differences between the combinatorial energy for each system and the average combinatorial energy are scaled so that the most negative result corresponds to 10, and the most positive to 0. The average energy values are obtained by integrating the area beneath the harmonic energy curve, and dividing it by the size of the selected pitch range.

$$R(s) = a\,(\sum E - (1 + E_{ø}\,\frac{s\,(s-1)/2 - 1}{s\,(s+1)/2} + b) \qquad (5)$$

R(s) = Relevance of given system
a,b = Scaling parameter
$E_{ø}$ = Average energy

$$= \int_{l_0}^{l_1} E(I)dI \;/(l_1 - l_0) \qquad (6)$$

l_0 = lower limit of range of investigation
l_1 = upper limit of range

Finally, the **quality** of a tempered system is given by:

$$Q(s) = \frac{\sum E_{irr}}{\sum E_{rat}} \qquad (7)$$

Fig. 7 displays the ***relevance*** and ***quality*** plots for tempered systems in different pitch-strength contexts. For a high pitch-strength context (W=15ct) the system with the lowest energy contains twelve equally spaced pitches per octave: Western standard tuning. Other systems with high relevance ranks are known in music theory: 6, 17, 19, 22, 31, 41 and 53 (Barbour, 1953).

In contrast, for a low pitch-strength context (W=40ct) the equidistant pentatonic system appears to be the most stable system. "High ranks" are obtained by equidistant systems with 3 (augmented triad), 6 (whole tone scale), 7 (heptatonic system), and--again--12 steps.

Occasionally, the quality value for a particular system exceeds 100% (e.g the whole tone scale in low pitch-strength context), which implies that the tempered system has a lower energy level before rationalization. As in pure tuning systems, unstable intervals are created by combinations of stable scale degrees. It is very likely that the combinatorial energy of such a system reaches its global minimum for an interval set which is neither equidistant, nor composed by SIPRs.

V. **Discussion**

The theory of harmonic energy based on Barlow's quantitative evaluation of harmonic consonance coincides with the actual evolution of Western music over hundreds of years. Yet, this fact applies only to a high pitch-strength context in which 12-tone equal temperament is **by far** the most stable system.

It is purely speculative and scarcely demonstrable, that in a low pitch-strength context the preponderance of the equidistant pentatonic (and to a certain extent, the heptatonic) system might correspond to an early or more primitive musical stage. Only the prevalence of pentatonic systems both in early and ethnic cultures, and the fact that the pentatonic scales can be labeled even by small children, provides suggestive hints that pitch categories develop phylogenetically and ontogenetically according to the external and internal factors of pitch perception mentioned in section III. In contrast, the limitation of the available pitch set may lead to the refinement of different musical parameters, such as rhythm, and/or pitch bending.

The probability that the harmonic energy content of a tuning system may reach its global minimum for intervals which were not taken into consideration by older theories (the intervals are not equally spaced and may not be assigned to SIPRs) could imply the existence of systems like pelog and to slendro whose interval sets exhibit the above mentioned characteristics. Future work has to consider the effect of multiple vitual pitches, audible in inharmonic spectra, on the evolution of tuning systems.

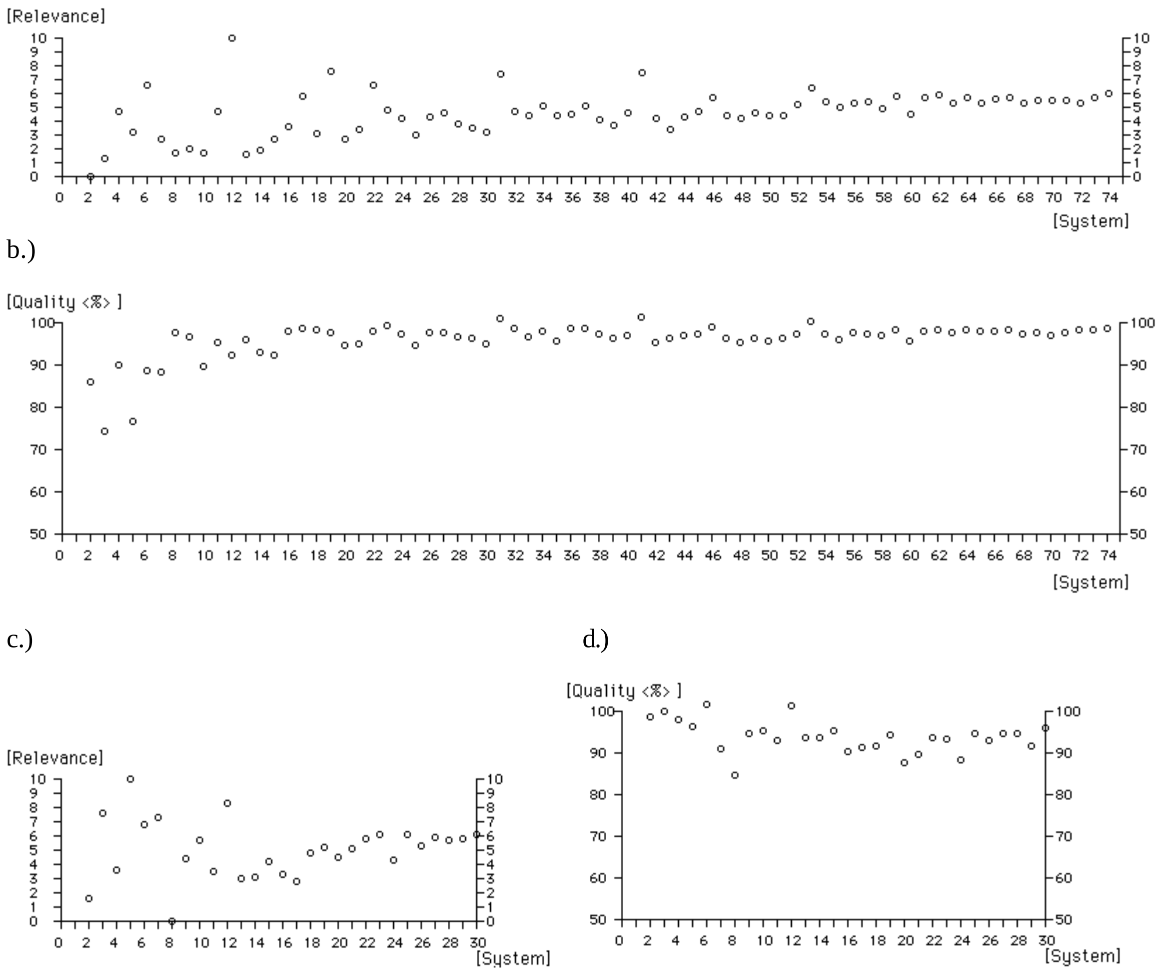


Fig. 7. *Relevance* and *quality* for W=15ct (a,b) and W=40ct (c,d).

VII. Notes

[1]For a scientific approach it is indispensable to make abstractions from single historical events (like the invention of a particular tuning system). Evolution of systems rather follow the Darwinistic principles of mutation/invention and selection, trial and error.

[2]In order to avoid terminological inconsistencies due to psychoacoustical phenomena as interval stretching (Terhardt, 1971), an ideal pitch ratio (rather than frequency ratio) is assumed. The author is aware of the fact that in tunings that employ tones with inharmonic spectra, the spacing of partials gains relative importance over SIPRs. For simplicity's sake, the theory at its present stage restricts itself to harmonic tones.

[3]Both, the consonance/dissonance phenomenon and the difference tone phenomenon occur in harmonic (vertical) musical situations (H context - exception: reverberant environments). Since early musical systems are more likely based on monody or heterophony (M context) rather than on harmony, both phenomena are secondary. They gained importance during the development of polyphony and homophony.

[4]But it comes close, which has greater importance for physiological systems.

[5]Another model for mental representation of ratios is based on the concept of harmony as special case of polymeter (Barlow, 1987). Risset (personal communication) was able to show common principles for the perception of meter and pitch by constructing analogous auditory illusions.

[6]In tonal music, the perfect-fourth interval class is affected by the increasing harmonic consonance values for the major-third pitch class. 1Surpassed by the major third in the third octave, the

perfect fourth turns into a "dissonance" with leading tone function towards the third scale degree. This might explain the varying treatment of the perfect fourth in music history and theory.

[7]A similar aspect of Euler's theoretical approach is treated by his substitution theory. According to this, complex acoustic events are substituted by simpler mental representations.

[8][In]harmonicity of spectra, temporal envelopes/transients.

[9]The root function and the value of the constant have been selected arbitrarily; future research comparing experimental data with theoretical values will eventually lead to a refinement of the pitch strength/pitch width correlation.

[10]In later adjustments of the model additional aspects will be considered: 1. pitch strength as function of register and timbre, 2. occurrence and frequency of occurrence of pitches, 3. linear pitch perception vs. fundamental-based pitch perception.

VII. **References**

Barbour, J.M. (1953). *Tuning and Temperament*. East Lansing: Michigan Sate College Press.

Barlow, C. (1981). Bus journey to Parametron. *Feedback Papers*, *21-23*.

Barlow, C. (1987). Two essays on theory. *Computer Music Journal*, *11*, 44-60.

Békésy, G. von (1960). *Experiments in Hearing*. New York: McGraw-Hill.

Busch, H.R. (1970). *Leonhard Eulers Beitrag zur Musiktheorie*. Regensburg: Bosse.

Fricke, J.P. (1973). Die Relativität optimaler Intervallgrößen. *Kgr.-Ber. Bonn 1970* (pp.397-400). Kassel: Bärenreiter.

Hajdu, G. (1989). Harmonische Energie - Motor bei der Evolution 'relevanter' Tonsysteme. *Feedback Papers*, *35*, 1-25.

Hajdu, G. (1992). *17 Töne*. Köln: Thürmchen.

Hall, D.E. & Hess, J.T. (1984). Perception of musical interval tuning. *Music Perception*, *2*, 166-195.

Hindemith, P. (1940). Unterweisung im Tonsatz. Mainz: B. Schott's Söhne.

Hornbostel, E.M. von (1905-1906). Die Probleme der vergleichenden Musikwissenschaft. *Zeitschrift der Internationalen Musikgesellschaft ,7, Vol.3*, 85-97.

Krumhansl, C.L. & Shepard, R.N. (1979). Quantification of the hierarchy of tonal functions within diatonic context. *Journal of Experimental Psychology*, *5*, 579-594.

Partch, H. (1949). *Genesis of a music*. Madison: University of Wisconsin Press.

Plomp, R. & Levelt, W.J.M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, *38*, 548-560.

Rasch, R. (1985). Perception of melodic and harmonic intonation of two-part musical fragments. *Music Perception*, *2*, 441-458.

Shepard, R.N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, *36*, 2346-53.

Shepard, R,N., Kilpatric, D.W. & Cunningham, J.P. (1975). The internal representation of numbers. *Cognitive Psychology*, *7*, 115-118.

Terhardt, E. (1971). Die Tonhöhe Harmonischer Klänge und das Oktavintervall. *Acustica*, *24*, 126-136.

Terhardt, E., Stoll, G., and Seewann, M. (1982). Pitch of complex signals according to virtual pitch-theory: tests, examples, and predictions. *Journal of the Acoustical Society of America*, *71*, 679-88.

Vos, J. (1988). Subjective acceptability of various regular twelve-tone tuning systems in two-part musical fragments. *Journal of the Acoustical Society of America*, *83*, 2383-92.