Title

Pitch Structure of Melodic Lines: An Interface between Physics and Perception

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN

1069-7977

Authors

Useche, Jorge Eduardo
Hurtado, Rafael German

Publication Date

2011

Peer reviewed

Pitch Structure of Melodic Lines: An Interface between Physics and Perception

Jorge E. Useche (jeusecher@unal.edu.co)
Conservatorio de Música & Departamento de Física, Universidad Nacional de Colombia, Carrera 30 No. 45-03
Bogotá, Colombia

Rafael G. Hurtado (rghurtadoh@unal.edu.co)
Departamento de Física, Universidad Nacional de Colombia, Carrera 30 No. 45-03
Bogotá, Colombia

Abstract

This work studies melodic lines of western art music tradition from frequency of occurrence of their constitutive elements. The model of analysis uses pairs of successive sounds as the minimal structural elements of a melody. Each pair of possible sounds in a musical instrument is associated to a quantity related to the difference of acoustic energies of the sound waves. This quantity expresses consonance properties that have been studied in experiments about the perception of combinations of sounds. We find statistical distributions of this quantity that show the existence of preferences for certain elements in a given melodic line. This preference can be interpreted as a consequence of the use of both formal musical theory rules and the creativity of the composer in order to create pleasant sensations in the listener.

Keywords: Consonance; Melody; Music composition; Perception; Cognition; Tempered scale.

Introduction

George Kingsley Zipf explored the relationship between the usefulness of a word and its frequency of occurrence in a text (Zipf, 1932). He defined *rank* (*r*) as the position of a word in an ordered list that goes from 1, for the most frequent word, to the number of words that differ "phonetically". The relation between rank and frequency of occurrence (*f*) of a word in the Ulysses from James Joyce resulted into a power law with exponent -1, that can be expressed as $\boldsymbol{r} \times \boldsymbol{f} = \boldsymbol{c}$ where *c* is a constant. Zipf observed the same phenomenon in other human systems and enunciated a law of human behavior driven by a Principle of Least Effort (Zipf, 1949) that he obtained in analogy to mechanics in physics.

Zipf also studied the case of melodic lines and reported a power law relationship between the length of musical intervals, defined as the number of notes between two notes that are played successively (Zipf, 1949), and the frequency of occurrence of each interval. In other words, the probability of finding an interval in a melodic line is inversely proportional to its length. He found the same result for ascending and descending intervals and combined both of them in the analysis.

Benoît Mandelbrot (1966) extended Zipf's work and found a more general form of power law that also describes other phenomena as Pareto's rule. Most recently, Statistical mechanics has been used to describe the statistical properties of many systems, ranging from financial markets and wealth distribution to complex networks, which present either exponential or power law distributions. Power law distributions are frequently associated to self-organized criticality and scale invariance, while exponential distributions are associated to equilibrium processes and to the occurrence of an absolute scale that is similar to temperature in Statistical physics.

Further studies of musical pieces using probabilistic aspects of musical styles and uncertainty in musical communication have been done in Information theory (Abdallah & Plumbley, 2009; Cohen, 1962; Cox, 2010; Meyer, 1957). These studies capture fundamental aspects of perception as sensation and meaning. This perspective needs some basic mathematical assumptions such as stochasticity, ergodicity, stationary and Markov consistency of the source (i.e. the piece of music itself), as well as an infinite memory capacity of the encoder (i.e. the human brain) (Cohen, 1962). These studies use quantities such as the mean and the variability of statistical distributions of sequences of sounds.

Recent studies show that some statistical properties of musical compositions can be treated using concepts and analytical tools developed for studying complex systems by applying Statistical physics. Liu, Tse and Small (2010) found that successive notes in musical pieces can be represented as complex networks that exhibit power law distributions for some connectivity properties. Gündüz and Gündüz (2010) studied melody formation in musical pieces and found that during the progress of a melody the entropy grows with each new note until it takes a limiting value that is smaller than the entropy of a random composition.

This paper explores the relation between microscopic properties of consonance in a melodic line and the macroscopic properties resulting from the composition process. Its aim is to establish links between pleasant sensations at the microscopic level with the ordering process which is needed to compose a musical piece. It presents the physical model for describing consonance, probability distributions for pitch structures and a discussion about some implications.

Paths between cognition and physics

A piece of music is a mentally constructed entity that is usually described as segmented in unites of all sizes. In order to ascribe some sort of reality to the internal structure of musical pieces one must treat them as mental products imposed or inferred from physical signals (Lerdahl &

Jackendoff, 1983). Physics and music deal with sound from two different perspectives that cross each other in the cognitive processes related to perception and creativity. Pythagoras posted that two sounds produced simultaneously by strings of equal tension and density and with lengths (i.e. frequencies) related by small Natural numbers produce a pleasant sensation (Rossing, 1989). Pythagoras' postulate relates subjective qualities to physical parameters and allows to define consonance as a subjective appreciation about pleasure (Plumb & Levelt, 1965). The theory of sound in physics states that a musical note is characterized by a fundamental frequency, timbre, loudness and duration. The fundamental frequency is the lowest frequency present in the sound. Timbre depends primarily on the spectrum of the stimulus. Loudness depends mainly on sound pressure and duration refers to the interval of time (Rossing, 1989). When two sound waves are superposed the physical properties of the resulting wave depend upon the frequency structure of the original waves (Helmholtz, 1862). Then, the level of consonance of two sounds played simultaneously (harmony) or in a rapid succession (melody) can be treated formally in terms of physical quantities.

Statistical properties of systems with many parts follow new laws different from those of mechanics. The study of these systems requires only of the knowledge of system composition and statistical distributions of the properties of the parts (Landau & Lifshitz, 1980). These applications involve social, natural and artificial systems and range from financial markets to complex networks and music.

On this new framework that involves both Information theory and Statistical mechanics to study musical pieces, physics acquires a new role in cognitive sciences by exploring both perception and the composition process based in formal composition rules and creativity of the composer.

Method

Melodic lines result from formal voice conduction and rhythm rules established in Music theory as well as from creativity of the composer (Aldwell & Schachter, 1989). Voice conduction rules involve both harmonic and melodic motion in a specific musical style, while rhythm involves time structure (bar) and the organization of the beats. These rules are the formal constraints for the composer. From this perspective, creativity of the composer is embedded in the freedom left by voice conduction and rhythm rules.

Voice conduction rules and the rhythmic structure of the formed melodic lines cannot be considered strictly as independent (Korsakov, 1930), however as rhythm has to do mainly with the duration and the intensity of sounds (beats) while voice conduction rules are related to perception of combinations of sounds, by ignoring the bar (i.e. the organization of the beats) we decouple the problem into pitch and rhythm figures problems. Our analysis centers in the pitch structure of melodic lines. Pitch is a subjective quality of sound related strongly with the fundamental frequency of a note and weakly related with pressure, spectrum, duration and envelope, all of them physical parameters (Rossing, 1989). On this work we use the fundamental frequency of a note in order to describe pitch.

The pitch structure is studied in musical theory through the succession of notes (Korsakov, 1930; Liu, Tse & Small, 2010). A note contains information about pitch and duration that can be distinguished in the score of a piece. At the microscopic level, Paul Hindemith (1942) used the length of musical intervals to characterize the consonance properties of simultaneous musical notes (i.e. related to harmony) and rapid successions (i.e. related to melody). The analysis made by Zipf (1949) uses the length of musical intervals between pairs of successive notes to characterize a melody at the macroscopic level.

Consonance has been studied formally in Music theory (Aldwell & Schachter, 1989) as well as in physics (Helmholtz, 1862; Plumb & Levelt, 1965; Rossing, 1989). From physics perspective and using physiological arguments, the difference and the ratio between simultaneous pure tones (i.e. sounds with just one frequency) have been used to study consonance (Helmholtz, 1862; Plumb & Levelt, 1965; Rossing, 1989). Plumb & Levelt (1965) made detailed experiments to relate the difference of frequencies of two simultaneous pure tones $(f_j - f_i)$ with the sensation of pleasure associated to consonance.

A natural extension of Zipf's analysis is to distinguish between all possible frequency transitions that can be related to each musical interval. Thus, any ordered pair of notes must be associated to only one transition containing information about the difference of frequencies and tone heights, as both are important for the listener's perception (Patterson, 1990; Plumb & Levelt, 1965). The product $(f_j - f_i) \cdot (f_j + f_i)$ contains information about consonance and tone heights. It is positive for ascending transitions and negative for descending ones. This quantity is extremely appealing, as it corresponds to the difference of average sound energy density carried by two waves with the same amplitude and frequencies f_i, f_j, thus $(f_j - f_i) \cdot (f_j + f_i) = f_j^2 - f_i^2$. The average sound energy density is $\varepsilon = (1/2)\rho_0 f_i^2 \eta_m$, where ρ_0 is the density of air and η_m is the maximum amplitude of the displacement of particles (Pain, 1992). In the case of simultaneous pure tones the superposition of two waves with frequencies $f_i < f_j$ produces a wave with a fast frequency $(f_j + f_i)$ and a modulation frequency $(f_j - f_i)$. It means that this quantity is meaningful for both melodic and harmonic intervals.

In order to perform the data analysis we transformed the MIDI codes to a sequence of frequencies in the Tempered scale as presented by Rossing (1989). Then, we found the set of values for the transitions $f_j^2 - f_i^2$ and, in order to find the distributions for this quantity, we defined bins or steps Δf^2 of equal size in such a way that any of them would contain at least one transition $f_j^2 - f_i^2$ for the full combination of the notes in the Tempered scale. The reason to define the bins in this way is twofold: First, to account for the *tessitura* of the instrument. Second, in most cases the number of possible transitions $f_j^2 - f_i^2$ for an instrument in the Tempered scale is larger or of the same order as the actual number of transitions in a musical piece. For comparison purposes, we normalize all distributions from the frequency of occurrence to probability distributions.

Experimental data

The selected pieces for this study are characterized by the internal coherence of their melodic lines. For all the pieces we used reduced Musical Instrument Digital Interface (MIDI) tables, obtained either from the scores or from available MIDI files. The only treatment made to the data was cutting melodic lines at the end of a section or when a rest (interval of silence) is found.

Missa Super Dixit Maria (Hans Leo Hassler)

This is a polyphonic composition for four voices (*soprano*, *contralto*, *tenor* and bass). Data acquisition was carried directly from the score.

Brandenburg *Concerto* No. 3 in G major, BWV 1048 (Johann Sebastian Bach)

The polyphonic texture of this piece shows a relative independence between melodic lines, each one played by a different musical instrument. This piece has eleven instruments. Data acquisition was carried directly from the score.

Suite No. 1 in G major BWV 1007 and Suite No. 2 in D minor BWV 1008 (Johann Sebastian Bach)

These pieces were written for a melodic instrument (Cello) and the melodic lines can be distinguished without any ambiguity. Data acquisition was carried directly from the score.

First movement of *Partita* in A minor, BWV 1013 (Johann Sebastian Bach)

This piece has just one melodic line for a melodic instrument (flute). Data acquisition was carried directly from the score.

Arrangement for flute of *Entr'acte* of the Act 4 of "Carmen" (Georges Bizet)

This arrangement for flute has just one melodic line. The MIDI file was downloaded the 6th of July of 2010 from http://www4.osk.3web.ne.jp/~kasumitu/eng.htm.

Results

All probability distributions for individual melodic lines fit well to exponential functions of form: $g(x) = A_1 e^{-x/t_1}$. We have taken the absolute value for descending transitions. For the superposition of all melodic lines in the *Missa* Dixit Maria the probability distribution also follows an exponential function, however for Brandenburg *concerto* it describes a power law.

Missa Super Dixit Maria

Figures 1 and 2 show the probability of occurrence of the ascending transitions for the *contralto* and *soprano* voices. In Figure 1 the bin size is 15000 Hz^2 and R^2=0.993. In Figure 2 the bin size is 16000 Hz^2 and R^2=0.992. The difference of 1000 Hz^2 in the bin size is due to the change in the *tessitura* between *contralto* and *soprano*, as transitions for *soprano* are larger than for *contralto*. Figure 3 shows that ascending and descending transitions have similar distributions for this piece. Figure 4 combines ascending and descending transitions for the four voices. Bin size in Figures 3 and 4 is 15000 Hz^2.

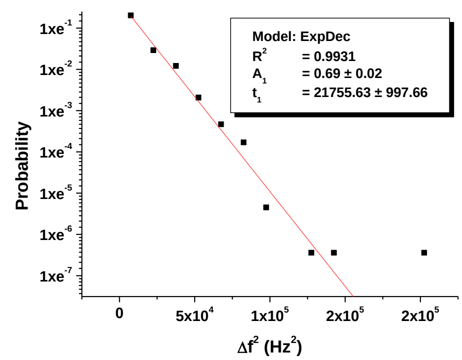

Figure 1: Probability distribution for the ascending transitions of the *contralto* voice of the *Missa* Super Dixit Maria.

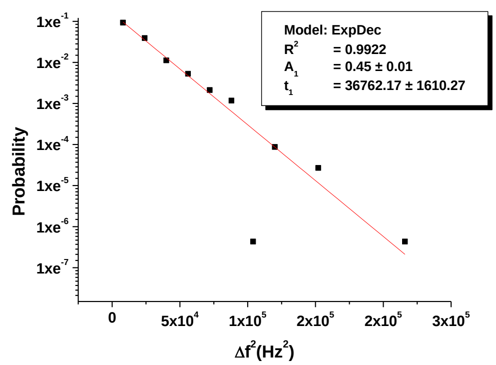

Figure 2: Probability distribution for the absolute value of the descending transitions of the *soprano* voice of the *Missa* Super Dixit Maria.

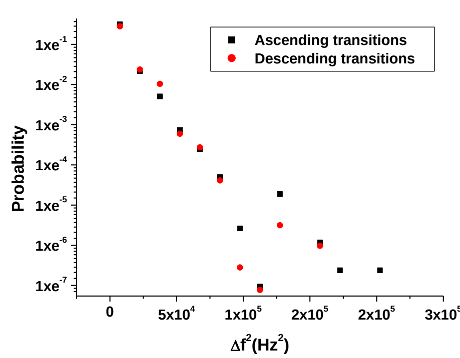

Figure 3: Probability distribution for the ascending (squares) and descending (circles) transitions for the superposition of the four melodic lines of the *Missa* Super Dixit Maria.

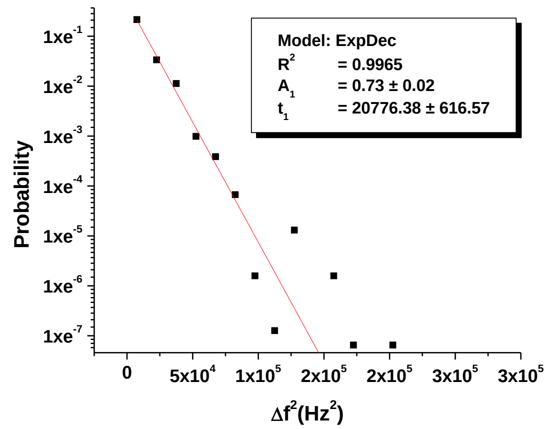

Figure 4: Probability distribution for the combined ascending and descending transitions for the four melodic lines of the *Missa* Super Dixit Maria.

Brandenburg *Concerto* No. 3 in G major, BWV 1048 (Johann Sebastian Bach)

In Figure 5 we have the exponential distribution for the combination of ascending and descending transitions of the melodic line for the first violin. The bin size is 50000 Hz^2. Table 1 contains the relevant parameters of the fit for each melodic line of the *Concerto* for the combination of both ascending and descending transitions.

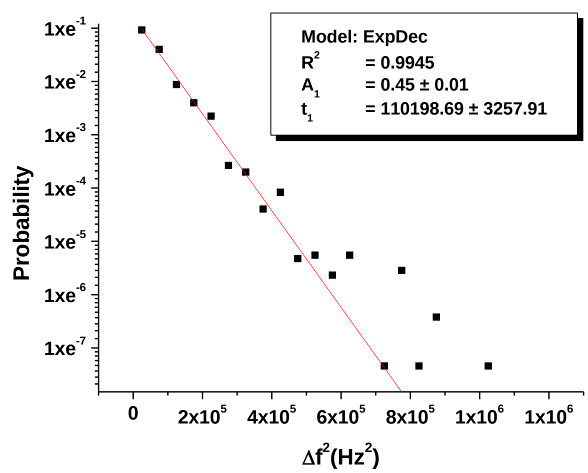

Figure 5: Probability distribution for combined ascending and descending transitions in the melodic line of the first violin of the Brandenburg *Concerto* No. 3.

Table 1: Relevant parameters for the probability distributions of the Brandenburg *Concerto* No. 3.

Instrument	t_1	R^2	Bin (Hz^2)
Violin 1	110198,69 ± 3257,91	0,995	50000
Violin 2	99858,85 ± 4438,80	0,989	50000
Violin 3	79525,72 ± 2619,57	0,995	50000
Viola 1	44936,96 ± 2601,48	0,984	20000
Viola 2	33832,24 ± 1847,97	0,987	20000
Viola 3	32656,70 ± 1564,20	0,990	20000
Cello 1	7656,90 ± 344,27	0,991	5000
Cello 2	7648,09 ± 335,76	0,991	5000
Cello 3	7426,44 ± 324,27	0,991	5000
Violone	1573,07 ± 63,94	0,992	1200
Harpsichord	6527,52 ± 279,40	0,992	5000

Figure 6 contains the distribution for combined ascending and descending transitions for the eleven melodic lines of the Brandenburg *Concerto* No. 3. The bin size is $16000\ Hz^2$. This distribution can be analyzed in two parts: The first one goes from the smallest transitions to $1{,}84 \times 10^5$ and the second one from 2×10^5 to the end. The first part exhibits a power law of the form: $p(x) = C/x^D$ and the second part shows a fat tail. This behavior is frequent in complex systems (Liu, Tse & Small, 2010).

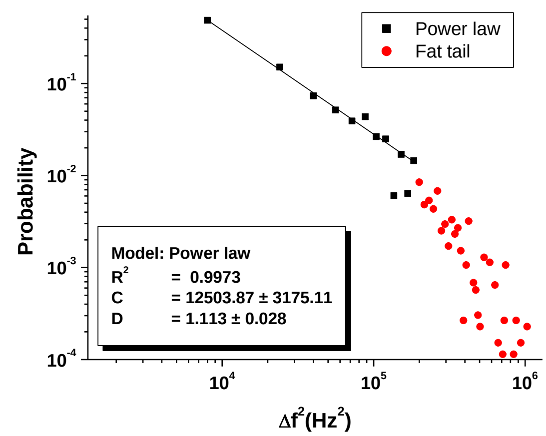

Figure 6: Probability distribution for combined ascending and descending transitions for the eleven melodic lines of the Brandenburg *Concerto* No. 3.

Suite No. 1 in G major BWV 1007 and Suite No. 2 in D minor BWV 1008 (Johann Sebastian Bach)

Figure 7 shows the probability distributions for combined ascending and descending transitions for Suites Nº1 y Nº2 of J.S. Bach. The bin size is 8000 Hz^2. The fit parameters are A_1 =0,53±0,03, t_1=15692,28±1140,61 and R^2=0,977 for Suite N°1, and A_1=0,60±0,01, t_1=13372,73±224,60 and R^2=0,999 for Suite Nº2.

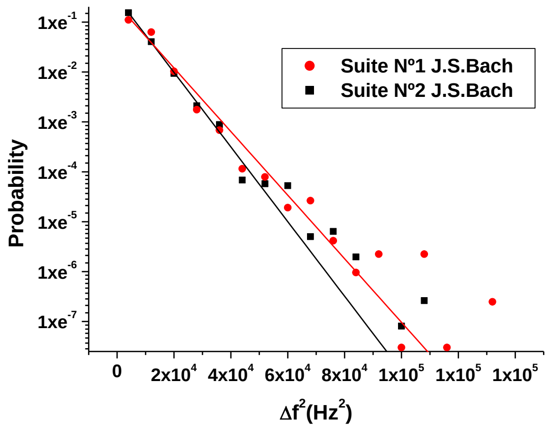

Figure 7: Probability distribution for combined ascending and descending transitions for Suites N°1 and N°2 of Johann Sebastian Bach.

First movement of *Partita* in A minor & Arrangement for flute of *Entr'acte* of the Act 4 of "Carmen"

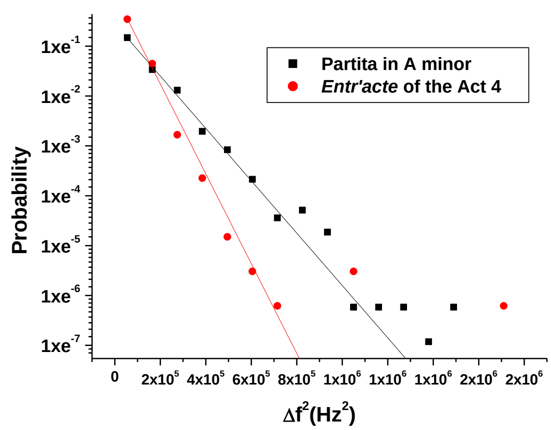

Figure 8: Probability distributions for combined ascending and descending transitions for *Partita* en A minor of J.S. Bach and Arrangement for flute of *Entr'acte* of the Act 4 of "Carmen".

Figure 8 shows the probability distributions for the combined ascending and descending transitions for *Partita* in A minor of J.S. Bach and Arrangement for flute of *Entr'acte* of the Act 4 of "Carmen". The bin size is $1{,}1x10^5$ Hz^2. Fit parameters to exponential functions are $A_1=0{,}58\pm0{,}01$, $t_1=190443{,}48\pm4989{,}61$ and $R^2=0{,}997$ for *Partita* in A minor, and $A_1=1{,}05\pm0{,}03$, $t_1=111013{,}48\pm3598{,}58$ and $R^2=0{,}997$ for *Entr'acte* of the Act 4.

Discussion

Results show that all melodic lines exhibit exponential behavior for probability distributions of transitions $f_j^2 - f_i^2$. For a first approach to this issue we analyze the effect of musical scale in a composition. For Tempered scale the distribution of musical intervals becomes "uneven" if they are expressed in terms of frequencies and it transforms into an exponential distribution. This is due to the fact that there are more $f_j^2 - f_i^2$ in a given Δf^2 for some small $f_j^2 - f_i^2$ in contrast to the for large ones. Figure 9 contains the distribution that results from all possible Δf^2 of frequencies that Bach used for the first violin of the third Brandenburg *concert*o as well as the actual distribution for melodic line. The effect of the Tempered scale contributes to the measured distribution of the melodic line but it does not explain the results observed. We interpret the contribution from the tempered scale as coming from its natural consonance properties. We hypothesize that the difference between both distributions must be related to formal composition rules and creativity of the composer, both of them are relevant to obtain a pleasant melodic line.

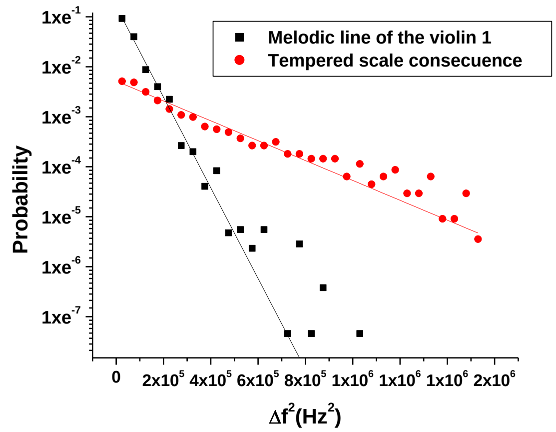

Figure 9: Probability distribution for a melodic line of the first violin of third Brandenburg *concerto* (squares), and the effect of Tempered scale (circles).

Figure 3 shows that ascending and descending transitions have the same behavior. Zipf (1949) obtained the same result for musical intervals. Figure 6 shows a power law distribution for the combination of the eleven melodic lines of third Brandenburg *Concerto*. Chu-Shore, Westover and Bianchi (2010) studied the conditions for the formation of a power law distribution from three exponential functions. In order to compare their analysis, we group melodic lines based on the *tessitura* of the instruments. The first group contains three violins, the second one violas and the third one cellos, violone and harpsichord. Figure 10 shows the frequency distribution for the three groups and for the full composition. Figures 7 and 8 compare some pairs of melodic lines in the same *tessitura*. We observe that the two Suites for cello of J.S. Bach have a similar behavior (Figure 7), while the *Partita* in A of J.S. Bach and *Entr'acte* of the Act 4 of "Carmen" exhibit clear differences.

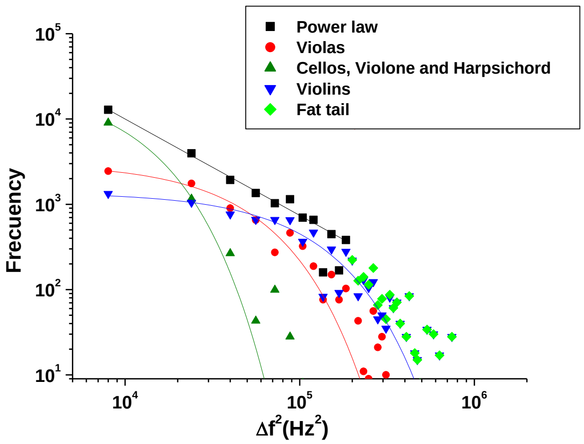

Figure 10: Power law for the combination of the eleven melodic lines of the third Brandenburg *concerto*, and for the three groups of melodic lines coming from the *tessitura* of instruments.

Conclusions

The physical quantity $f_j^2 - f_i^2$, defined for transitions between successive notes instead of musical intervals used by Zipf (1949) involves information that is related to the consonance properties contained on the difference between frequencies $(f_j - f_i)$ and the "tone height" of both notes $(f_j + f_i)$. Since $f_j^2 - f_i^2 = (f_j - f_i) \cdot (f_j + f_i)$, this can be related to both, melody (succesive sounds) and harmony (simultaneous sounds) as well. This quantity comes from the difference of the sound energy density of two waves with equal amplitudes and therefore it is a relevant quantity for musical analysis rooted in physics.

All the studied melodic lines are characterized by exponential distributions of quantity $f_j^2 - f_i^2$. On this work we have not discussed about this result from the perspective of Statistical physics however it has to be stated that there is a close relationship between this type of distributions and the conservation laws.

Consonance properties of Tempered scale are reflected in the fact that small values of the quantity $f_j^2 - f_i^2$ are more frequent than the large ones. This phenomenon is reinforced by the composer in all studied pieces and indicates that it is related to the sensation produced in the listener by the ordering process involved in the selection of the elements in a musical piece. The difference between the effect of the Tempered scale and the actual distributions for pieces studied can be explained by the use of formal composition rules and creativity of the composer. Both related to the cognitive process involved in the composition of pleasant melodic lines.

Acknowledgments

J. Useche thanks *Fundación Mazda para el arte y la ciencia* for the scholarship for undergraduate studies in Piano.

References

Abdallah, S. & Plumbley, M. (2009) *Information dynamics: Patterns of expectation and surprise in the perception of music*. Connection Science, 21(2), 89–117.

Aldwell, E. & Schachter C. (1989) *Harmony and voice leading*. Harcourt Brace Jovanovich, Publishers, second edition.

American National Standards Institute (1960) *USA Standard Acoustical Terminology*. S1.1-1960

Cohen, Joel E. (1962), *Information Theory and Music*. Behavioral Science, 7:2, 137.

Cox, G. (2010) On the Relationship Between entropy and meaning in Music: An Exploration with Recurrent Neural Networks. *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society*, 2010.

Fan Liu, X. Tse, C. & Small, M. (2010) *Complex network structure of musical compositions: Algorithmic generation of appealing music*. Physica A 389, 126-132.

Günduz, G. & Günduz, U. (2005). *The mathematical analysis of the structure of some songs*. Physica A 357 (2005) 565–592

Helmhotz, H. (1862) *On the Sensations of Tone as a Physiological Basis for the Theory of Music.* Dover Publications Inc, 2nd edition.

Hindemith, P. (1942) *The Craft of Musical Composition, Book I*. Associated Music Publishers Inc., New York.

Korsakov N. (1930) *Practical Manual of Harmony*. Carl Fischer, LLC. First English edition.

Landau, L. & Lifshitz, E.M. (1980) Statistical physics. Third edition. Pergamon Press Ltd. Hungary

Lerdahl, F. & Jackendoff, R. (1983). *A generative theory of tonal music*. The MIT press

Mandelbrot, B. (1966). *Information Theory and Psycholinguistics: A Theory of Words Frequencies*. In Readings in Mathematical Social Science. P. Lazafeld and N. Henry, Editors, Cambridge MA, MIT Press

Meyer, L.B. (1957) *Meaning in music and information theory*. The Journal of Aesthetics and Art Criticism, 15(4), 412–424.

Pain, H.J. (1992) *The physics of vibrations and waves.* John Wiley and Sons, 4th edition.

Patterson, R.D. (1990). *The Tone Height of Multiharmonic Sounds*. Mus. Percept. 8**,** 203–214.

Plumb, R. & Levelt, W.J.M. (1965) *Tonal consonance and critical bandwidth*. J. Acoust. Soc. Am. 46:409.

Rossing T.D. (1981). *The Science of sound*. Adisson-Wesley Publishing Company, 2nd edition.

Chu-Shore J., Westover M.B. & Bianchi M.T. (2010) *Power Law versus Exponential State Transition Dynamics: Application to Sleep-Wake Architecture*. PLoS ONE 5(12): e14204. doi:10.1371/journal.pone.0014204

Zipf, G. K. (1932). *Selected Studies of the Principle of Relative Frequency in Language*. Harvard University Press. (Mentioned in Zipf, G.K. (1949))

Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort*. Cambridge, Mass. Addison-Wesley Press.