A Singular Measure of Dissonance

or how

Music provides an empirical support in Number Theory

Alexandre Guillet

May 14, 2018

Musical *sonance* can be described by a singular measure, based on Minkowski's question mark function. This arises from a refinement of the continued fractions, which extends the analysis of commensurability from numbers (musical intervals) to functions (stationary sounds' spectrum). An application to musical scales characterization brings a strong empirical support for its musical relevance. The Riemann zeta function appears in this context to describe the influence of sound timbres. This yields a musical interpretation of its famously conjectured property, weaving together number theory and solfeggio. Methods for numerical application are discussed. The minimal framework of this analysis is extracted to be exported to various commensurably rich objects such as rhythm. A formal connection to the statistical physics of the primon gas is also evoked.

Contents

Context

The entanglement of Maths and Music can be tracked back to the Pythagorean school, twenty five centuries ago. With the eye guided by the compass and the ear by the monochord, Pythagoreans' thoughts about commensurate quantities has lead both to a method to tune musical instruments and to the discovery of irrational numbers. This intertwining sounds and looks more mysterious at the time of Euler's *gradus suavitatis*. The suavity degree involves an additional prime number to describe intervals, 5, as compared to the Pythagorean approach. This leads to the conception of the *Tonnetz*. This abstract tones network is the common structure behind the many dodecaphonic scales that were developped to improve the limitations of the Pythagorean scale. It reflects the fundamental difficulty to deal with irrationality in musical structures. Some hearing intuition was then given back to Maths with the development of the Fourier analysis, which made Helmholtz design sophisticated psycho-acoustic experiments. He considered the influence of overtones, that constitutes the sound timbre, in consonance and tuning. His work has yielded the construction of a dissonance curve, probably the first harmony measure embedded on a continuum of musical intervals and smoothing out irrationality issues.

The recent Math&Musical works around these issues as well as the tools developped in the following will be refered to these three conceptions. William Seather dissonance curves, methods, influence of timbre on dissonance.........openly Helmholtzian

Paul Ehrlich harmonic entropy is another exemple of dissonance curve, smoothing out irrationality in another way. (interesting analysis of triads which could constitute a generalization of the work here). Harmonic entropy generalized... in the community of microtonality developers Xenharmonic.......... knowledge of the musical relevance of the zeta function reflecting the overtone importance in dissonance.....but missing piece for drawing the associated dissonance curve, presented here.

The Eulerian inheritance is sensible... in developers of algebraic structures such as generalized Tonnetz (neo Riemanian...things...?) as well as prime decomposition based approached. Work of Albert Gräft is representative of development of Euler idea. Euler's $g(p) = p - 1$ is the $\epsilon = 1$ version of $g(p) = \frac{p^\epsilon - 1}{\epsilon}$, the $\epsilon = 0$ version is simply $g(p) = \log p$ yeilding $g(q) = \sum_p |a_p| \log p$ termed the complexity function by F. Faure, thought as a distance function on the generalized tonnetz, and the .... measure in the Xentonality community. This object with $\epsilon = 0$ appears in this work (multiplicative? additive? nb theoretical function)

The use of the Stern Brocot tree or continued fractions in the work of..........can be termed of Pythagorean flavour. Surprisingly, the same author explored the thightly related Minkowski's question mark, whithout however bringing musical intuition. In the same way, the mathematician but also music composer G. Alkauskas has developped a great part of the mathematics associated with the Minkowski's question mark function as a measure, some of which are directly relevant to music as explained below.

.......two ingredients : spectrum of relation (find discrete version in Emmanuel Amiot's ....) and singular measure (to my knowledge never applied to music) which is the central contribution of this paper.

1. One possible mathemusical genesis of harmony

1.1. Consonance and commensurability of frequencies

Musical intervals can be perceptually ordered by degree of consonance, which seems closely related to the commensurability of their vibration frequencies, as observed by Pythagoras. But Mathematics puts usually the stress on the rationality or the irrationality of the number, whereas Music just cares about the simplicity —up to a correct approximation— of the frequency ratio (more simply called interval). The partial ordering of intervals by degree of simplicity proposed empirically by the musical harmony is well reproduced by the iterative construction of the Stern-Brocot binary tree, which provides a natural enumeration of the rational numbers in reduced form $\frac{a}{b}$. The vertical growth of this tree, by means of "mediants", represents the decreasing simplicity of rational numbers while preserving the natural ordering horizontally.

$$\text{generators: } \frac{0}{1}; \frac{1}{0} \qquad \text{mediants: } \frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$$

The displacement in this research tree from the root $\frac{1}{1}$ (first generated mediant) towards the value of any positive real number constitutes an increasingly precise and complex approximation. This depicts graphically the continued fraction algorithm. The trajectory in the tree can be represented by a sequence $s_n \in \{0; 1\}$, where 0 means *less* and 1 means *more* at the next step. The research in the tree of all rational numbers will not stop if the approximated number is irrational. Hence, we can unify the binary sequence representation of any number by completing the finite sequence by 100000... (or equivalently 011111...) where the first bit makes us go one step away from the exact rational value, but the infinite following bits make us converge eventually to the exact value again (by above or by below). The continued fraction is a condensed representation where the successive identical bits are replaced by a block size sequence α_n (where the first block size stands for a block of one): $1^{\alpha_1}0^{\alpha_2}1^{\alpha_3}0^{\alpha_4}.... \to \{\alpha_1;\alpha_2;\alpha_3;\alpha_4;...\}$. The sum of all finite block sizes for a number $x \in [0,\infty]$ is the information needed or the number of stages of the tree to find the exact value of the number. We call it its *harmonicity* $H(x)$.

1.2. The answer to Minkowski's question mark function?

Interpreting the binary sequence as the decimals of a binary number, we resume these previous representations in a unique number, which represents the horizontal position in the tree when rational numbers are equally spaced at each stage:

$$\text{𝄞}(x) = \sum_{n>0} \frac{s_n}{2^n} = (0.s_1s_2s_3...)_2 = (0.1^{\alpha_1}0^{\alpha_2}1^{\alpha_3}...)_2$$

This is a non decreasing bijective map between the non-negative real numbers and the unit interval. We call it the *musical function*, for reasons that will become clearer. It has the following inversion symmetries, self-similarity and mediant property:

$$\text{𝄞}\Big(\frac{1}{x}\Big)+\text{𝄞}(x) = 1 \quad ; \quad \text{𝄞}\Big(\frac{1}{1+\frac{1}{x}}\Big) = \frac{1}{2}\text{𝄞}(x) \quad ; \quad \text{𝄞}\Big(\frac{a+c}{b+d}\Big) = \frac{\text{𝄞}(\frac{a}{b}) + \text{𝄞}(\frac{c}{d})}{2} \; \forall |ad-bc| = 1$$

We recognise the pattern of the iterative expansion in continued fraction, as well as the growth of the harmonic (Stern-Brocot) tree. Its has many interesting properties that have already been studied for the essentially identical function $?(x)$ called *Minkowski's question mark function* (see [1] [2]).

$$?(x) = 2\text{𝄞}(x) \quad \forall x \in [0,1] \quad \Leftarrow \quad \text{𝄞}(x) = ?\Big(\frac{1}{1+x^{-1}}\Big) \quad \forall x \in [0,\infty]$$

The property of mapping quadratic irrational to rational numbers is the reason why it has been defined by Minkowski (original article [3]). The simplest example writes:

$$?(\varphi^{-1}) = \frac{2}{3} \quad \Leftrightarrow \quad \text{𝄞}(\varphi^{-1}) = \frac{1}{3} \quad ; \quad \text{𝄞}(\varphi) = \frac{2}{3}$$

where $\varphi = \dfrac{1+\sqrt{5}}{2}$ denotes the golden ratio.

We will explain in the next sections why the implicit question set more than one century ago in the glyph of this exotic function is answered by Music.

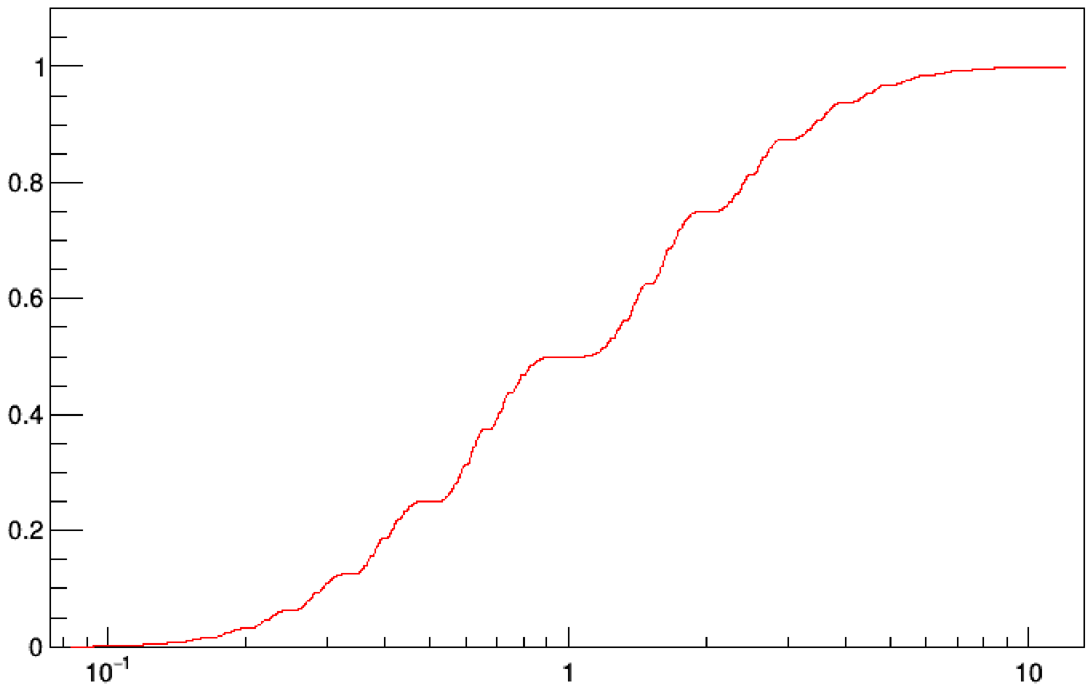

Figure 1: Musical function $\mathcal{G}(x)$ with a logarithmic X-axis. Note the size of the stair steps at simple rational number: largest for the unison $\frac{1}{1}$, followed by the octave $\left(\frac{1}{2}; \frac{2}{1}\right)$, the fifth $\left(\frac{2}{3}; \frac{3}{2}\right)$ and fifth plus octave $\left(\frac{1}{3}; \frac{3}{1}\right)$, etc.

1.3. Requirements for a dissonance function

The harmonicity $H(x)$ could constitute a rough dissonance function for the number x. A more common choice consists in $C\left(\frac{a}{b}\right) = ab$, called the complexity of the rational number $\frac{a}{b}$. We can also consider the simplicity $C(\frac{a}{b})^{-1} = \frac{1}{ab}$ which has an interesting property: its sum over all rational number in the same stage of the harmonic tree is one. This sketches the order of magnitude $C \sim 2^H$, despite big variations ranging from $C = H$ for $\frac{N}{1}$ to $C \sim \frac{\varphi^{2H+1}}{5}$ for $\frac{F_{N+1}}{F_N}$ where F_N are large Fibonacci numbers and φ is the golden ratio. We can also consider the logarithm of the complexity and use it as a graph distance in a generalised *Tonnetz* based on the prime number decomposition [4].

Variations around $C\left(\frac{a}{b}\right) = ab$ do not change the relative ordering of the intervals. Moreover, they satisfy an important property: their are all invariant under inversion of the rational number $q \leftrightarrow q^{-1}$. This translates in Music as "the dissonance of a rising interval (going from a lower tone to a higher one) is the same dissonance as the falling interval". Its additive form is the most convenient one from a musical point of view: $n \leftrightarrow -n$, where $n = \frac{\log_2 q}{N}$ is in octave, semi-tones or cents for $N = 1, 12, 1200$. Not all musical concepts have this invariance: it turns any Major chords into minor chords.

The complexity fonction has an acoustical interpretation. The spectrum of a tone (which constitutes its timbre) are ideally discrete such that each partial has a frequency which is an integer multiple of the fundamental frequency. Given 2 tones of commensurable fundamental frequencies $f_1, f_2 = \frac{a}{b} f_1$, their greater common divisor is $f_0 = \frac{f_1}{a} = \frac{f_2}{b}$. It is the highest frequency such that the 2 timbres/disrete spectra can be contained in a unique discrete spectrum. Partials frequencies coincide at $nbf_1 = naf_2 = nabf_0 \; \forall n \in \mathbb{N}$ and their is a clear correspondance between a high "density" of coincidences and a high consonance. Hence, the relevant measure of these coincidences is the interval between the first coincidence and the characteristic frequency f_0:

$$\frac{bf_1}{f_0} = ab = C\left(\frac{a}{b}\right)$$

This measure has the required invariance properties of being independent on musical transposition $f_1, f_2 \leftarrow cf_1, cf_2$ and on the inversion $\frac{f_1}{f_2} \leftarrow \frac{f_2}{f_1}$. The bigger this interval, the less coinciding the tones spectra, the bigger the dissonance.

1.4. The problem of defining a dissonance for a real sound

A major difficulty of modelling musical intervals by rational numbers consists in assigning it a realistic dissonance: what should we chose for $\frac{2001}{1000}$? It is a fairly perfect octave to the ear and something goes wrong when we assign it a very high complexity. This may be fixed by various means at the cost of introducing an arbitrary parameter. For instance, we can chose the best rational approximation within a maximum harmonicity or a relative uncertainty (see [5] for another inspiring example).

The next step to measure the dissonance of real sounds is to take into account all its simultaneous frequencies, which seems at first sight a complex algorithmic task. This is motivated by the observation that the timbre content of tones may change their

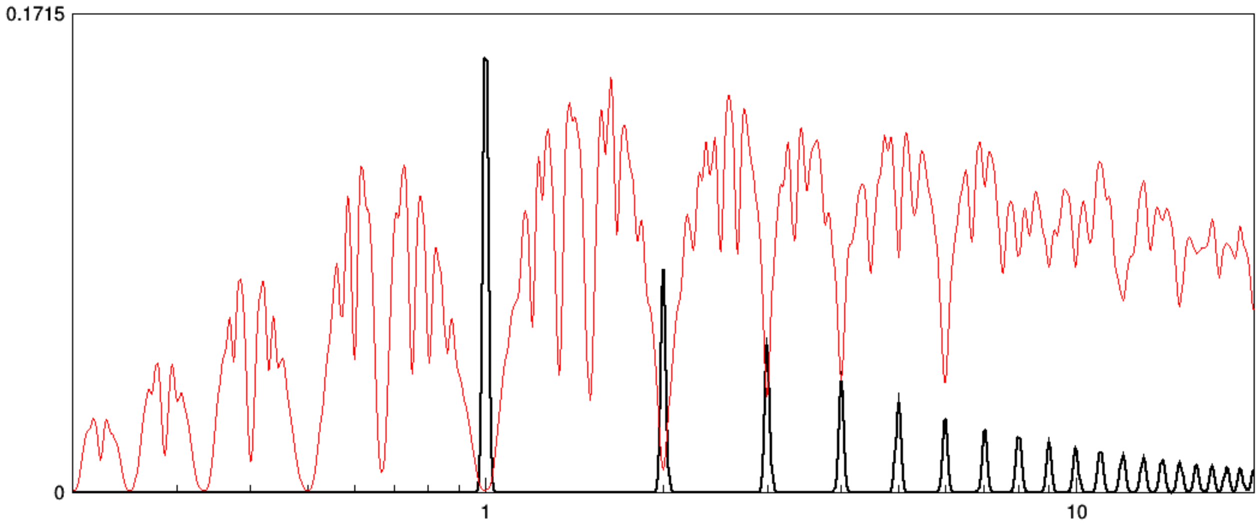

Figure 2: Arbitrary spectral density wich models the timbre of a musical sound (black) and its dissonance with a single mode/"pure tone" of varying frequency (red), with a logarithmic X-axis. The position of the unison, the octave, the fifth, etc are recognisable from the positions of the deep local minima.

dissonance ability [6]. The reduction of the representation of a musical sound to its fundamental frequency is then an approximation which is often justified by the perceptual fusion of the partials into the fundamental tone sensation (see Seebeck Ohm controversy and the missing fundamental paradox). We are then seeking such a dissonance for a continuous spectrum of ratio.

1.5. A perfect job for $𝄞(x)$

This task is fullfilled by the use of the musical function $𝄞(x)$ as the cumulative distribution function of a (singular) measure. Although the function has a diverging derivative at each irrationnal and vanishing at each rational number, it is nevertheless continuous everywhere and monotonically increasing from 0 to 1. The dissonance of a single frequency f with respect to an arbitrary sound (positive) spectrum S is defined as:

$$\mathcal{D}[S](f) = \int_0^\infty S(f')d𝄞\left(\frac{f'}{f}\right)$$

From this, we can derive the expression of the dissonance between two real sounds:

$$\mathcal{D}(S_1, S_2) = \int_0^\infty \mathcal{D}[S_1](f)S_2(f)\frac{df}{f} = \int_0^\infty \int_0^\infty S_1(qf)S_2(f)\frac{df}{f}d𝄞(q)$$

As a remark, we could try to study the topological space of sounds induced by the following distance-like function:

$$d(S_1, S_2) = -\log \frac{\mathcal{D}[S_1, S_2]}{\sqrt{\mathcal{D}[S_1, S_1]\mathcal{D}[S_2, S_2]}}$$

It is a "symmetric premetric" from the following properties:

$$d(S_1, S_2) = d(S_2, S_1) \quad ; \quad S_1 = S_2 \Rightarrow d(S_1, S_2) = 0$$

As derived in the appendix A, the convergent for the density function of the measure has the following asymptotic behaviour:

$$\frac{\Delta_H \text{𝄞}(x)}{\Delta_H \log x} \equiv \frac{\text{𝄞}(q_H^+) - \text{𝄞}(q_H^-)}{(q_H^+ - q_H^-)/\sqrt{q_H^+ q_H^-}} = \frac{\sqrt{C(q_H^+)C(q_H^-)}}{2^H} \quad \forall x \in \mathbb{R}^+$$

$$\sim C(x)\frac{H}{2^H} \quad \forall x \in \mathbb{Q}^+ \mid H \gg H(x)$$

Here, $q_H^\pm$ denote the closest (greater and lower) rational numbers of harmonicity H. Thus, the measure $d\text{𝄞}(x)$ inherits all the properties of the complexity function $C(x)$. The rational number q —which was modelling the interval between the fundamental frequencies of two tones— is replaced by function of the tones' full spectra of the form:

$$Q(q) = \int_0^\infty S_1(qf)S_2(f)\frac{df}{f}$$

plus an arbitrary parameter chosen to be a maximum harmonicity H.

2. General framework for the 𝄞-analysis

2.1. Discretisation of the measure

For all "𝄞-interesting" objects, there exists a function $Q(q)$ of a positive real number q, that we call the *spectrum of relations*, and that is reduced to a positive number when measured musically:

$$\text{𝄞}[Q] = \int_0^\infty Q(q)d\text{𝄞}(q)$$

The natural choice of discretization of this integral is the sequence of consecutive rational numbers of harmonicity bounded by H so that $\Delta_H \text{𝄞}(q) = 2^{-H}$. This avoids the algorithmic computation of $\text{𝄞}(x)$ (2^H times) and the measure up to H is simply:

$$\text{𝄞}_H[Q] = 2^{-H} \sum_{q \in \mathbb{Q}_H} Q(q) \xrightarrow[H \to \infty]{} \text{𝄞}[Q]$$
$$\text{where } \mathbb{Q}_H = \{q \in \mathbb{Q}^+ | H(q) \leq H\}$$

The implementation of this sum requires an explicit bijection between natural numbers and rational ones with a control over the harmonicity. This is provided by the Calkin-Wilf sequence of rational numbers, generated as follow:

$$q_{n+1} = \frac{1}{\lfloor q_n \rfloor + \lceil q_n \rceil - q_n} \quad ; \quad q_1 = \frac{1}{1}$$

For an efficient computation, this formula is decomposed into integer operations on the numerator and the denominator separately. All the rational numbers present in another binary tree, the Calkin-Wilf tree, are enumerated stage by stage. This tree is dual to our previous Stern-Brocot tree in the sense that the sequence of bits that encodes the path to reach a number in one tree has to be read backwards to reach the same number in the other tree. It follows that all rational numbers of harmonicity H are in the stage H of both trees (even though the numbers are not ordered in this new tree), so that the evaluation of the measure up to harmonicity H stops before computing $q_n = \frac{1}{H+1}$ for $n = 2^H$.

The convergence of the measure in the continuous limit for a regular enough $Q(q)$ (for instance interpolated from data) shows that the previously introduced arbitrary parameter is not always a feature of the dissonance assigning procedure: The final measure depends on the resolution of the ear (if the 𝄞-interesting object is a sound), of the spectrum and the intrinsic uncertainty of the phenomenon. When the ear is limiting, a model of perception requires a realistic H. When the data is limiting, the convergence of the measure gives the best estimate given the data. But when the intrinsic variability of the sound is limiting, then the computed dissonance is not only unique, it is also realistic.

2.2. Properties of the spectrum of relations

The spectrum of relations captures the relations between two "𝄞-related" objects through the generic form of a multiplicative cross-correlation:

$$Q_{1,2}(q) = \int_0^\infty S_1(x) S_2(qx) \frac{dx}{x}$$

where the objects S_1, S_2 can be thought as commensurably rich spectral densities (thus positive) that need to be defined depending on the context. Their argument must be an absolute quantity. For instance, a moment in time is a relative quantity, whereas a duration (between two moments) is an absolute quantity.

The spectrum of relations defined as above is usually not symmetric $Q_{2,1}(q) = Q_{1,2}(q^{-1})$. But the musical measure is symmetric, as required by the dissonance interpretation 𝄞$[Q_{1,2}] =$ 𝄞$[Q_{2,1}]$

When the compared object do not have their parameter scaled in the same way (eg. chromatic transposition, or change of tempo in Music) or even dimensionally different, we can replace $q \leftarrow \alpha q$. The value of α which minimises the musical measure gives the factor $\alpha_{\text{opt}} x_1 = x_2$ that compensates the unknown transposition/change of scale, and has the required dimension. This turns the musical measure into a multiplicative cross-correlation with the singular dissonance density, that we call the *dissonance curve*:

$$\text{𝄞}[Q_{1,2}](\alpha) = \int_0^\infty Q_{1,2}(\alpha q) d\text{𝄞}(q)$$

But it is not symmetric with respect to S_1, S_2: $\text{𝄞}[Q_{1,2}](\alpha) = \text{𝄞}[Q_{2,1}](\alpha^{-1})$

An important special case is when $S_1(x) = \delta(1-x)$. Indeed, it extracts the *dissonance potential* of the object $S_2(x)$, since consonance happens at local minima. We rebuild the dissonance curve of any real spectrum S_1 with S_2 by convolving S_1 with the dissonance potential of S_2:

$$S_1 * \text{𝄞}[Q_{\delta,2}](\alpha) = \text{𝄞}[S_1 * Q_{\delta,2}(\alpha)] = \text{𝄞}[Q_{1,2}](\alpha)$$

The convolution of a spectrum with its own dissonance potential is the musical measure of its multiplicative autocorrelation $Q(\alpha q)$:

$$Q(\alpha q) = \int_0^\infty S(q\alpha x) S(x) \frac{dx}{x}$$

then we will call $\text{𝄞}[Q](\alpha)$ the *self-dissonance curve* associated to the spectral density S, as explained in the following.

Note that the spectrum of self-relations have the invariance $Q(q) = Q(q^{-1})$, then $\text{𝄞}[Q] = ?[Q]$ the musical measure coincides exactly with the question mark's measure in this case.

Addition of two signals: $S_1 + S_2 \quad \Rightarrow \quad Q_{12,12} = Q_{1,1} + Q_{2,2} + Q_{1,2} + Q_{2,1}$ where $Q_{2,1}(q) = Q_{1,2}(q^{-1})$.

2.3. The canonical example of the musical harmony

Consider 2 sounds represented equivalently by their time signal $s_1(t), s_2(t)$ (t relative quantity) or by their Fourier transform $\tilde{s_1}(\omega), \tilde{s_2}(\omega)$ (ω absolute quantity). We would like to know whether they are musically related, therefore we prefer $\tilde{s_1}(\omega), \tilde{s_2}(\omega)$, fonction of an absolute quantity.

Note that the relativeness of the parameter in the direct (time) space translates as an arbitrary phase $e^{i\phi}$ which multiplies the function in the reciprocal (frequency) space. The absoluteness is recovered by considering only the norm or the square of the Fourier transform, so that our density-like positive functions of an absolute quantity are $|\tilde{s_i}(\omega)|^\nu$. Here we introduce ν to denote different conventions that will be discussed in various cases.

Finally, this has to be converted from an additive density to a multiplicative one:

$$|\tilde{s_i}(\omega)|^\nu d\omega = S_i(\omega)\frac{d\omega}{\omega} \quad \Rightarrow \quad S_i(\omega) = \omega|\tilde{s_i}(\omega)|^\nu \quad i = 1, 2$$

Then, the spectrum of relations is:

$$Q(q) = \int_0^\infty |\tilde{s_1}(\omega)\tilde{s_2}(q\omega)|^\nu q\omega d\omega$$

The convention $\nu = 1$ consists in taking the norm of the Fourier spectrum, whereas $\nu = 2$ corresponds to considering power spectral densities. At this stage, we can not see which choice describes best the musical consonance. We will see in the following several arguments in favour of the convention $\nu = 1$, the first of which is the physical dimension of $Q(q)$. As a generalisation of the relation (ratio) between two quantities, $Q(q)$ does not have a dimension, which means that $S(\omega)$ has no dimension neither. Considering that the signal $s(t)$ is normalised (by the square root of its "energy" time density) and has thus no dimension, the time dimension of its Fourier transforms is compensated by the frequency factor only for $\nu = 1$. This observation provides a useful constraint on the construction of other spectrum of relations.

2.4. Simplified notation and summary

All the previous formula can be rewritten more simply by noticing that the previous expressions are more familiar after a multiplicative to additive change of variables: they are convolutions and cross-correlations. Let us make clear this change of variables by defining what is sometimes called the Lamperti transform:

$$\begin{aligned} \mathcal{L}_a : \; & L(\mathbb{R}) \to L(\mathbb{R}^+) \\ & g(u) \mapsto h(x) = x^a g \circ \log(x) \end{aligned}$$

$$\begin{aligned} \mathcal{L}_a^{-1} : \; & L(\mathbb{R}^+) \to L(\mathbb{R}) \\ & h(x) \mapsto g(u) = e^{-au} h \circ \exp(u) \end{aligned}$$

where $L(\mathbb{K})$ is an unspecified space of functions or distributions defined on $\mathbb{K}$. The exponent $a = -1$ is here used to take into account the Jacobian when the variable of a probability distribution function is changed. Otherwise $\mathcal{L}$ will simply denote $\mathcal{L}_0$.

This means that the functions defined on positive real numbers must be composed with an exponential function to extend its support to all real numbers and use the regular cross-correlation. **Let us slide a bit the notation in this way $S(x) \leftarrow S(2^x)$.** The base 2 exponential is natural in a musical context since it turns the multiplicative view of frequency ratio into an additive "tempered" fraction of the octave. For instance, the fifth (which is the seventh among the twelve half-tones forming an octave) has this remarkable and fundamental coincidence:

$$\frac{3}{2} \approx 2^{\frac{7}{12}}$$

Then, a spectrum of relations is the cross-correlation of 2 𝄞-interesting objects $Q = S_1 \star S_2$. And the interval dependent 𝄞-mesure (dissonance curve) can be formally written as a cross-correlation as well:

$$\text{𝄞}[Q](\alpha) = Q \star \text{𝄞}'(\alpha)$$

$$\text{𝄞}[Q] = Q \star \text{𝄞}'(0)$$

Note that the null interval $\alpha = 0$ in the additive setting corresponds to 1 in the multiplicative one. This particular value of the cross-correlation can be interpreted as a scalar product or a projection of the spectrum of relations on the musical —wild and singular— distribution. The essence of the musicality / commensurability is this distribution which can be expressed explicitly in term of Minkowski's question mark function as follows (keeping track of the successive transformations):

$$\text{𝄞}'(x) = \frac{d?\left(\frac{2^x}{2^x+1}\right)}{dx}$$

which takes an infinite value when 2^x is irrational and an infinitesimal value when it is rational. The most general of the above formula is the mutual dissonance curve between the two objects S_1, S_2:

$$(S_1 \star S_2) \star \text{𝄞}'(\alpha) = (S_2 \star S_1) \star \text{𝄞}'(\alpha) \quad \text{since} \quad \text{𝄞}'(\alpha) = \text{𝄞}'(-\alpha)$$

The (self-)dissonance curve associated to sound S is:

$$(S \star S) \star \text{𝄞}'(\alpha)$$

where $(S \star S) \star \text{𝄞}'(0)$ is the intrinsic dissonance of S.

The dissonance potential of S_1 is:

$$(\delta \star S_1) \star \text{𝄞}'(\alpha) = S_1 \star \text{𝄞}'(\alpha)$$

which is used as follows:

$$S_2 * (S_1 \star \text{𝄞}')(\alpha) = (S_1 \star S_2) \star \text{𝄞}'(\alpha)$$

3. Dissonance of pure tone, rich timbre sounds and coloured noise

The Fourier spectrum of a pure tone is a Dirac delta, so that $Q(q) = \int_0^\infty \delta(\omega - 1)\delta(q\omega - 1)q\omega d\omega = q\delta(q-1) = \delta(\log q)$. This the most consonance relation, which scales as $\text{𝄞}[\delta(\log q)] \sim H2^{-H}$, where H is the highest possible harmonicity of the approximation, chosen to correspond to the width of the sigle peak around $q = 1$.

For a timbred sound of parameter σ, we get:

$$Q(q) \propto \sum_{m,n=1}^{+\infty} (mn)^{-\sigma}\delta(qm - n)qm \propto \delta\big(\log q - \log \frac{a}{b}\big)$$

Therefore the dissonance is of order $\text{𝄞}[Q] = H2^{-H} \sum_{q \in \mathbb{Q}_H} C(q)^{1-\sigma}$, which is of order $H^2 2^{-H}$ at $\sigma = 2$ and H at $\sigma = 1$. This means that very energetic timbres $\sigma < 1$ have much more dissonance that weaker timbres $\sigma > 1$.

The Fourier spectrum of a coloured noise of parametre $c = 2\sigma/\nu$ is a power law over a large domain $|\tilde{s}(\omega)|^2 \sim \omega^{-c}$. Therefore, the spectrum of relations has the form:

$$Q(q) \simeq \int_{\omega_0}^{\infty} |\tilde{s}(\omega)\tilde{s}(q\omega)|^{\nu} q\omega d\omega \propto q^{1-\sigma}$$

We can fix the proportionality constant to 1 by requiring $Q(1) = 1$. The dissonance is then the integral transform of the musical distribution:

$$\text{𝄞}[Q] = \int_0^{\infty} q^{1-2\sigma} d\text{𝄞}(q)$$

This transform can be understood as the moment of order $2\sigma - 1$ of 𝄞 (Mellin transform) or as the Laplace transform of 𝄞$\circ$exp. From the symmetry of 𝄞, the less dissonant noise is therefore $\sigma = \frac{1}{2} \Rightarrow c = \frac{1}{\nu}$, with a dissonance equal to 1 (normalisation of 𝄞). This corresponds to the pink noise in the convention $\nu = 1$. This result favours the convention $\nu = 1$ because the pink noise is defined such that the amount of energy per octave is constant ; it is then the good reference for spectral densities in octave unit. The white noise $\sigma = 0$ has the dissonance equal to the first moment of 𝄞, which is known to be $\frac{3}{2}$.

4. Study of rhythm

4.1. What is a rhythmic signal

The construction of the spectrum of relations is identical. Assume that we have a rhythmic signal $r(t)$, positive function of a relative quantity (time parameter). Then:

$$Q(q) = \int_0^{\infty} |\tilde{r}(\omega)\tilde{r}(q\omega)|^{\nu} q\omega d\omega$$

The question is rather: how to build the (slowly fluctuating) rhythmic profile $r(t)$ from a raw (rapidely fluctuating) musical signal $s(t)$? We require its value to be causal at all time (no influence of future signal on the present value of the rhythmic profile) and to be increased by any frequency change. We propose to build it from a contrast function, which maps the frequency changes at all time using a causal wavelet capturing about n oscillations (show derivation):

$$\Xi_n(t,\omega) = \int_0^{\infty} s(t)s(t-\tau)\cos(\omega\tau)e^{-\frac{\omega\tau}{2\pi n}} \frac{\omega d\tau}{2\pi n}$$

where n is a scale parameter. Then the rhythmic signal is:

$$r(t) = \int_0^{\infty} |\Xi_n(t,\omega)|^{\mu} \frac{d\omega}{\omega}$$

The absolute values means that negative contrasts (caused by fading tones) can not compensate the contrast of an appearing tone and even participate positively to the rhythmic signal. Note that it has no dimension, a scale parameter and we might tune the contrast influence from an exponent μ.

Other possibilities if we forget the deterministic windowing function: use Wigner distribution...or other wavelet decompositions.

4.2. Pulse model

Because of uncertainties relations between the time and frequency domains, sharp tone changes will be a bit smoothed. We can nevertheless represent them as peaked functions using Dirac delta functions as a first approximation.

The trivial case is the single pulse, which is entirely analogous to the pure tone. Consider the extrem case of an absolutly regular sequence of pulses:

$$r(t) \propto \sum_{n=-\infty}^{+\infty} \delta(t-n)$$

$$Q(q) \propto \sum_{m,n=1}^{+\infty} \delta(qm-n)qm \propto \sum_{\frac{a}{b}\in\mathbb{Q}_N} \delta(\log q - \log \frac{a}{b})$$

It is analogous to a sound that would have a timbre parameter $\sigma = 0$ (unphysical). From this, it is obvious that for $Q(1) = 1$ we have a dissonance $\text{𝄞}[Q] = 0$. We will therefore call it consonant. To compare this spectrum of relations with other consonant ones, we may compute it using the convergent formula:

$$\text{𝄞}_H[Q] \sim H2^{-H} \sum_{\frac{a}{b}\in\mathbb{Q}_H} ab$$

The opposite case would be an absolutly random sequence of pulses, which is best modeled by a Poisson process. Indeed, probability to find a pulse per unit time interval is constant and the pulses are independent. Then, the expected autocorrelation of a large sequence of such pulses is proportional to the integral for all time t of the probability to find a pulse at time $t+\tau$ given there is a pulse at t, which is flat for all non zero τ. Then $Q(q) = q$ once normalised, which yields the dissonance $\text{𝄞}[q] = \frac{3}{2}$. This makes it analogous to the white noise case in the study of noise dissonance.

$$R(t) \propto \sum_{n=-\infty}^{+\infty} \delta(t-T_n) \quad \text{where } \mathbb{P}(T_n \in [t;t+dt]) \propto dt$$

5. Study of musical scales

We claim that the Fourier transform $\mathcal{F}[(S \star S) \star \text{𝄞}'] = |\tilde{S}|^2 \cdot \widetilde{\text{𝄞}'} \equiv \text{𝄢}$ is useful to classify all the tempered scales. 𝄢 extracts the periodicity behaviour of the dissonance curve in the additive view. Then we expect 12 (notes per octave) to be a local minimum of this function, reflecting of a certain universality of the 12 tones tempered scale. Interestingly, the maxima would correspond to the less musical/consonant tempered scales. As a consequence, we can call this object the *musical scale potential*, specific to the instrument whose timbre is described by S.

5.1. The timbre-less description

A first approximation consists in considering single frequency sounds, in which case $|\tilde{S}|^2$ is constant. This is a good approximation of sounds with poor timbre (fast decaying partials) and also the only model which is independent of the characteristic of the sound S. Thus, the study will just consists in evaluating numerically $\widetilde{\text{𝄞}'}$.

Since the musical distribution is symmetric, its Fourier transform is real:

$$\widetilde{\text{𝄞}'}(n) = \int_{-\infty}^{+\infty} e^{-i2\pi nx}\,\text{𝄞}'(x)dx = \int_{-\infty}^{+\infty} \cos(2\pi nx)\,\text{𝄞}'(x)dx = \int_{0}^{1} \cos(2\pi n \log_2 y) d?(y)$$

Among different possibilities for discretising it, the straightforward Riemann sum from the last expression gives the best result (others lead to artifacts or aliasing).

This yields a very interesting result. First we notice that the spectrum of this wild function converges to a smooth curve. Then the normalisation is verified by the value 1 on the Y-axis. Finally, the wells of this potential turns out to deserve its musical quality: we observe large negative peaks around 5, 7 and 12 which are the most common number of notes per octave. Furthermore, the width of the peaks contains an important information: peaks at 5 and 7 are wide compared to 12, which reflects that harmonious scales with 5 or 7 notes are irregular (eg. with tones and semi-tones for the heptatonic scales), whereas the really thin peaks at 12 expresses a very regular scale (typically tempered, hence strictly regular in a modern fashion).

We also observe higher large negative peaks, which corresponds to known microtonal scales. Conversly, we deduce that dissonant scales are indicated by large positive peaks: for instance we can produce especially dissonant 6-notes scales.

However, the position of the peak for the dodecaphonic scale is slightly shifted $n \simeq 12.156$. Considering a regular scale of elementary addition interval $\frac{\log 2}{12.156}$, we get a better tuning for the major thirds but the other intervals are going out of tune. This optimum might be perceptually true for pure frequencies, nevertheless this suggests to include a timbre in our model to account for the common musical experience.

5.2. Power-law harmonic timbres or the musical approach to Riemann Hypothesis

Let us now consider the correction $|\tilde{S}|^2$ for real musical sounds. Assume first that the sound is stationnary, more specifically its signal $s(t)$ is assumed to be periodic (with period 1) so that its spectrum is discrete and regular. Then, the simplest model is a polynomial decay of the intensities of the partials:

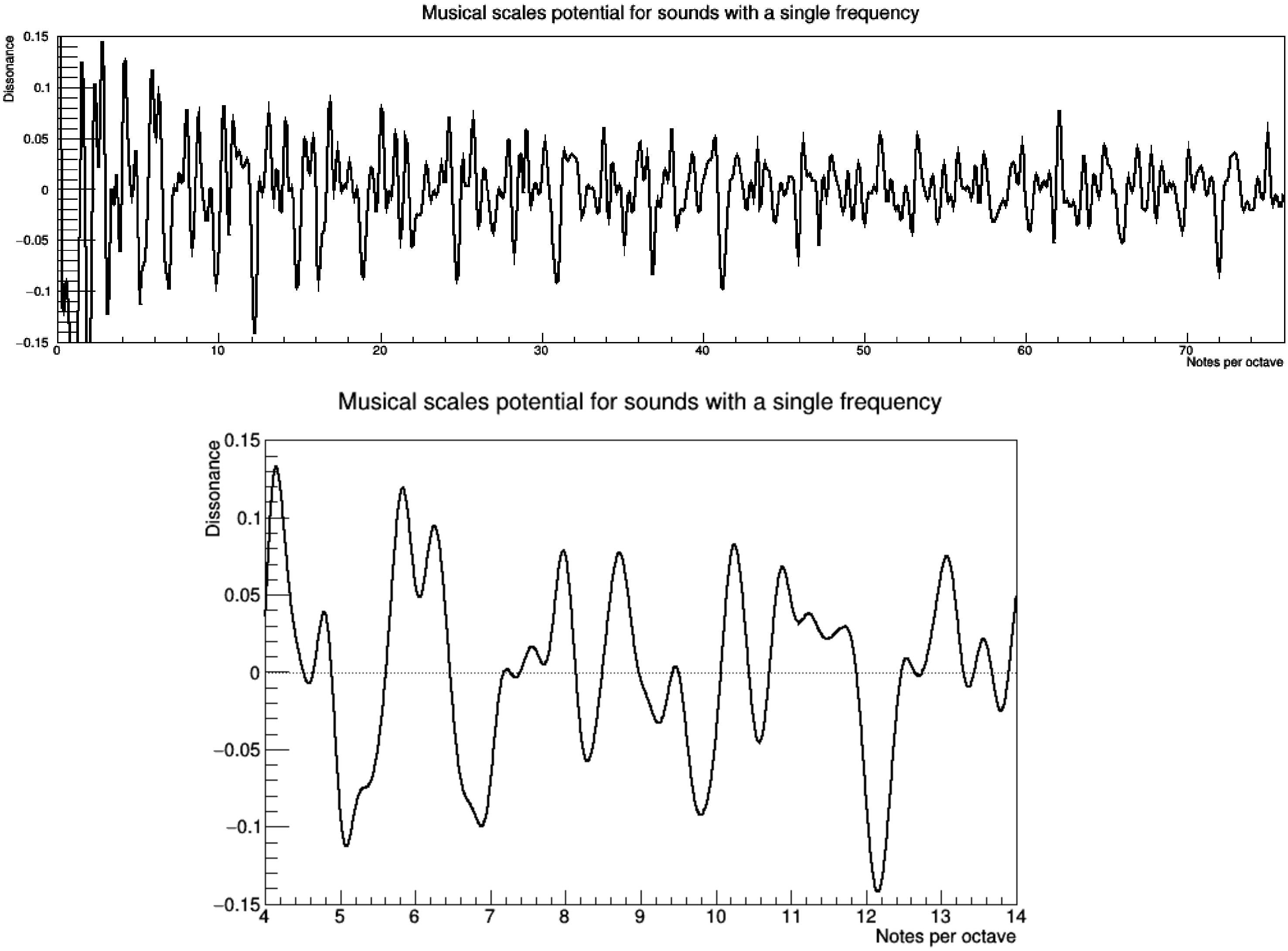

Figure 3: Plots of $\widetilde{\text{𝄞}}'(n)$.

$$
\begin{aligned}
\tilde{s}(f) &\propto \sum_{m=1}^{\infty} m^{-\sigma/\nu} e^{i\phi_m} \delta(f-m) \\
\tilde{s}_m &\propto m^{-\sigma/\nu} e^{i\phi_m} \quad \text{so that} \quad |\tilde{s}_m|^{\nu} = m^{-\sigma}
\end{aligned}
$$

where the first formula is the Fourier spectrum of an infinitly long signal, whereas the second formula are the coefficients of the Fourier series over on perdiod. Note that the negative part of the spectrum is not considered here because it is symmetric (in a Hermitian way). A finite energy of the sound requires $\sigma > \frac{\nu}{2}$ (the signal is square integrable):

$$
\begin{aligned}
E = \sum_{m=1}^{\infty} |\tilde{s}_m|^2 &= \sum_{m=1}^{\infty} m^{-2\frac{\sigma}{\nu}} < +\infty \\
&\propto \int_{-\infty}^{+\infty} |\tilde{s}(f)|^2 df
\end{aligned}
$$

This suggests to normalise this typical sound spectrum with the prefactor $\zeta(2\sigma/\nu)^{-\frac{1}{2}}$ (times the large number $\delta(0)$ in the infinite time formulation), introducing the Riemann zeta function. In the following, we will adopt the convention $\nu = 1$.

In the additive view (in octave unit), define $S(x)dx = |\tilde{s}(2^x)|d2^x$, hence:

$$
\begin{aligned}
S(x) &= \zeta(2\sigma)^{-\frac{1}{2}} \sum_{m=1}^{\infty} m^{-\sigma} \delta(x - \log_2 m) \\
S \star S(y) &= \zeta(2\sigma)^{-1} \sum_{n,m=1}^{\infty} (nm)^{-\sigma} \delta\big(y - \log_2 \frac{n}{m}\big) \\
\tilde{S}(n) &= \int_{-\infty}^{+\infty} S(x) e^{-i2\pi n x} dx \\
&\propto \sum_{m=1}^{\infty} m^{-\sigma - i\frac{2\pi}{\log 2} n} = \zeta\big(\sigma + i\frac{2\pi}{\log 2} n\big)
\end{aligned}
$$

Therefore we get the following expression for the musical scale potential:

$$
\text{𝄢}(n, \sigma) = \frac{\big|\zeta\big(\sigma + i\frac{2\pi}{\log 2} n\big)\big|^2}{\zeta(2\sigma)} \widetilde{\text{𝄞}'}(n)
$$

The Fourier transform of the spectrum of relations is juste 1 in the pure tones case and is expressed in terms of the zeta function. In the timbre-less limit, we recover:

$$
\text{𝄢}(n, +\infty) = \widetilde{\text{𝄞}'}(n)
$$

Note that we think of the timbre parameter σ in terms of energy distribution among partials, so that $2\sigma = \frac{1}{T}$ is an inverse temperature. The low temperature limit is the pure tone (all the sound energy is contained in the fundamental frequency) and the energy of the sound diverges at a critical temperature $T_c = 1 \Leftrightarrow \sigma = \frac{1}{2}$. This critical value corresponds to the *critical line* of the Riemann zeta function, which is famously conjectured to contain all the non-trivial (non-real) zeros of ζ.

We can try to translate this conjecture in musical terms:

Musical formulation of the RH: *We do not expect to change qualitatively the musical scales dissonance when we increase the energy of the timbre, but only quantitatively.* The quality of the scales, which is the sign of their potential (positive for dissonant, zero for neutral and negative for consonant) is entirely contained in the pure tones potential $\widetilde{𝄞}'(n)$.

For physical sounds with a timbre $\sigma > \frac{1}{2}$, the correction — which modulates quantitatively the musical scales property— is strictly positive: $|\zeta(\sigma + it)|^2/\zeta(2\sigma) > 0$. This excludes zeros of ζ with a real part greater than $\frac{1}{2}$ (which exludes them also from $\sigma < \frac{1}{2}$ using a symmetry in the functional equation of the zeta function). The energy divergence yields an intuitive reason for the criticality at $\sigma = \frac{1}{2}$, but note that the correction is identically zero $\forall t$ at $\sigma = \frac{1}{2}$. That means that consonant and dissonant scales are less and less distinguishable as we approach the critical sound of diverging energy.

Note that this agrees with the soundness of the convention $\nu = 1$, which has been now supported in various contexts.

Let $t = \frac{2\pi}{\log 2}n$. The asympotic behaviour of the correction for a poor timbre is:

$$\Big|\sum_{m=1}^{\infty} m^{-\sigma-it}\Big|^2 = \Big|\sum_{m=1}^{\infty} \frac{\cos(t\log m)}{m^\sigma} - i\sum_{m=1}^{\infty} \frac{\sin(t\log m)}{m^\sigma}\Big|^2$$
$$\sim 1 + \cos(2\pi n)2^{1-\sigma} \quad \text{as} \quad 1 < \sigma \longrightarrow +\infty$$

Thus, the presence of a timbre will favour scales with an integer number of notes per octave 2^n (asymptotic maximum) at leading order (and the next orders are favouring just intervals through reinforcments at m^n, $m = 3, 4, 5$, etc). As expected, the large negative peaks of the potential 𝄢 which corresponds to dodecaphonic scales will get closer to $n = 12$ than in the timbre-less model.

That correction can be easily estimated numerically as long as $\sigma > 1$. In the region $\frac{1}{2} < \sigma < 1$ where the potential has been regularised by analytical continuation, the estimation is much more difficult. Even though, the literature is very prolific for the critical sound $\sigma = \frac{1}{2}$ where the correction is called the Riemann-Siegel Z-function (squared):

$$\Big|\zeta\big(\frac{1}{2} + it\big)\Big|^2 = Z(t)^2$$

where $t = \frac{2\pi}{\log 2}n$ keeps track of the musical interpretation.

It has already been discovered and rediscovered by different means that the Z-function has some interest in classifying musical scales [7] [8]. The last reference defines the real part of $Z(t)$ as an estimate of the amount of coincidences. The imaginary part happens to vanish close to the maxima of the real part, called Gram points, which are themselves close to integer multiples of $\frac{2\pi}{\log 2}$. This confirms, on the critical line, the property of stabilising the optimal scales close to an integer number of notes per octaves. Interestingly, it is also presented as being sufficient to find the best scales.

The numerical estimation of 𝄢 $(n, \sigma = 1)$ illustrates the radical effect of the timbre over the musical scale potential. The dissonant scale peaks are tamed whereas the musically interesting ones are much more peaked and close to an integer number of notes per octave.

The optimum of the dodecaphonic scale is situated arround $n = 12.0568$, which is more realistic than the $\sigma = +\infty$ model. Even though this choice for an equal temperament scale still looks a bit awkward, we notice that it decreases the mistuning of the fourths and of the Major thirds compared to the $n = 12$ equal temperament. However, this potential should not be understood as characterising strictly equal temperament scales but rather as a first approximation of the potential for all scales with a given density of notes.

Beyond the classical scales $n = 5, 7, 12$, the observation of large negative peaks suggests that the scales densities $n = 19, 31, 41, 72, ...$ are also very good. We can also notice the following relations : $5 + 7 = 12$, $7 + 12 = 19$, $12 + 19 = 31$, as well as $10 + 12 = 22$, $19+22 = 41$ and $31+41 = 72$ (where $n = 10, 22$ are smaller negative peaks). This suggest that scales of higher densities are simply composed of the superposition of simpler scales. For instance, the 5 notes of the pentatonic scales (the black keys on a piano keyboard) are intercalated between the 7 notes of heptatonic (major or natural minor) scale (the white keys of the keyboard) ; all of them form the chromatic scale with 12 notes per octave. This can inspire a recursive scale construction method in order to generalise the harmony to a microtonal solfeggio:

- Given the structure of the 5-scale and the 7-scale embedded in an equally tempered 12-scale, we perturb it to suit the thiner equal note spacing of the 19-scales
- We find the best possible choice (or two symmetric best choices) such as the complementary notes also has the structure of an interersting lower density scale. (See the scale phase plot later to help doing this by hand.)
- Iterate by perturbing the regular structure of the 19-scale to embed it inside the equal temperament 31-scale.
- For this, we possibly need to find the structure of the complementary scale (22-scale) by embedding in it lower density scales with a known structure.

In conclusion, the observation of the positions of the deeper peaks may be sufficient to find interesting approximate scale structures. From this recursive method, we could deduced the best subscales irregular structures within a background equal temperament scale without additional extensive optimisation computations. The trade-off between the translation possibility allowed by the equal temperament scales and the higher consonance in some irregular scales take the following form: the chose background scale is

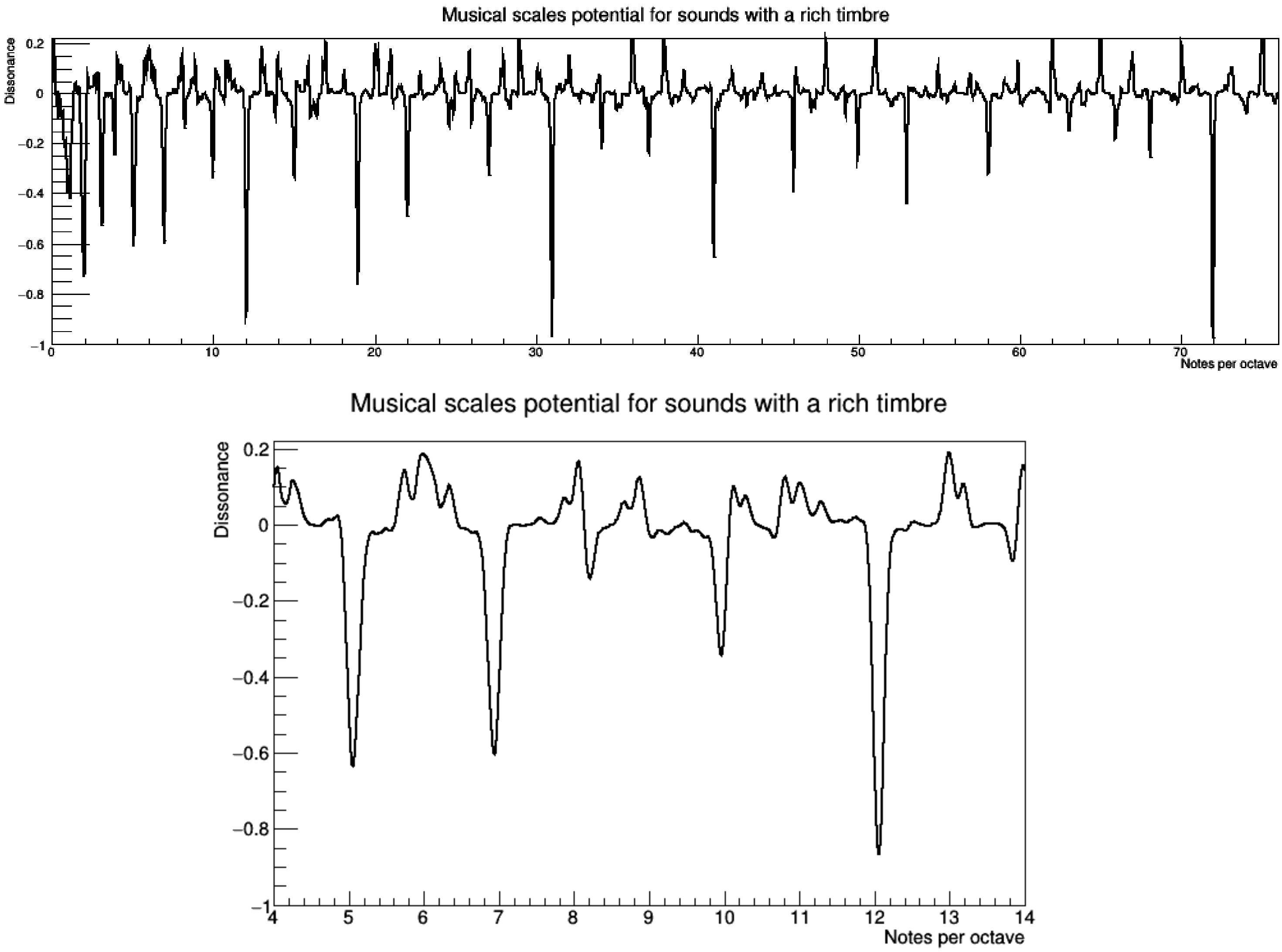

Figure 4: Plots of 𝄢 $(n, 1)$.

always equally tempered, but the sub and subsubscales are less and less tempered but more and more consonant.

5.3. Clarinet-like timbres

The timbre of a clarinet does not have any even partial. This odd spectrum is modelled easily with the zeta function:

$$\sum_{m=1}^{\infty}(2m-1)^{-s} = \sum_{m=1}^{\infty} m^{-s} - \sum_{m=1}^{\infty}(2m)^{-s} = (1-2^{-s})\zeta(s)$$
$$\Rightarrow \quad \tilde{S}_{cla}(n) = \big((1-2^{-2\sigma})\zeta(2\sigma)\big)^{-\frac{1}{2}}\big(1-2^{-\sigma-i\frac{2\pi}{\log 2}n}\big)\zeta\Big(\sigma+i\frac{2\pi}{\log 2}n\Big)$$

The sound of a clarinet compares to other full timbres as follows:

$$\frac{𝄢_{cla}(n,\sigma)}{𝄢(n,\sigma)} = \frac{|1-2^{-\sigma-i\frac{2\pi}{\log 2}n}|^2}{1-2^{-2\sigma}} = \frac{1-2^{1-\sigma}\cos(2\pi n)+2^{-2\sigma}}{1-2^{-2\sigma}}$$

The influence of the prime number 2 is removed in this way, and it can be generalised to higher prime numbers (although the resulting timbre is not natural). We notice that the timbre of the clarinet reduces a bit the consonance of the common scales ($n = 5, 7, 12, ...$). However, it improves the consonance of scales with a non-integer number of notes per octave but an integer number of notes per "tritave" $n' = \frac{\log 2}{\log 3}n = 13, 26, 39, ...$ The first of these scales —13 notes per tritave— is known as Bohlen-Pierce scale for clarinets (with a bit more that 8 notes per octave), and the numerical evaluation of the potential suggests to divide its elementary interval in 2 or 3 to improve the overall consonance ($n' = 26$ or 39).

5.4. Piano-like anharmonic timbres

Vibrating strings have some anharmonicity, which has to be taken into acount for instruments with thin strings such as piano or harp. This yields partials with a frequency $f_n = n\sqrt{1+\epsilon n^2} f_1$ where ϵ is a small coefficient (see Wikipedia). A perturbative analysis could be achieved.

The simple observation of the resulting scales potential reveals that the consonance of high order microtonal scales ($n = 31, 41, 72, ...$) vanishes with an increasing anharmonicity, whereas low order/common scales are robust. Thus, the degree of anharmonicity of an acoustical instrument limits the order of the possible microtonal scales.

An increasing shift of the position of the potential peaks towards lower values suggests a slight dilatation of the intervals, in accordance with the specificity of the piano tuning. Though, the exact structure of the tuning is not specified in this analysis.

5.5. A potential for exactly equal temperament?

The previous results might be refined when we impose a fixed elementary interval (equal temperament): the dissonance curve is no more projected on the sloppily tempered cosine potential (leading to Fourier transform of the dissonance curve ie. the musical scales potential), but on a sharp "Dirac comb" distribution $\sum_k \delta(x-kx_0) = 1/x_0 \sum_k e^{i2\pi kx/x_0}$:

$$\int_{-\infty}^{+\infty} (S\star S)\star \text{𝄞}'(x) \sum_{k=-\infty}^{+\infty} \delta\Big(x-\frac{k}{n}\Big)dx = \sum_{k=-\infty}^{+\infty} (S\star S)\star \text{𝄞}'\Big(\frac{k}{n}\Big)$$
$$= n\sum_{m=-\infty}^{+\infty} |\tilde{S}|^2 \cdot \widetilde{\text{𝄞}'}(mn)$$

where x is in octave unit, hence $x_0 = \frac{1}{n}$. Since the dissonance curve is even and the $m=0$ is irrelevant for a potential (constant), the equal temperament scales potential is:

$$n\sum_{m=1}^{\infty} \text{𝄢}\,(mn,\sigma)$$

A cut-off for high m means a "thick comb", which could result from hearing limitation (duration, resolution), whereas a high k cut-off stands for the range of the instrument.

The resulting potential is highly fluctuating and slowly (or probably not) converging for several notes per octaves which makes it hard to interprete in this region. However, it has a deep minimum at one note per octave, 2 secondary peaks at one note per tritave and per fifth, and we recognise smaller peaks at the intervals with higher harmonicities. Although interesting scales are hardly readable from this curve (suggesting that equal temperament is far from the most consonant scale structure), we can draw from it the good periods for building scales.

5.6. Structure of periodic scales

As observed previously, the octave is good period for a scale, as well as the tritave or the fifth (the last of which is indeed used for piano and orchestra and known as Cordier temperament). We can use a similar analysis introducing a phase in the periodic distribution to distinguish the most consonant tones:

$$\int_{-\infty}^{+\infty} (S\star S)\star \text{𝄞}'(x) \sum_{k=-\infty}^{+\infty} \delta(x-(k+\phi)x_0)dx = \sum_{k=-\infty}^{+\infty} (S\star S)\star \text{𝄞}'((k+\phi)x_0)$$
$$= \frac{1}{x_0}\sum_{m=-\infty}^{+\infty} |\tilde{S}|^2 \cdot \widetilde{\text{𝄞}'}\Big(\frac{m}{x_0}\Big)e^{i2\pi m\phi}$$

This yields the following object, that we can call the periodic scale structure:

$$\text{𝄡}(\phi,b,\sigma) = \frac{1}{\log_2 b}\sum_{m=1}^{\infty} \text{𝄢}\,\Big(\frac{m}{\log_2 b},\sigma\Big)\cos(2\pi m\phi)$$

where we have set $x_0 = \log_2 b$ for a "b-tave", and the phase is $\phi \in [0, 1[$. The previously studied period potential is thus $\text{𝄡}(0, b, \sigma)$.

5.7. Unifying formalism for these series

We can get neater expressions in the form of series over all positive rational numbers (that we denote here simply $\mathbb{Q}$):

$$
\begin{aligned}
S(x) &= \zeta(2\sigma)^{-\frac{1}{2}} \sum_{m=1}^{\infty} m^{-\sigma} \delta(x - \log_2 m) \\
\Rightarrow \quad S \star S(y) &= \zeta(2\sigma)^{-1} \sum_{n,m=1}^{\infty} (nm)^{-\sigma} \delta\big(y - \log_2 \frac{n}{m}\big) \\
&= \sum_{q \in \mathbb{Q}} C(q)^{-\sigma} \delta(y - \log_2 q)
\end{aligned}
$$

where $C(q) = ab$ is the complexity of the rational number $q = \frac{a}{b}$ in its reduced form. The normalisation cancels out with the multiple counting caused by non-reduced forms, underlying the relevance of the energetic normalisation choice. This remark is summarised in the following formula:

$$
\sum_{q \in \mathbb{Q}} C(q)^{-\sigma} = \frac{\zeta(\sigma)^2}{\zeta(2\sigma)} = \text{𝄢}\,(0, \sigma) \quad \forall \sigma > 1
$$

which extends to complex values as follows:

$$
\sum_{q \in \mathbb{Q}} C(q)^{-\sigma} \cos(t \log q) = \sum_{q \in \mathbb{Q}} C(q)^{-\sigma} q^{it} = \frac{|\zeta(\sigma + it)|^2}{\zeta(2\sigma)}
$$

Justification:

$$
\begin{aligned}
|\zeta(\sigma + it)|^2 &= \left(\sum_{m=1}^{\infty} m^{-\sigma - it} \right) \left(\sum_{n=1}^{\infty} n^{-\sigma + it} \right) \\
&= \sum_{m,n=1}^{\infty} (mn)^{-\sigma} \left(\frac{n}{m} \right)^{it} \\
&= \sum_{m,n=1}^{\infty} (mn)^{-\sigma} \cos\left(t \log \frac{n}{m} \right) \\
&= \sum_{q \in \mathbb{Q}} \sum_{l=1}^{\infty} \big(C(q) l^2 \big)^{-\sigma} \cos(t \log q) \\
&= \zeta(2\sigma) \sum_{q \in \mathbb{Q}} C(q)^{-\sigma} \cos(t \log q)
\end{aligned}
$$

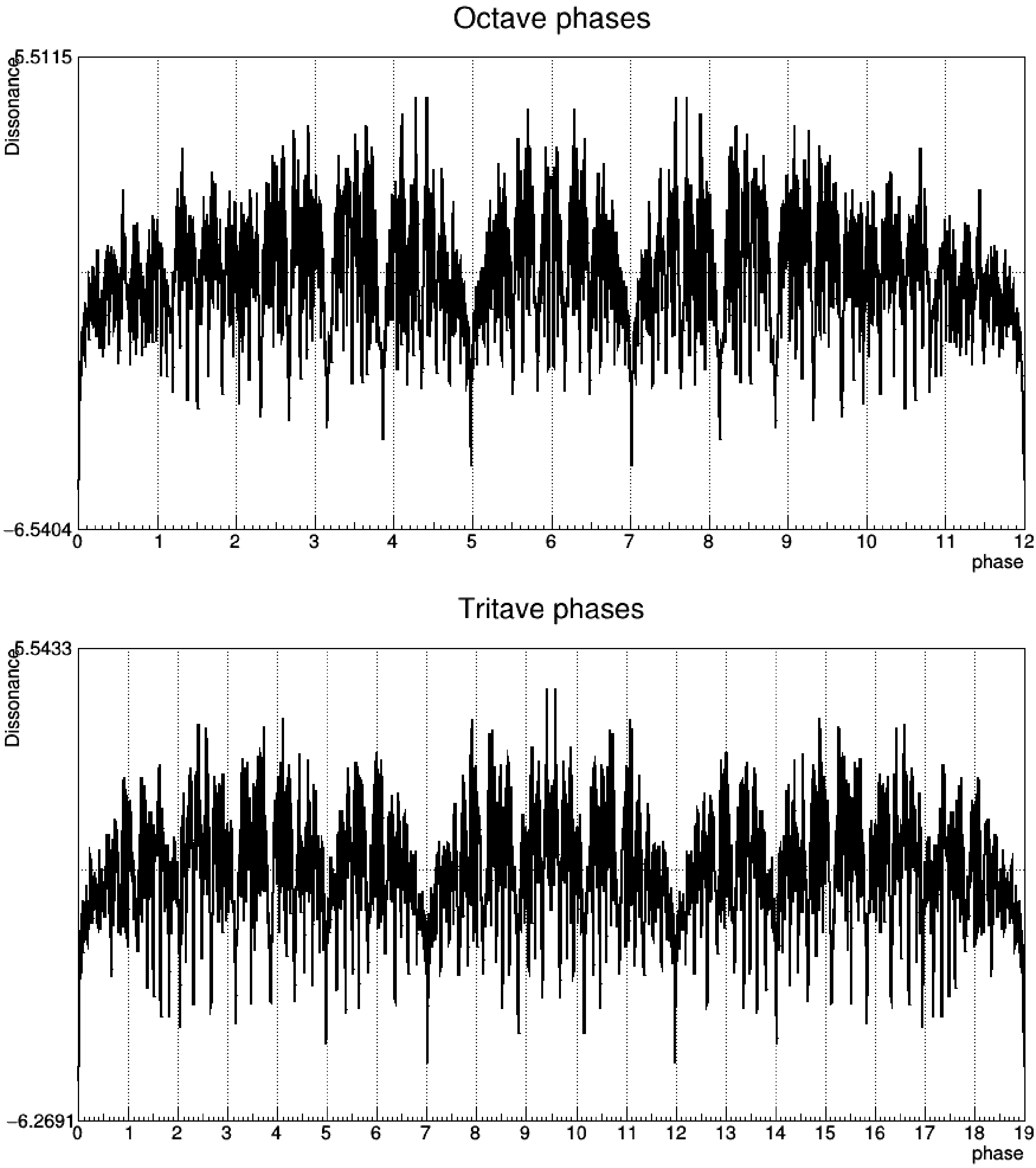

Figure 5: Plots of 𝄡$(\phi, 2, 1)$ and 𝄡$(\phi, 3, 1)$ against 12ϕ and 19ϕ respectively, with a cut-off at $m = 1000$. In this way, we display the proximity of the tones in the $n \simeq 12$ equal temperament with the most consonant phases for the octave and tritave period.

where we have considered $n = al$, $m = bl$, $q = \frac{a}{b}$ irreducible fraction and $C(q) = ab$.

$$\begin{aligned}
\widetilde{\text{𝄞}'}(n) &= \int_{-\infty}^{+\infty} \cos(2\pi n x) d\text{𝄞}(x) \\
&= \lim_{H\to\infty} \sum_{q\in\mathbb{Q}|H(q)=1}^{H} \cos(2\pi n \log_2 q)\big(\text{𝄞}(q_H^+) - \text{𝄞}(q_H^-)\big) \\
&= \lim_{H\to\infty} 2^{-H} \sum_{q\in\mathbb{Q}|H(q)=1}^{H} \cos(2\pi n \log_2 q)
\end{aligned}$$

where $q_H^{\pm}$ denote the two neighbouring rational numbers of q of harmonicity $H+1$. Note that this is not proportional to $\frac{|\zeta(i\frac{2\pi}{\log 2}n)|^2}{\zeta(0)}$, despite the series matching. Indeed, this last formula is famously known to have no root whereas $\widetilde{\text{𝄞}'}(n)$ has many zero crossings between consonant and dissonant scales.

This allows us to write the following expression for the scale potential:

$$\begin{aligned}
\text{𝄢}(n,\sigma) &= \frac{\left|\zeta\left(\sigma + i\frac{2\pi}{\log 2}n\right)\right|^2}{\zeta(2\sigma)} \widetilde{\text{𝄞}'}(n) \\
&= \lim_{H\to\infty} 2^{-H} \sum_{p,q\in\mathbb{Q}|H(q)=1}^{H} C(p)^{-\sigma} \cos(2\pi n \log_2 p) \cos(2\pi n \log_2 q) \\
&= \lim_{H\to\infty} 2^{-H} \sum_{q\in\mathbb{Q}|H(q)=1}^{H} \sum_{p\in\mathbb{Q}} C(p)^{-\sigma} \cos\left(2\pi n \log_2(pq)\right)
\end{aligned}$$

Eventually, we can rewrite the periodic scale structure:

$$\begin{aligned}
\text{𝄡}(\phi, b, \sigma) &= \frac{1}{\log_2 b} \sum_{m=1}^{\infty} \text{𝄢}\left(\frac{m}{\log_2 b}, \sigma\right) \cos(2\pi m \phi) \\
&= \log_b 2 \lim_{H\to\infty} 2^{-H} \sum_{q\in\mathbb{Q}|H(q)=1}^{H} \sum_{p\in\mathbb{Q}} C(p)^{-\sigma} \sum_{m=1}^{\infty} \cos\left(2\pi m \log_b(b^\phi pq)\right)
\end{aligned}$$

Although we can evaluate a profile for this function introducing a cut-off $m < M$, this formal expression may be divergent.

6. Further links with number theory

Consider the dissonance ”Hamiltonian” $D(q = \frac{a}{b}) = \log C(\frac{a}{b}) = \log(ab)$ where a and b are relatively prime integers, and the partition function (characteristic function):

$$Z(\sigma, t) = \sum_{q\in\mathbb{Q}^+} e^{-\sigma D(q)} \cos(t \log q) = \sum_{\frac{a}{b}\in\mathbb{Q}^+} (ab)^{\sigma} \left(\frac{a}{b}\right)^{it}$$

The above series is not absolutely convergent, but we observe:

$$\begin{aligned}
|\zeta(\sigma+it)|^2 &= \Big(\sum_{m=1}^{\infty} m^{-\sigma-it}\Big)\Big(\sum_{n=1}^{\infty} n^{-\sigma+it}\Big)\\
&= \sum_{m,n=1}^{\infty} (mn)^{-\sigma}\Big(\frac{n}{m}\Big)^{it}\\
&= \sum_{m,n=1}^{\infty} (mn)^{-\sigma}\cos\Big(t\log\frac{n}{m}\Big)\\
&= \sum_{q\in\mathbb{Q}^+}\sum_{l=1}^{\infty}\big(C(q)l^2\big)^{-\sigma}\cos(t\log q)\\
&= \zeta(2\sigma)\sum_{q\in\mathbb{Q}} C(q)^{-\sigma}\cos(t\log q)
\end{aligned}$$

Therefore we have the absolute convergence for $\sigma > 1$:

$$Z(\sigma>1,t) = \frac{|\zeta(\sigma+it)|^2}{\zeta(2\sigma)}$$

From numerical simulations, the ambiguity in the order of summation is obvious for $\sigma < 1$.

Let us consider two objects related to this series. The first one uses this order of summation (with $\sigma > \frac{1}{2}$ for sounds of finite energy):

$$\mathbb{Q}_N = \{q = \frac{a}{b} \in \mathbb{Q}^+ | a, b \leq N\}$$

$$\sum_{q\in\mathbb{Q}_N} e^{-\sigma D(q)}\cos(t\log q) \xrightarrow[N\to\infty]{} \frac{|\zeta(\sigma+it)|^2}{\zeta(2\sigma)}$$

and I wonder if this limit above is correct (maybe for all σ)? This is equivalent to considering:

$$\sum_{n=1}^{N} n^{-\sigma+it}$$

The second object is the generating function of the probability distributions defined as:

$$\mathbb{Q}_H = \{q \in \mathbb{Q}^+ | H(q) \leq H\}$$

$$\lim_{H\to\infty}\frac{\sum_{q\in\mathbb{Q}_H} e^{-\sigma D(q)}\cos(t\log q)}{\sum_{q\in\mathbb{Q}_H} e^{-\sigma D(q)}} = \lim_{H\to\infty}\frac{\sum_{\frac{a}{b}\in\mathbb{Q}_H}(ab)^{-\sigma}\left(\frac{a}{b}\right)^{it}}{\sum_{\frac{a}{b}\in\mathbb{Q}_H}(ab)^{-\sigma}}$$

where $H(q)$ is the sum of the term in the continued sequence of q. More conveniently, it is the length of the path to reach q in the Stern-Brocot or Calkin-Wilf binary trees. In this sense, this series is a path integral. All I know for $\sigma \leq 1$ are the following facts:

- The normalization for $\sigma = 1$ is H
- The normalization for $\sigma = 0$ is 2^H
- The probability distribution for $\sigma = 0$ has the cumulative distribution function $?\left(\frac{x}{1+x}\right)$ where ? is Minkowski's question mark function

From numerical evaluations, it seems that it is converging (at least for $\sigma = 0$), smooth and continuous, with many zero crossings. However, I have no idea how to get these properties analytically.

Note that $\mathbb{Q}_N$ refers to the Farey enumeration of rational numbers, whereas $\mathbb{Q}_H$ refers to Stern-Brocot or Calkin-Wilf enumeration.

We find another series:

$$\begin{aligned}
\frac{|\zeta(\sigma+it)|^2}{\zeta(2\sigma)} &= \sum_{\frac{a}{b}\in\mathbb{Q}_+} (ab)^{-\sigma}\left(\frac{a}{b}\right)^{it} \\
&= \sum_{q\in\mathbb{Q}^+} C(q)^{-\sigma}\cos(t\log q) \\
&= \sum_{n=1}^{\infty} \frac{A_n(t)}{n^\sigma} \\
\text{where } \; A_n(t) &= \sum_{q;C(q)=n} q^{it} \\
&= 2^{\omega(n)} \prod_{i=1}^{\omega(n)} \cos(t\log p_i^{v_i})
\end{aligned}$$

where we have used the prime number decomposition of $q = \prod_i^{\omega(n)} p_i^{v_i}$ and $n = C(q) = \prod_i^{\omega(n)} p_i^{|v_i|}$, and $\omega(q) = \omega(n)$ is their number of distinct prime factors. Note that $q \in \mathbb{Q}^+; C(q) = n$ is equivalent to the apparently more used "unitary interaction" $d \in \mathbb{N}; d|n, (d, \frac{n}{d}) = 1$ d divisor of n such as it is relatively prime to $\frac{n}{d}$. It also converges absolutly for $\sigma > 1$. Is it the right analytic continuation for $\sigma > \frac{1}{2}$? Is it equivalent to the Farey enumeration $\mathbb{Q}_N$?

The case $t = 0$ has an equivalence in the model of "primon gas" (mixing statistical physics formalism and number theory): it is the mixing of two gas, equivalently two boson gas interacting in a unitary way, or a boson and a fermion gas non interacting.

References

[1] L. Vepstas, *On the Minkowski measure* ; arXiv:0810.1265, 2008.

[2] G. Alkauskas, *The moments of Minkowski question mark function: the dyadic period function* ; arXiv:0801.0051, 2007.

[3] H. Minkowski, *Zur Geometrie der Zahlen*, ∼ 1900.

[4] Faure Frédéric, Mezzadri Malik et Ratchov Alexandre, *Analyse et jeu musical en tempérament juste adaptatif* , 2015.

[5] Paul Erlich, see article *Harmonic Entropy* on Xenharmonic, ∼ 2000.

[6] H. von Helmholtz, *Théorie physiologique de la musique : fondée sur l'étude des sensations auditives.* Masson, 1868.

[7] Gene Ward Smith, *The Riemann Zeta Function and Tuning*, see the Xenharmonic Wiki.

[8] Peter Buch, *Favoured Cardinality of Scales*

A. Asymptotic analysis of the convergent for the derivative

A.1. The case of rational numbers

Consider the rational number $q = q_n^H = \frac{a_n^H}{b_n^H}$ where H denotes its harmonicity, such that:

$$𝄞(q_n^H) = \frac{n}{2^H} = (0.s_1 s_2 ... s_{H-1})_2$$

The arbitrary parameter evoked previously is implemented as a maximum harmonicity $H_{\max}$. Then we can approximate the derivative $𝄞'(q)$ at an increasing degree of precision with a finite difference considering the surrounding numbers of harmonicity $H + m = H_{\max}$:

$$(0.s_1 s_2 ... s_{H-1}01...1)_2 = \frac{2^m n - 1}{2^{H+m}} < 𝄞(q_n^H) = \frac{n}{2^H} < \frac{2^m n + 1}{2^{H+m}} = (0.s_1 s_2 ... s_{H-1}10...0)_2$$

Denotes these numbers by $q_m^{\pm} = q_{2^m n \pm 1}^{H+m}$. Note that we suppose $H < H_{\max}$. Then, we find the following convergent (using properties of continued fractions):

$$\begin{aligned} 𝄞'_{H_{\max}}(q_n^H) &= \frac{𝄞(q_m^+) - 𝄞(q_m^-)}{q_m^+ - q_m^-} = \frac{2}{2^{H+m}} \frac{b_{2^m n+1}^{H+m} b_{2^m n-1}^{H+m}}{2m+1} \\ &= \frac{(m b_n^H + b_n^H)(m b_n^H + b_{\lfloor \frac{n}{2} \rceil}^{H-1})}{(m + \frac{1}{2}) 2^{H+m}} \quad \text{where } \left\lfloor \frac{n}{2} \right\rceil \text{ can denote } \left\lfloor \frac{n}{2} \right\rfloor \text{ or } \left\lceil \frac{n}{2} \right\rceil \\ \Rightarrow \quad 𝄞'_{H_{\max}}(q_n^H) &\sim_{m \to \infty} \frac{m}{2^m} \frac{(b_n^H)^2}{2^H} \end{aligned}$$

This proves that $𝄞'(q) = 0 \; \forall q \in \mathbb{Q}^+$.An important remark is that this is not invariant under inversion $q \leftarrow q^{-1}$. In fact, we have seen that the correct way of seeing it is:

$$q 𝄞'(q) = (𝄞 \circ \exp)'(\log q) = (𝄞 \circ \exp)'(-\log q) = q^{-1} 𝄞'(q^{-1})$$

which means that $\dfrac{d\text{𝄞}(q)}{d\log q}$ is the relevant function (although non-differentiable and discontinuous). Then, its discrete convergent is:

$$\left.\frac{\Delta_{H_{\max}}\text{𝄞}(q)}{\Delta_{H_{\max}}\log q}\right|_{q=\frac{a}{b}} \sim_{H_{\max}\to\infty} \frac{m}{2^m}\frac{ab}{2^H} = \frac{m}{2^m}\frac{C(q)}{2^{H(q)}} = C(q)\frac{H_{\max}-H(q)}{2^{H_{\max}}}$$

$$q\text{𝄞}'(q)_{H_{\max}} \sim C(q)\frac{H_{\max}}{2^{H_{\max}}} \quad \forall q\in\mathbb{Q}^+$$

This reveals how $\text{𝄞}(q)$ is related to the harmonicity $H(q)$ and most importantly to the complexity $C(q)$ function, which confirms its relation to dissonance. We can even write the equality:

$$\frac{p\text{𝄞}'(p)}{q\text{𝄞}'(q)} = \frac{C(p)}{C(q)} \quad \forall p,q\in\mathbb{Q}^+$$

A.2. The case of irrational numbers

Irrational numbers deserve the same asymptotic analysis. Consider an irrational number x. From its coordinates in the harmonic tree (or equivalently for $\text{𝄞}(x)$ in a binary form), we can extract a sequence of converging rational intervals containing x: the two bounds are the rational number corresponding to the sequence of bits coordinates truncated after a complete block of bits of the same sign and just one bit before completing the block. For instance:

$$0.1 \le 0.1100 \le 0.1100 \le 0.1100100 \le 0.110010011 \le 0.110010011100... = \text{𝄞}(x)$$

$$0.11 \ge 0.110 \ge 0.11001 \ge 0.110010 \ge 0.1100100111 \ge 0.110010011100... = \text{𝄞}(x)$$

Let us denotes these rational numbers, which are consecutive at a certain level of harmonicity H (strictly increasing in the sequence): $q_n^H = \frac{a_n^H}{b_n^H} = \text{𝄞}^{-1}(\frac{n}{2^H})$ and q_{n+1}^H. Their consecutiveness in the tree translates as the following property (inherited from the underlying symmetry of the modular group): $a_{n+1}^H b_n^H - a_n^H b_{n+1}^H = 1$

From this, we deduce a "convergent" for the derivative (which can be proved to diverge at all irrational numbers):

$$\text{𝄞}'_H(x) = \frac{\frac{n+1}{2^H} - \frac{n}{2^H}}{q_{n+1}^H - q_n^H} = \frac{b_{n+1}^H b_n^H}{2^H}$$

Therefore, the relevant function approximated for x known to harmonicity at least H is:

$$\left.\frac{\Delta_{H_{\max}}\text{𝄞}(q)}{\Delta_{H_{\max}}\log q}\right|_{q=x} = \sqrt{q_{n+1}^H q_n^H}\,\text{𝄞}'_H(x) = \frac{\sqrt{a_n^H b_n^H a_{n+1}^H b_{n+1}^H}}{2^H} = \frac{\sqrt{C(q_n^H)C(q_{n+1}^H)}}{2^H}$$

We can get the asymptotics in the special case of the golden ratio $\varphi = \dfrac{\sqrt{5}+1}{2}$ (or its inverse), which is an important special case since it is the most badly approximable irrational number. The sequence of consecutive rationals is $\{\frac{F_n}{F_{n-1}}, \frac{F_{n+1}}{F_n}\}$. Using Binet's fomula for Fibonacci numbers $F_n = \frac{\varphi^n - \varphi^{-n}}{\varphi - \varphi^{-1}}$, we get:

$$\left.\frac{\Delta_{H_{\max}} 𝄞(q)}{\Delta_{H_{\max}} \log q}\right|_{q=\varphi} = \frac{(F_H)^2}{2^H} \sim \frac{1}{5}\left(\frac{\varphi^2}{2}\right)^H$$

We may ask whether each irrational number can be characterised by two positive numbers $\alpha > 0$ and $\beta \in [1, \frac{\varphi^2}{2}]$ such that:

$$\frac{\Delta_{H_{\max}} 𝄞(q)}{\Delta_{H_{\max}} \log q} \sim \alpha\beta^H$$

There are probably "goodly" approximable numbers ($\beta = 1$) which are not following this exponential divergence but rather a polynomial one.