Why sines?

Author: Anonymised

Abstract
Sines (and cosines) seem to crop up fairly often in music theory. However, even in the seminal case of sound signal Fourier decomposition, their use can be questioned, as alternative bases for decomposition exist both theoretically and practically. Nonetheless, sines appear in many other musical contexts too. This paper answers the seemingly simple question whether their use is a commodity grounded on common usage or are there fundamental, systemic reasons for it.

Keywords: sine, cosine, Fourier decomposition, convolution, autocorrelation.

A good question

This paper was born during a talk in the SMCM convention in London, 2015 [27]. The orator has expounded some fascinating developments in the then nascent theory of Discrete Fourier transforms of musical structures (usually pc-sets[1] in the context). In this theory, n-periodic musical objects are decomposed into sums of sines:[2]

$$\mathbf{1}_A(t) = \frac{1}{n}\sum_k c_k \exp(2ik\pi t/n)$$

where the sines are another expression for the exponential: $\exp(2ik\pi t/n) = \cos(2k\pi t/n) + i\sin(2k\pi t/n)$.
At the end of the talk, a distinguished member of the audience asked the following fundamental question:

> Why do you choose the [orthonormal] basis of sines and cosines, while there are infinitely many other bases to choose from?

He was actually re-framing a question already written down in his own magnum opus [22], and in infinite dimensional context. Indeed, even if originally sines and cosines (or complex exponential functions) have been used to provide decompositions of **periodic** sounds, this theoretical case appears to be ill suited to actual musical sound, if only because the latter is not made of periodic signals – to begin with, few musical pieces have infinite duration![3] Actually, *windowed* Fourier transform, wavelets (Haar, Daubechie, Tchamichian, etc.) with finite support (or quick decay) appear better suited at practical harmonic analysis, and are effectively used for example in many file compression standards[4].

Faced with an admittedly embarrassing question, the orator found a fairly convincing answer, arguing the simplicity of the differential equations whose solutions are these functions (harmonic oscillators). This simplicity or even minimality argument has a lot going for it, adhering to William of Occam's famous razor. But was there more to it than that? At least in the field of mathematical music theory? Or was it another instance of the

[1] A *pitch-class set* is a collection of notes considered modulo octave. One can view these pitch-classes as the names of notes, and model them as integers modulo 12. This fits in the decomposition in sines because the pitch-class sets are 12-periodic.

[2] The formula is up to some constants, depending on the precise definition adopted for DFT, for instance $1/\sqrt{n}$ instead of $1/n$. Also I leave aside groupings and symmetries occurring when the signal is real-valued. Throughout we will consider interchangeably (co)sines and exponentials of imaginary quantities.

[3] Even Satie's *Tango perpétuel* or Chopin's mazurka in F minor must have a beginning in time and all interpretations (so far) have had an ending of sorts.

[4] Recently, *Hadamard transform* was used on recursively binary rhythms with finer results than Fourier transform or wavelets, see [29].

famous saying[5]: 'When your favourite tool is a hammer, all problems look like nails'? This is the topic of this paper.

We will sum up recent results of Discrete Fourier transform of musical objects, explore the importance and spontaneous emergence of Sines in many mathematical situations, as an idiosyncratic feature of the scientific mind, which are fairly well known and understood, and eventually focus on musically specific processes.

1 DFT of rhythms or pc-sets

See [2] for complete references or [6] for an introduction.

Replacing a set (say a pc-set) by its characteristic function provides the advantage of working in a smooth linear space, often called *space of distributions.*

For instance the C major triad, or pc-set $\{0, 4, 7\}$, can be seen as the vector $(1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0) \in \mathbb{R}^{12}$. Natural operations on these distributions are perhaps more important for computer implementations than for actual musical practice:[6] the intersection of two pc-sets can be done by termwise multiplication of their distributions, but their union is only the sum of distributions when the intersection is void. Values other than 0 or 1 can be used to model loudness or weight or redoublings of a pc, though the meaning of a complex value is anyone's guess (see however section 4.4). Hereafter we present objects with a fixed period of n, which can be, for instance, periodic rhythms with a period of $n = 16$ eighth notes, or pitch-class sets reduced to an octave containing $n = 12$ notes.

The **Fourier transform** of a distribution $s = (s_0, s_1, \dots s_{n-1}) \in \mathbb{C}^n$ is defined[7] by another, complex-valued, distribution $\widehat{s}$:

$$\forall k \in \mathbb{Z}_n \quad \widehat{s}_k = \sum_{\ell \in \mathbb{Z}_n} s_\ell \, e^{-2ik\ell\pi/n}.$$

All indexes can be taken modulo n, as elements of the cyclic group $\mathbb{Z}_n$. All sums run on the whole of the cyclic group $\mathbb{Z}_n$ unless specified. When $s = \mathbf{1}_A$ is the characteristic function of (say) a pc-set A, this formula reduces to

$$\forall k \in \mathbb{Z}_n \quad \widehat{\mathbf{1}}_A(k) = \widehat{a}_k = \sum_{a \in A} e^{-2ika\pi/n} = \sum_{a \in A} \cos(2ka\pi/n) + i \sin(2ka\pi/n).$$

For instance when A is the C major triad as given above, one gets

$$\widehat{\mathbf{1}}_A(2) = \widehat{a}_2 = e^0 + e^{-2i2\times 4\pi/12} + e^{-2i2\times 7\pi/12} = 1 + e^{-4\pi/3} + e^{-7\pi/3} = 1.$$

One can express the initial signal as a sum of complex exponentials (which are combinations of sines and cosines, and the mathematician's choice) by *inverse Fourier transform*:

$$s_\ell = \frac{1}{n} \sum_k \widehat{s}_k e^{+2ik\ell\pi/n} = \frac{1}{n} \sum_k \widehat{s}_k \big(\cos \frac{2k\ell\pi}{n} + i \sin \frac{2k\ell\pi}{n}\big).$$

Is this a gratuitous choice of basis for expressing the distribution s? This is the question that we are studying.

A first, though partial answer, is that we enjoy many fascinating properties and interpretations of the values of the $\widehat{s}_k$, the Fourier coefficients.

Of course, since $t \mapsto e^{2ikt\pi/n}$ has period n/k, one can see at a glance on the Fourier coefficients whether the signal has periods smaller than n (for instance, if all odd index coefficients are nil, then $n/2$ is a period), which is more or less the first motivation for Fourier decompositions in general. It is perhaps counterintuitive that *discrete* periodic sequences are still sums of (discrete) sines, see for instance Fig. 1.

[5] Known as *Maslow's Hammer*, see `https://en.wikipedia.org/wiki/Law_of_the_instrument`.

[6] See however [8] and section 4.2 on convolution. Also intersection of sequences was often used by Xenakis,which he called *sieving*, while union of disjoint translates of a set is involved in tilings.

[7] Other possible definitions differ by a multiplicative constant. See footnote 2.

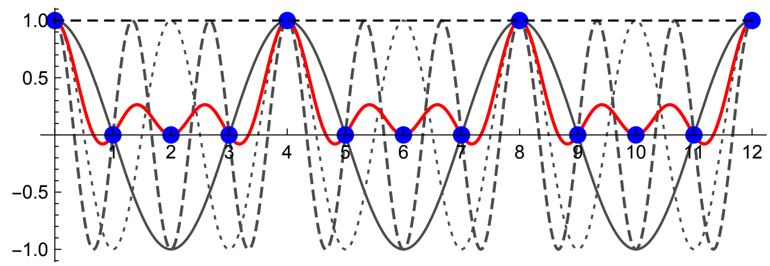

Figure 1: Characteristic function of the diminished seventh $\{0, 3, 6, 9\}$ as $\dfrac{1 + 2\cos(2\pi t/3)}{3}$ (solid, red). The signal is the sum of its harmonics (dotted lines) but only the values on integers are considered (blue dots).

Less trivially, for the characteristic function of a pc-set A, this pc-set is *Maximally Even*[8] if and only if some precise Fourier coefficient has maximal magnitude.[9] For instance, any diatonic scale, i.e. any translate of pc-set $\{0, 2, 4, 5, 7, 9, 11\}$ has the magnitude of its fifth (or seventh) coefficient equal to $2 + \sqrt{3} \approx 3.73205$, which is the maximal value for this coefficient among all seven-note scales. This can be understood as expressing that this scale is closest to a regular heptagon in the chromatic circle, as shown on Fig. 2. Indeed, in the simpler case of a 4-notes scale, the largest $4th$ coefficient is found for a diminished seventh, which is drawn on the chromatic circle as a perfect square; for 8 notes the winner is the octatonic scale, the complement of the latter. Similarly, rhythms like the tresillo `XooXooXo` or cinquillo `XXoXXoXo` are Maximally Even (and hence *Euclidean*, and *Deep* in the sense of Toussaint in [16]), being closest to regular 3- or 5-gons. The opposite of *deep* is *flat*, such as $\{0, 1, 4, 6\} \subset \mathbb{Z}_{12}$ of $\{1, 2, 4\} \subset \mathbb{Z}_7$, and this can also be seen directly on the DFT ([2], sec. **4.3**).

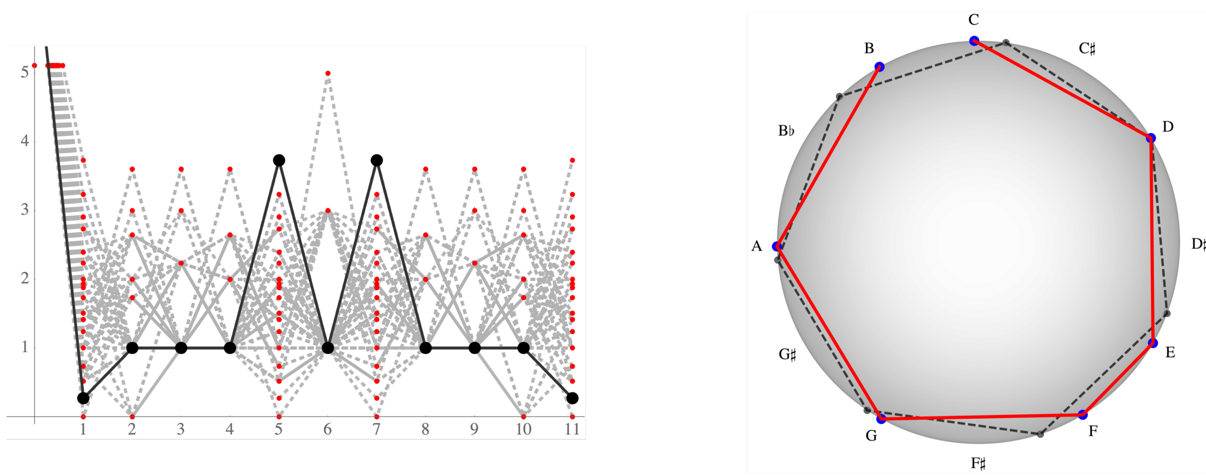

Figure 2: Left: Diatonic scales (black) have maximal fifth coefficient among all 7-note scales (gray, dotted), and are closest to regular heptagons (Right).

Hence the size/magnitude of a specific (complex) Fourier coefficient measures a definite musical character, see Fig. 3: for pc-sets, diatonicity, chromaticism, even octatonicity[10] (a contentious notion for slavic composers, see [7]), etc. For rhythms, *cinquillos* and *tresillos* exhibit a maximal ternary character which dancers appreciate in Caribbean and South-American musics, like Tango.

Also this size is invariant by transposition or inversion of a pc-set, which means that looking at Fourier coefficients magnitudes *factors out* the group action of the dihedral group T/I, leaving only information about shape, which

[8]Informally, Maximally Even Sets or ME-sets ([13]) are the most even distributions of k elements on a circle equally divided in n, like the chromatic circle. For instance, as proved by Douthett and Kranz, 7 electrons placed on 12 sites on the circle will find a stable configuration if distributed as a diatonic scale.

[9]As compared with the same Fourier coefficient for pc-sets with the same cardinality. This was discovered by Ian Quinn [25] in 2005, see also [3].

[10]Respectively considering $|\widehat{a}_5|, |\widehat{a}_1|, |\widehat{a}_4|$.

is extremely convenient (especially when one does not have perfect pitch) and musically significant[11].

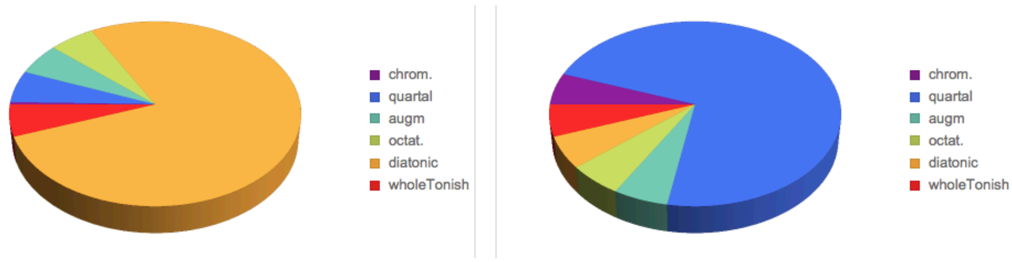

Figure 3: Pie-chart diagrams of Fourier coefficients magnitudes measure musical characters. Left: average characters in Mozart's *Sonata facile*, extremely diatonic; right: Berg's Sonata, rich in tritonic harmonies.

The other dimension of these complex coefficients, the phase, also conveniently embodies musical ideas and is actively studied in recent research, see [28] among many others. For instance, taking the phases of a_3, a_5 as coordinates yields a picture containing most pc-sets, including both Euler's Tonnetz and its dual, with their usual shape and topology and much more, see Fig. 4.

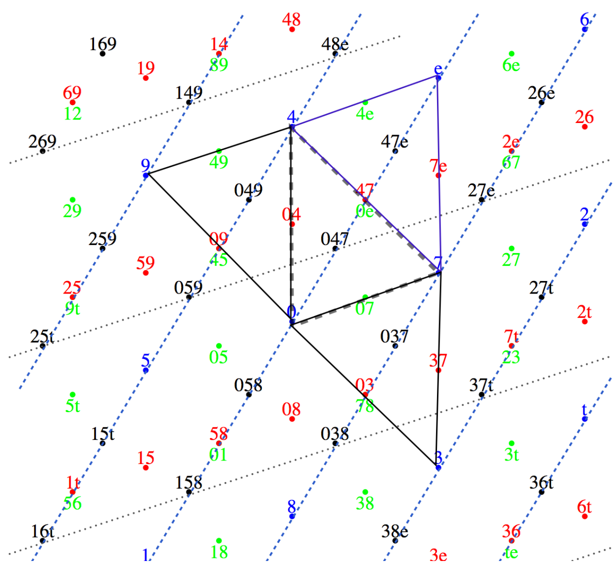

Figure 4: This *Tonnetz* of phases embeds both pitch-classes as single figures grouped as consonant triads, most dyads, the triads themselves as 3-digits labels, and most pc-sets (omitted for legibility). It is periodical both horizontally and vertically.

On the whole, many interesting interpretations arise when reading musical objects in the sine basis of Fourier decomposition. This raises the question whether there exists some *other* basis even more appropriate for musical purposes. Let us explore for the time being the first appearance of Fourier decomposition (as infinite series) in solving sound equations.

2 Between music and their maths: vibrations

2.1 Harmonic oscillator

Newton's second law implies that any physical phenomenon wherein the forces exerted depend on the system's state can be described by second order differential equations, since the sum of forces is proportional to the

[11] As will be argued below, these magnitudes encapsulate no more and no less than the intervallic information of a pc-set, its IFunc.

acceleration. A very well known case is the harmonic oscillator (we will leave aside dampening factors for concision):

$$y'' = -\omega^2 y.$$

Here y measures some physical quantity varying in time (or sometimes space), and y'' is its second derivative whereas ω is some real constant depending on the physical data (for instance, the length of a pendulum). It can be illuminating to consider the solutions to the equivalent equation $y = -y''/\omega^2$ as fixed points of the differential operator $y \mapsto -y''/\omega^2$, or more loosely as a comparison between a signal and its second derivative, a perfect correlation (see section 4.1) up to a constant.

As some readers will probably remember, the space of solutions of this equation is precisely the vector plane generated by sines and cosines with pulsation ω:

$$y(t) = a\cos(\omega t) + b\sin(\omega t).$$

The simplicity of this equation, leading unequivocaly to sine and cosine solutions, is the Occam's razor argument provided as answer in the MCM 2015 conference, as related in the preamble. The use of this principle is prevalent in the elaboration and discussion of scientific theories because simple models are easier to test (falsifiability in the sense of Karl Popper).

But a devil's advocate would argue that the differential equation itself is a very simplified model of real situations which are far more complex (to begin with, producing musical sounds involves movements of more than a single point as involved in the pure harmonic oscillator). Could it be that sine solutions are simply products of this oversimplification? And that music, that most complex contrivance of the human mind, deserves more sophisticated explanations?

2.2 D'Alembert's vibrating string equation

It has been known allegedly since the Pythagorean school that reeds or strings (as in a monochord) produce harmonic spectra, i.e. superpositions of sounds whose frequencies are multiples of the fundamental one. This was modellised and proved with the following equation, which is derived in simplistic situations (infinitely thin strings with 'small' oscillations, for instance):

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = 0.$$

Here x is a position and t measures time, v is a constant comparable to a velocity; $u(x,t)$ measures the swelling of the string at abscissa x and instant t, see Fig. 5 which should be understood as changing with time.

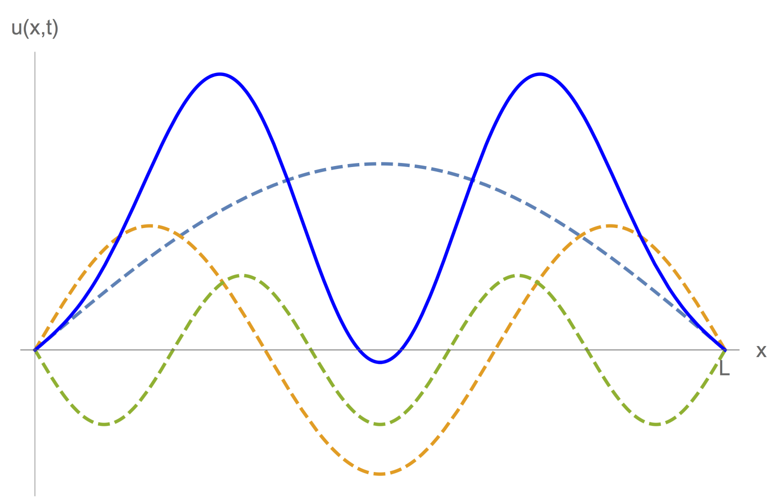

Figure 5: A musical string's shape (strong line) follows a sum of sine vibrations (dotted).

This equation was stated by Jean le Rond d'Alembert in 1729.[12] It is a *partial differential equation* (PDE), meaning it involves derivatives with respect to both time t and space x. Even *stating* such equations had only recently been made possible by the invention of differential calculus (Leibniz, Newton). As became clearer when the study of these equations slowly developed, much depends on initial and/or boundary conditions. They impose, more or less, the method outlined below. Here, considering a metallic thin string attached at both ends [there is no essential difference for an air column like in the case of a wind instrument], the value of u there is invariable in time, i.e. (for some adequate choice of constants)

$$\forall t \quad u(0,t) = u(L,t) = 0.$$

Even in this basic situation, a rigorous derivation of the most general solution was unattainable at the time.[13] A customary method for such (linear) PDE involves the *superposition principle*: a sum (or linear combination) of solutions is still a solution. As a first step, one searches for solutions with *separated variables*, i.e. the (very) particular case when

$$u(x,t) = W(x) \times T(t).$$

Injecting this in the wave equation, one finds that W and T must satisfy harmonic oscillator equations. Taking into account physical constraints and boundary conditions, one gets eventually for some *integer* n

$$W(x) = A \sin \frac{n\pi\, x}{L},$$

and

$$T(t) = B\sin(\frac{n\pi v t}{L}) + C\cos(\frac{n\pi v t}{L}) = D\sin(\frac{n\pi v t}{L} + \varphi).$$

(A, B, C, D are arbitrary real constants)
Let us point out the most illuminating points of this resolution:

- Spatial and temporal frequencies are proportional – to quote Wagner's Gurnemanz, "Zum Raum wird hier die Zeit".
- There is discretization of the spectrum, which is *harmonic*: the n^{th} frequency is equal to n times the first one, or *fundamental*. This is easily understood for the spatial part, where the nodes (the points where the string does not move) must divide the whole length by some integer (see Fig. 5), but the corollary that the temporal spectrum is harmonic vindicates the pythagoreans' discovery two millenaries before.
- The resolution boils down to the simplest case of the harmonic oscillator, with its plain sine solutions, fixed points of the squared differential operator.
- The more general solution can be rebuild from adding up elementary solutions for all values of n:[14]

$$u(x,t) = \sum_{n\in\mathbb{N}} a_n \sin\frac{n\pi x}{L} \sin(\frac{n\pi v t}{L} + \varphi_n).$$

Of course, we could again argue that such a simplistic equation is a biased argument for the use of sine functions. Indeed, the spectrum of many instruments (say, the piano) is only approximately harmonic. We feel compelled to examine a less frugal model.

[12] Learned readers will remember, or notice, that $1729 = 10^3 + 9^3 = 12^3 + 1^3$.

[13] The more complicated Navier-Stokes PDE governing the behaviours of fluids are still very much unsolved, so much so that even a partial resolution would be rewarded by one of the seven 1-million prizes of the Cray foundation.

[14] Leaving aside problems of convergence for the infinite sum, mostly irrelevant in the physical world where it can be shown generally that all solutions can be expressed in this way.

2.3 A more complicated case

It can be argued that d'Alembert's equation is oversimplified. It is in the nature of a model to stand up insofar as it has not shown its limits, and then to be supplanted by one more refined, with perhaps brand new objects as solutions.

Not so in this case. A very interesting illustration can be made with the analysis of [open-ended] tubular bells (also called chimes) published recently in [24] (among other recent papers on the topic, with similar results). Taking into account that a tubular bell has non-zero width, the authors eventually derive a more complicated, fourth order PDE:

$$\frac{\partial^4 u}{\partial x^4} + \frac{\partial^2 u}{\partial t^2} = 0$$

where x, t are reduced variables (obtained by adequate choices of length and duration units). We concentrate on the nature of the solutions, not on the details which are well covered in the paper. In substance, it appears that the spectrum is not harmonic[15] (or only approximately), but the expression of the solutions still unequivocally involves sines.

Computations similar in the method but more arduous lead to elementary solutions $W(x)T(t)$ where the spatial component $W(x)$ is a combination of trigonometric and hyperbolic (co)sines, and $T(t) = \sin(\xi^2 t)$ where ξ satisfies the *discretizing condition*

$$\cos\xi \cosh\xi = 1.$$

Solutions to this equation are still discrete, $\xi_0 = 0 < \xi_1 = 4.73004 < \xi_2 = 7.8532 < \dots$, but no more harmonic: $\xi_n - \xi_{n-1}$ is not a constant (though this quantity quickly converges[16] when n grows to infinity, but high-order, inaudible harmonics are physically irrelevant). The temporal frequencies are no more proportional to the ξ_n, but instead to the ξ_n^2.[17]

This study indicates that even in complicated cases, when all else fails, sines endure: the superposition method will always yield sines, because these functions are the solutions of linear ODEs with constant coefficients, which generalise the harmonic oscillator to any conceivable order but whose solutions still are [complex] eigenvalues of differential operators, hence sines.[18] The MCM Occam argument actually applies to all solutions of all linear problems (roughly speaking, any oscillating system) in any dimension.[19]

It can be argued that *discrete* musical structures should not be treated as continuous objects – indeed the lines on Fig. 1 between the relevant blue dots are more confusing that helpful! However, even before the successes related in Section 1, counterparts to the continuous sines that we found in (partial) differential equations have shown their importance in discrete mathematics in general.

3 Linear equations and geometric sequences

3.1 Linear recursions

One distinctive feature of Discrete Fourier transform, and even of Fourier series decomposition, is that the Hilbert basis of maps involved, the $t \mapsto e^{nit}$ for $n \in \mathbb{Z}$ (or the $t \mapsto e^{2ik\pi t/n}$ for $k \in \mathbb{Z}_n$ in the finite case), are *powers of one of them:*

$$e^{nit} = \left(e^{it}\right)^n \quad \forall n \in \mathbb{Z}.$$

This makes sense because an integer power of a complex number is well defined. The sines are barely hidden in this formula through Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$.

[15] This is very well validated by measurements, and vindicates the common aural experience of a bell's spectrum, discrete but not quite harmonic. More or less similar results hold for Tibetan bowls.

[16] Actually $\xi_n \approx (n + \frac{1}{2})\pi$. It follows easily from the discretising condition.

[17] Significantly, ξ_5^2/ξ_4^2 is very close to a just fifth, which is quite perceptible in the actual sound of a tubular bell and helps in tuning them together.

[18] The only *bounded* solutions, that is to say the only ones physically possible in the temporal domain.

[19] I leave aside here non-linear equations (or linear with non constant coefficients) such as those modelling drums, which involve other useful families of solutions, like Bessel functions: this is relevant in signal processing, but so far almost never in music theory.

This echoes through centuries the generation of secondary modes of vibration for a string.[20]

Powers sometimes occur in other decompositions in Hilbert bases, but in less straightforward ways. For instance, many wavelets are derived from a 'mother' wave by *scaling:*

$$\varphi_{n,\tau}(t) = \varphi\Big(\frac{t-\tau}{2^n}\Big).$$

and the daughter wavelet is not a power of its mother.

Moreover, complex geometric sequences play a vital part in solving linear recursions. Of course, $(\lambda^n)_{n\in\mathbb{N}}$ is the basic solution of the simplest possible recursion equation, that is to say $u_{n+1} = \lambda u_n$. But even when a, b are real constants, the solutions of

$$u_{n+2} = a u_{n+1} + b u_n$$

are almost always linear combinations of $(\lambda^n)_{n\in\mathbb{N}}$ and $(\mu^n)_{n\in\mathbb{N}}$ wherein λ, μ may very well be complex, non real.[21] This holds, *mutatis mutandis*, for higher-order recursions. Where are the sines? Say $a = 2, b = -2$, then $\lambda = 1 + i = \sqrt{2}e^{i\pi/4}$ and $\mu = \sqrt{2}e^{-i\pi/4}$ so that powers of both are $2^{n/2} \times e^{\pm ni\pi/4}$. It could be said that scaling, or normalization (cancelling by the amplitude $2^{n/2}$) yields again a Fourier basis. In some cases the normalization is superfluous: the solutions of $u_{n+2} = u_{n+1} - u_n$ are $u_n = \alpha\cos(n\pi/3) + \beta\sin(n\pi/3)$. This holds in other spaces, too, as we will see in the next, slightly digressive, section.

3.2 Anatol Vieru, Neptune, difference equations, and powers of eigenvalues

In most situations, the iteration of a transformation *asymptotically* reduces to a geometric progression. For instance, the Lukas sequence (ℓ_n) (cousin to the famous Fibonacci sequence):

$$2, 1, 3 = 2 + 1, 4 = 1 + 3, 7, 11, 18, 29, 47\ldots$$

has for nth element the *integer* closest[22] to $\Big(\frac{1+\sqrt{5}}{2}\Big)^n$.

Generally speaking, this points out the largest eigenvalue of the linear transformation generating the sequence, that is to say

$$\begin{pmatrix}\ell_{n-1}\\ \ell_n\end{pmatrix} \mapsto \begin{pmatrix}\ell_n\\ \ell_{n+1}\end{pmatrix} = \begin{pmatrix}\ell_n\\ \ell_{n-1}+\ell_n\end{pmatrix}$$

whose matrix is $\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}$ and eigenvalues $\dfrac{1\pm\sqrt{5}}{2}$. The larger one, the Golden ratio, appears as the limit of the (rational) quotient ℓ_n/ℓ_{n-1}, as Leonardo da Pisa found out around 1200.

In the same vein, in the 1840's, astronomer Urbain Leverrier determined to three decimal places the largest eigenvalue of a 4×4 matrix by studying the asymptotic behavior of its powers, a crucial step in his computation of the orbit of Neptune. His ground-breaking idea (years before the notion of diagonalization was even considered) was that for a given matrix A, and almost all vectors X, one has

$$A^n X \approx \lambda^n.X$$

where λ is the largest eigenvalue of A.[23]

Returning to music, in discrete situations the approximation by a geometric sequence is often exact after a few iterations, because the vanishing part actually disappears:

- Take the Lukas sequence above, modulo 12: it is now periodic with period 24 and can be expressed with powers of "numbers" lying in a quadratic extension of the ambient ring, namely $\mathbb{Z}_{12}[X]/(X^2 - X - 1)$, and whose order is 24, like $e^{\pm i\pi/12}$ in the complex field.

[20] Or less romantically the fact that the algebra of characters on periodical functions is monogenous.

[21] Whenever $a^2 + 4b < 0$. When this quantity is nil, the solutions are more complicated.

[22] $\Big(\frac{1+\sqrt{5}}{2}\Big)^7 = 29.0344\cdots$ while $\ell_7 = 29$.

[23] The equation is exact when X is an eigenvector. When several eigenvalues share the same magnitude, we get a more complicated combination of their powers.

- Romanian composer Anatol Vieru iterated the difference operator on cyclic lists of pc-sets and observed that this procedure produces a periodic sequence after a few steps. For instance, numbering pitch-classes from 0 to 11 and starting from (0, 0, 7, 7, 4, 4, 3, 3, 4, 4, 7, 7) one gets

$$(0-0=0, 0-7=5, 7-7=0, \dots) = (0,5,0,3,0,1,0,11,0,8,0,7) \mod 12,$$

  and this process soon reaches $(4,4,0,8,8,0,4,4,0,8,8,0)$ which the difference operator permutes circularly. See [9] for a detailed analysis. I mention this sequence beacause again, it could be expressed with powers of elements of a quotient ring, some of them nilpotent (which may or may not be a good idea).

- Composer Tom Johnson has explored *autosimilar melodies*: when extracting from some periodic sequence one note out of (say) 3, sometimes one finds the initial melody.[24] In [5] the question is thoroughly explored mathematically, eventually studying the effect of such an extraction on a random (not "autosimilar") sequence. The result is essentially equivalent to Vieru's: the sequence converges to something autosimilar [here a link to the sound file Attracteur autosimilaire.m4a].

- The most recent case known to the author is the iteration of a *discrete average operator* [17], with similar results albeit the transformation is not exactly linear.

These examples (approximately for the last one) mostly exemplify the general behaviour of all linear operators, governed by the all-embracing:

Lemma 1 (Fitting Lemma) *For any linear transformation of a finite dimensional linear space, or finite type module, the space is a direct sum of two stable subspaces, on which the linear map is respectively nilpotent (it vanishes after some iterations) and bijective.*

If the space has a finite number of elements, this means that when the nilpotent part has vanished under iteration, what remains is periodical.[25]

The general idea underlying all these disparate cases is that **from a distance, iterating any linear transformation looks a lot like multiplying some fixed vector by a geometric sequence** λ^n, which generalises the Fourier basis case.

Admittedly, all of this is exceedingly general, mathematical and abstract. It proves nonetheless how natural it is for the human brain to use sine functions when solving Nature's problems – at least when assuming its equations are linear.[26] Again we may ask: what is so specific to music in that respect?

4 Musical operations: IFunc and convolution product

We shall now return to music theory and its abstract, cultural, and even perceptual, concepts. Remember how Vieru's process compares elements of a sequence to the previous ones; Johnson's to subsequences; Lukas or Fibonacci numbers to the two previous ones. It is an aesthetic longing in the human mind to *compare* the object it is considering to itself.

4.1 Some music at last: correlation

Periodicity, in some loose sense, is perhaps the most essential feature of music: most listeners will expect repetition of the same or partial recognition of the original theme, or of some feature at least – think of Fuga construction, recaps and motivic use in general, or local stability of a mode or tonality, or even stability of the tonal character through modulations. This involves essential cognitive capacities which are very useful in other situations, so

[24] Try (F **F** F D **F** F F **D** F F **F** D . . .) for a definitely Beethovenian case. There are even earlier examples.

[25] It also works in the messier setting of a module on a non-division ring, such as $\mathbb{Z}_{12}$, with nilpotent "numbers". This would be the case for Vieru's transformation but it has not been used so far in the literature. On the other hand, Fitting's Lemma also holds in a number of infinite and infinite dimensional cases.

[26] As is customarily done for the pendulum equation, quite improperly!

much so that recent research encourages therapeutic use of music as intellectual stimulation, from kindergarten kids to Alzheimer patients [11].
The degree of repetition after a delay τ of some musical object A can be measured by the intersection of A and $A+\tau$. This applies in the temporal domain as well as pitch, or both at the same time, for instance when the subject of a fuga is repeated later on and one fifth higher.
Such comparisons have been much discussed in the second half of the XXth century, with interval vectors, contents, and functions as described by Forte [14] or Lewin [19, 21] for instance.
For simplicity we focus first on the case of the internal IFunc, which is obtained by tallying the occurrences of all possible intervals inside a pc-set $A \subset \mathbb{Z}_n$[27]:

$$\mathrm{IFunc}_A(k) = \#\{(a,b) \in A \times A,\ b-a=k\}.$$

That it measures some kind of internal periodicity is better seen from the equivalent definition:

$$\mathrm{IFunc}_A(k) = \#(A \cap (A+k)).$$

For instance, an octatonic scale satisfies $\mathrm{IFunc}_{Oct}(3) = 8$, meaning that it is the same when transposed by a minor third; and a modulation to the dominant turns C major, or $\{0,2,4,\mathbf{5},7,9,11\}$, into G major $= \{0,2,4,\mathbf{6},7,9,11\}$, with 6 notes in common. Hence $\mathrm{IFunc}_{CM}(7) = 6$, which means that the scale is actually a sequence of fifths and also that a fifth transposition will change only one pitch class (F to F♯). Conversely, the notorious pc-set $\{0,1,4,6\}$ shares exactly one pitch-class with all of its transpositions (2 with the tritone transposition, 4 with itself). This notion has grown in importance for both analysis and composition, but seems to take us far away from the question of sines. It does not, because IFunc is a type of convolution.

4.2 Some maths: convolution of distributions

The IFunc function (whether between two pc-sets or the same one) is actually an instance of a *convolution product*. It is averred that David Lewin was aware of this already in his very first published paper [20]. The general definition of convolution between two vectors, or lists, or distributions, is

$$(s*t)_k = \sum_{x \in \mathbb{Z}_n} s_x t_{k-x}$$

where indexes are again considered modulo n for commodity.
When s, t are characteristic functions of pc-sets (or periodic rhythms), meaning that $s_k = \mathbf{1}_A(k) = \begin{cases} 1 & \text{when } k \in A \\ 0 & \text{else} \end{cases}$ and $t = \mathbf{1}_B$, one easily checks that

$$(s*t)_k = (\mathbf{1}_A * \mathbf{1}_B)(k) = \sum_{x \in \mathbb{Z}_n} \mathbf{1}_A(x) \times \mathbf{1}_B(k-x)$$

has 1's instead of 0's in the sum whenever $x \in A$ and $k - x \in B$, i.e. $x \in k - B$ (the retrogradation of B transposed by k), which happens when k is an interval between an element of $-B$ and an element of A: *this quantity tallies the elements of* $-B + k \cap A$. In other words, changing B to $-B$ we retrieve[28]

$$\mathrm{IFunc}_{A,B} = \mathbf{1}_A * \mathbf{1}_{-B}.$$

Notice that the complexity of this computation is *quadratic*: one has to effect n times n comparisons.
$*$ is of course fundamental in signal processing and its properties are well known.[29] As seen in the last paragraph under the guise of correlation, it has psychoacoustical value for comparison purposes, checking how well a copy of A translated in space corresponds with B, enabling to notice repetitions, recaps, periodicities and subtler phenomenons. Many, if not most, musical operations in the brain can be expressed mathematically as convolutions.
The most important link for our present topic is the following simplification theorem.

[27] Or up to a constant, the probability of occurrence of an interval when notes of A are played uniformly randomly.
[28] $\mathrm{IFunc}(A,B)$ tallies intervals from A to B.
[29] It is a 'good' multiplication: Associative, commutative, bilinear, with a neutral element $\delta = (1, 0, \dots 0)$.

Theorem 1 *(Discrete) Fourier transform turns convolution product into termwise product, i.e. for any two distributions s, t, in any index k, one has*

$$\widehat{s * t}_k = \widehat{s}_k \times \widehat{t}_k.$$

The proof, elementary,[30] can be found in any textbook on the subject. Notice that the left-hand term involves convolution whose computational complexity is n^2, whereas the termwise product of $\widehat{s}$ and $\widehat{t}$ is done in *linear time* (n multiplications), which cannot be bested for n-dimensional objects. This quality is so important that in many forms of convolutions, like multiplication of polynomials or of very large integers, are processed by first applying Fourier transform (or rather *Fast* Fourier Transform, which has $n \log n$ complexity), then multiplying termwise, then doing inverse Fourier Transform.
From Thm. 1 we get immediately that the Fourier transform of the IFunc is the squared magnitude of the Fourier transform of the set A:

$$\widehat{\text{IFunc}(A)} = \widehat{\mathbf{1}_A * \mathbf{1}_{-A}} = \widehat{\mathbf{1}_A} \times \widehat{\mathbf{1}_{-A}} = \widehat{\mathbf{1}_A}\,\overline{\widehat{\mathbf{1}_A}} = \left|\widehat{\mathbf{1}_A}\right|^2.$$

This is capital, showing that the magnitude of the Fourier coefficients is equivalent to the intervallic information.[31]
Another important formal information is that DFT itself – a termwise product of a distribution with a collection of exponentials – can be interpreted as correlation with a geometrical sequence of exponentials. This illuminates the characterization of ME sets by Quinn, which told us that (for instance) 4-notes pc-sets have maximal $4th$ coefficient when correlated with an arithmetic sequence with ratio 3, i.e. a diminished seventh. On these tricky interpretations, see [7].

We can now turn back to (co)sines.

4.3 A last, improbable cosine

Bill Sethares once communicated to the author the following *empirical observation* (!):

> Iterating and normalising autocorrelation on **any** signal converges (quickly) to a cosine[32].

This looks mysterious or even unbelievable at first glance, but experimentation indeed shows a sine wave with maximum value in 0, i.e. a cosine[33], after 5 or 6 iterations as can be seen on Figure 6 [online an animated gif can be provided].

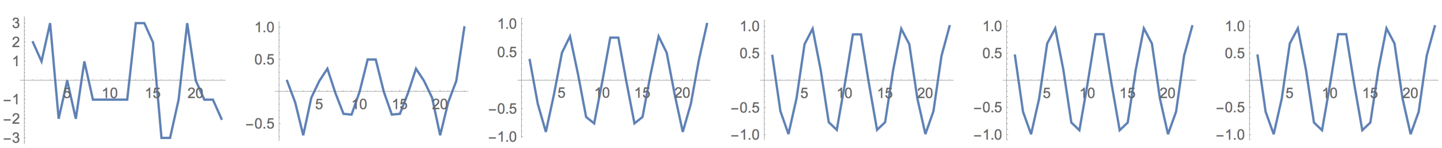

Figure 6: Iterating autocorrelation and normalization always produces a cosine.

Indeed this is easily explained and even proved (except on a nil measure set) by Thm. 1 and its corollary: we get $|\widehat{s}|^{2^n}$ for the Fourier transform of the autocorrelation of a distribution s taken n times. Assuming one of the Fourier coefficients is larger than the others, its $2^n th$ power will dominate all others. Dividing by this largest value (normalizing), there remains a 1 at its position and everywhere else, vanishing values[34] as is apparent on Fig. 7. By inverse Fourier transform one retrieves a cosine, and the quick convergence is explained by the exponential power 2^n.[35] This portends a stupendous interpretation, that the wired-in comparison (correlation)

[30] It is more difficult for the non-discrete case, but the problem is topological and not inherent to Fourier transform itself.

[31] This value is therefore the core notion for the study of *homometry*, or Z-relation as the American school once called it. See [1, 14] for historical and modern references, like the definition of the equivalent Patterson function.

[32] After inquiry it seems that this had been already noticed around the 50's by some engineers (in the general context of continuous signals), which is why Sethares did not publish about this. It may have been known earlier still.

[33] A fixed point of the operation, of course.

[34] Actually, there is another 1 at the symmetrical position because the distribution is real-valued. This is but a technical detail.

[35] Meaning that the number of 0's after the decimal dot, in the small components, doubles at each iteration.

operation in the brain spontaneously produces cosine waves! It is perhaps not a coincidence that the brain is hard-wired to perform Fourier transforms, as it does for pitch recognition.

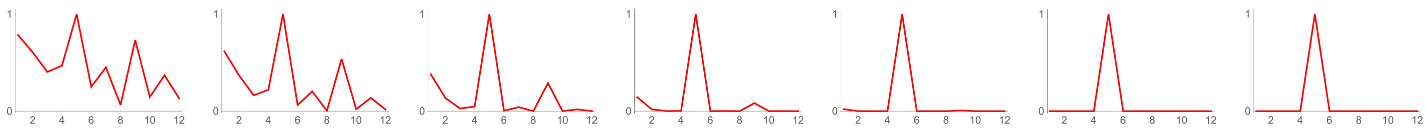

Figure 7: Fourier coefficients magnitudes for the above signals. All but the largest vanish in the iteration.

A more theoretical occurrence of sines and cosines is rooted both in the most abstract, theoretical physics, and the most practical experience (free jazz improvisation).

4.4 Mazzola's imaginary time vs. predictive power

Guerino Mazzola [23] has advocated the introduction of a second dimension to time, an imaginary axis purported to help parametrise the memorial dimension, "for a precise description of musical artistic presence", and notably the ongoing conscience during improvisation of the themes that have been developed by other jazz players.
Although his initial motivation had to do with the 'natural'(!) extension of the real world to complex vector spaces as is done in quantum mechanics, this is clearly a form of real-time correlation, doubling the complexity of perception/production of the musical flux. A second axis introduces a change of phase, enabling correlation across time, the decoupling of sine and cosine; or equivalently, introducing the basis of complex exponentials used in Fourier transform.

Quite independently, this imaginary dimension has been shown to endow musical meaning when one considers some specific *products of Fourier coefficients*, wherein the signum of the imaginary part allows for example to distinguish between major and minor,[36] or diatonic and pentatonic (recall these last two have the same magnitudes for Fourier coefficients) or the four different classes of pc-sets homometric[37] to $\{0, 1, 4, 6\}$, as recently explored in [26].
Another quite recent development in music theory may be a consequence of the brain's natural longing for sines: the analogy with quantum mechanics inspiring Mazzola was pursued rigorously by some researchers.

4.5 Quantum gauges

There is more than a fashionable trend in some recent efforts at bringing quantum mechanics concepts in the musical world. This can help for instance modelling the human mind's choices of interpretations in ambiguous situations as *collapsing a waveform* (e.g. when an musician plays a tritone, is it a diminished fifth or an augmented fourth? This can be determined by subsequent observation, see [15]). Another more involuted situation uses *gauge functions* for modelling the notion of tonal attraction (see [12, 10] for instance). Starting from physical quantities (energy, momentum, gravity) and their analogies with empirical/psychological musical perceptions (attraction, in tonal music, mostly) the authors postulate a Schrödinger-like equation admitting these quantities as solutions, more specifically in the form of cosine functions, with the nice consequence that they retrieve the discrete structure of pitch-classes. This may seem tautological (or even a vicious circle) since *quantum* mechanics purport to explain discrete changes between levels, and indeed the author must confess (to his shame) that he was at first unconvinced by these explorations. Also, the occurrence of sine solutions could be attributed to the simplified Schrödinger equation reducing to a harmonic oscillator. There was however some strong experimental evidence to support this line of research in Carol Krumhansl's famous work [18], which highlights cosines and Fourier transforms in its empirical results. Furthermore, this psycho-acoustical evidence is fully in phase with the theoretical aspects of discrete Fourier transform described in the first section. This promising field of exploration

[36] Considering the phase of $\hat{s}_2\hat{s}_3\hat{s}_7$.

[37] Two pc-sets are homometric if they share the same intervalic content, or equivalently when the magnitudes of their Fourier coefficients are equal.

could perhaps be subsumed under the heading of section 2.1 where the discretisation of sines is explained by limit conditions.

Although all the above argumentation is mathematically convincing, many among the given examples may well seem marginal. It is time to develop the most general reason why sines are legitimate in mathematical music theory.

4.6 The key theorem

Considering the musical (or even aesthetic, generally speaking) importance of the transformation theorem of convolution into termwise product, we will argue that the initial question, specifically in the domain of music theory, can be largely rephrased as:

> Are there any other transformations [than DFT] turning convolution into ordinary, termwise product?

This means achieving maximal reduction in computational complexity, surely a competitive (perhaps even evolutionary?) advantage for a brain feature. The surprisingly simple answer is, simply put: no.

Theorem 2 *Let $E = \mathbb{C}^n$ be the complex vector space of distributions endowed with either the convolution product $*$ or the termwise product $\times$. Let Φ be some linear isomorphism of E that turns $*$ into termwise product $\times$:*

$$\forall s, t \in E \quad \Phi(s * t) = \Phi(s) \times \Phi(t).$$

Then Φ is the DFT up to some permutations of coefficients: there exists a bijection σ of the indexes such that

$$\text{for all indexes } k,\ \Phi(s)_k = \widehat{s}_{\sigma(k)}.$$

For a detailed proof see [2, Thm. 1.11]. The gist of it is that by composing Φ with the inverse Fourier transform, one creates a “licorn”, a mathematical object with so many wondrous features that it collapses under its own constraints and must be trivial: it is an automorphism of E which preserves all operations, linearity and termwise product. Reasoning on members of the canonical basis, $(1, 0, 0 \dots), (0, 1, 0, 0 \dots)$ etc., which are together with their sums the fixed points of $s \mapsto s \times s$, is it fairly easy to prove that the image of a basis vector is another basis vector, hence they are permuted. Notice that once again the notion of fixed point is instrumental in the proof.[38]

Conclusion

As shown by the present (partial) survey, the power of Fourier transform in analysing musical phenomena is bewildering. The amount of evidence carries a lot of conviction in itself, though this does not suffice to prove logically that alternative bases of decomposition could not do better. However, DFT excels in two domains:

- Meaningfulness. The magnitudes of individual Fourier coefficients have direct, perceptible, immediate musical significance (for instance their collection carries essentially all that information that is invariant under transposition or inversion) and mirrors perceptual notions. So do their phases.
- Simplicity. Thm. 2 proves that it is optimal, for all manipulations of musical data reducible to a convolution operation: comparisons of repetitions, periodicities, similarities, predictability and other cognitive processes known to be involved in musical perception (and creation). This was our last and incontrovertible argument. Moreover, modelling or problem-solving is easier with economical tools, and sines (or complex exponential functions) are also optimal in that respect, being fixed points of differential operators and asymptotical limits of many natural processes, leading to unquestionable usefulness when solving ordinary or partial differential equations and discrete difference equations.

[38]Though this theorem is fairly elementary, I could not find any prior mention in the literature. Maybe it is because a more general phrasing, for non-discrete and integral Fourier transform, would be more difficult to spell out because of topological issues (which function space or spaces should be considered?).

The musical mind prefers Fourier transform because it optimises the computation of correlations, which is the essence of making sense out of our perceptions; while the mathematical mind interprets the world in sums of sines because of their "unreasonable efficiency", as fixed points of the most elementary operations – another kind of comparison. So perhaps these two minds are one and the same.

References

[1] Agon, C., Amiot, E., Andreatta, M., Ghisi, D., Mandereau, J., *Z-relation and Homometry in Musical Distributions*, J. Maths. Mus., **4** (5), Taylor and Francis (2011).

[2] Amiot, Emmanuel: Music in Fourier Space. Springer (2017).

[3] Amiot, Emmanuel, *David Lewin and Maximally Even Sets*, J. Maths. Mus. **1** (3), pp. 157-172, Taylor and Francis (2007).

[4] Amiot, Emmanuel: *The torii of phases*. In: Yust, J., Wild, J., Burgoyne, J.A. (eds.) Mathematics and Computation in Music, Fourth International Conference, 2013, LNCS, vol. 7937, pp. 1–18. Springer, Heidelberg (2013).

[5] Amiot, Emmanuel: *Autosimilar Melodies*, J. Maths. Mus. **2** (3) (2008).

[6] Amiot, Emmanuel: *The discrete Fourier transform of distributions*, J. Maths. Mus., **11** (2-3), pp. 76-100 (2017).

[7] Amiot, Emmanuel: *Interval Content vs. DFT*. In: Agustín-Aquino, O., Lluis-Puebla, E., Montiel, M. (eds) Mathematics and Computation in Music. MCM 2017. Lecture Notes in Computer Science, vol 10527. Springer, Cham (2017). DOI:10.1007/978-3-319-71827-9

[8] Amiot, E., Sethares, W., *An Algebra for Periodic Rhythms and Scales*, J. Maths. Mus. **5** (3), Taylor and Francis (2011).

[9] Andreatta, M., Vuza, D.T.: *On some properties of periodic sequences in Anatol Vieru's modal theory*. Tatra Mt. Math. Publ. **23**, pp. 1–15 (2001).

[10] Beim Graben, P., Mannone, M.: *Musical pitch quantization as an eigenvalue problem*, J. Maths. Mus., **14** (3), pp. 329-346 (2021). DOI 10.1080/17459737.2020.1716404

[11] Bigan, E., Tillman, B.: *La symphonie neuronale*. HumenSciences, Paris (2021).

[12] Blutner, R., beim Graben, P.: *Gauge models of musical forces*, J. Maths. Mus., **15** (1), pp. 17-36 (2021). DOI 10.1080/17459737.2020.1716404

[13] Clough, J., Douthett, J., *Maximally even sets*, Journal of Music Theory, **35**, pp. 93-173 (1991).

[14] Forte, Allen: *The Structure of Atonal Music*. Yale University Press, New Haven (1973).

[15] Fugiel, Bogusłav: *Quantum-like melody perception*; J. Maths. Mus. (2022).
DOI 10.1080/17459737.2022.2049383 DOI 10.1080/17459737.2022.2042410

[16] Gómez-Martín, F., Taslakian, P., Toussaint, G.: *Structural properties of Euclidean rhythms*, J. Maths. Mus., **3** (1), pp. 1-14 (2009). DOI: 10.1080/17459730902819566

[17] Hazama, Fumio: *Iterative method of construction for smooth rhythms*, J. Maths. Mus. (2021). DOI: 10.1080/17459737.2021.1924303

[18] Krumhansl, Carol: *Cognitive Foundation of musical pitch*, Oxford Psychology Series **17**, Oxford University Press (1990).

[19] Lewin, David : *Forte's interval vector, my interval function, and Regener's common-note function*, J. Mus. Theory **21** (2): pp. 194-237 (1977).

[20] Lewin, David: *Re: Intervallic Relations between Two Collections of Notes*. J. Mus. Theory **3**, pp. 298-301 (1959).

[21] Lewin, David: *Special Cases of the Interval Function between Pitch-Class Sets X and Y*. J. Mus. Theory **45**, pp. 1-29 (2001).

[22] Mazzola, Guerino, *The Topos of Music*, Birkhäuser, Basel (2003). Reedited 2016.

[23] Mazzola, G., Lubet, A., Pang, Y., Goebel, J., Rochester, C., Dey, S. *Imaginary Time*. In: Making Musical Time. Computational Music Science, Springer, Cham (2021).

[24] Oliver, D.L., Arsie, A.: *Effects of weighting the ends of a tubular bell on modular frequencies*, J. Maths. Mus., **13** (1), pp. 27-41 (2019). DOI: 10.1080/17459737.2018.1508617

[25] Quinn, Ian: *General equal-tempered harmony: parts two and three*. Perspectives of New Mus. **45** (1), pp. 4-63 (2006).

[26] Yust, J., Amiot, E.: *Non-spectral Transposition-Invariant Information in Pitch-Class Sets and Distributions*, in Agustín-Aquino, O., Montiel, M. (eds) Mathematics and Computation in Music, MCM 2022, Atlanta, GA. Springer, Cham (2022).

[27] Yust, Jason: *Applications of the DFT to the theory of twentieth-century harmony*. In: Collins, T., Meredith, D., and Volk, A., Mathematics and Computation in Music, Fifth International Conference, 2015, LNCS, vol. 9110, pp. 207-18. Springer, Heidelberg (2015).

[28] Yust, Jason: *Schubert's harmonic language and Fourier phase space*. J. Mus. Theory **59** (1), pp.121-181 (2015).

[29] Yust, Jason: *Hadamard transforms of pure-duple rhythms*, J. Maths. Mus. **16** (2) pp. 200-215 (2022).