Why sines?

Author: Anonymized

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, musical objects are decomposed into sums of sines :

$$\mathbf{1}_A(t) = \frac{1}{n}\sum_{k+0}^{n-1} a_k \cos(2k\pi t/n) + b_k \cos(2k\pi t/n) \text{ or equivalently } = \frac{1}{n}\sum c_k \exp(2ik\pi t/n).$$

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 [23], 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 and seminal 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 ! [2] 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 [3].

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, since William of Occam's famous razor. But was there more to it than that ? At least in the field of mathematical music theory ? Or is it another instance of the famous saying [4] that 'When your favourite tool is a hammer, all problems look like nails' ? This is the topic of this paper.

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 pc-sets are 12-periodic.
2. 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.
3. Recently, Hadamard transform was used on recursively binary rhythms with finer results than Fourier transform or wavelets, see [29].
4. Known as Maslow's Hammer, see `https://en.wikipedia.org/wiki/Law_of_the_instrument`.

1 Other occurrences

Though this may appear as digressive, we will begin by enumerating more evidence of sine functions in mathemusical publications. It appears in a surprising number of apparently unrelated situations.

1.1 Quantum gauges

There is more than a fashionable trend in some recent efforts at bringing quantum mechanics concepts in the musical world. One barely touched but promising notion is modeling the human mind's choices of interpretations in ambiguous situations (e.g. when an instrumentist strikes a tritone, is it a diminished fifth or au augmented fourth? this will be determined by subsequent observation, see [15]) as *collapsing a waveform*; another more involuted situation uses gauge functions for modelling the notion of tonal attraction (see [11, 12] 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) inasmuch *quantum* mechanics purports to explain discrete changes between levels, and indeed the author must confess (to his shame) that he was at first unconvinced by such explorations. There was however some strong experimental evidence to support this line of research in Carol Krumhansl's seminal work [19], which highlights cosines and Fourier transforms though out of empirical results. Furthermore, this psycho-acoustical evidence is fully in phase with the theoretical aspects of discrete Fourier transform described in the next section. Moreover, this field of exploration is promising and in the context of the present paper, could be subsumed under the heading of section 2.1 where the discretization of sines is explained by limit conditions.

1.2 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 [5] : 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 3.4). Hereafter we present objects with a fixed period of n, which could be for instance periodic rhythms with a duration of n eighth notes, or the seminal case of 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 [6] by another, complex-valued, distribution :

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

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

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

5. See however [8] and section 3.2 on convolution. Also intersection of sequences was often used by Xenakis,which he called *sieving*, while union of disjoint translates of a set are involved in tilings.

6. Other possible definitions differ by a multiplicative constant.

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

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

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 significations 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.

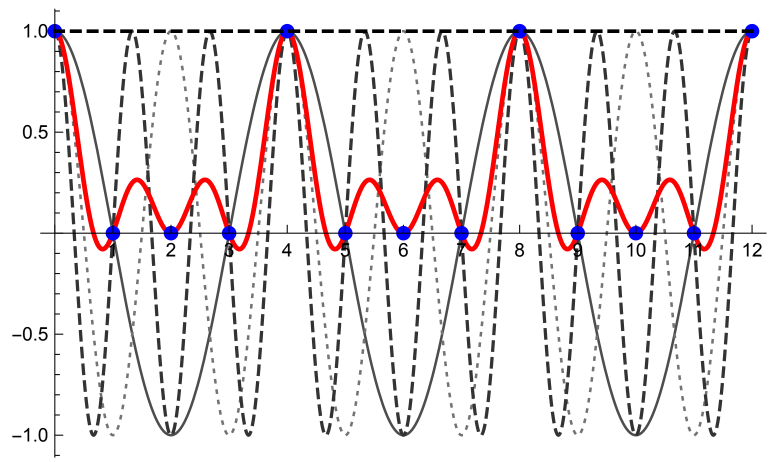

FIGURE 1 – Characteristic function of the diminished seventh $\{0, 3, 6, 9\}$ as $\dfrac{1 + 2\cos\frac{2\pi t}{3}}{3}$. Potential harmonics appear as dotted lines.

Less trivially, for the characteristic function of a pc-set A, this pc-set is Maximally Even [7] if and only if some precise Fourier coefficient has maximal magnitude [8]. 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 4*th* 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* in the sense of [16]), being closest to regular 3- or 5-gons. Again this will be revisited in section 3.1.

Hence the size of a specific Fourier coefficients is a measure of a definite musical character : for pc-sets, diatonicity, chromaticism, even octatonicity [9], a contentious notion for slavic composers, see [7]. For rhythms, the cinquillos and tresillos exhibit a maximal ternary character which dancers will

7. Informally, Maximally Even Sets or ME-sets are the most evenly 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 place on 12 sites on the circle will find a stable configuration if distributed as a diatonic scale.

8. As compared with the same Fourier coefficient for pc-sets with the same cardinality. This was discovered by Ian Quinn [26] in 2005, see also [3]. Such musical significance is in itself a good factual answer to our initial question, but it does not explain *why* this decomposition in sines is significant, or tell whether other decompositions might be as, or more, meaningful.

9. Respectively considering $|\widehat{a}_5|, |\widehat{a}_1|, |\widehat{a}_4|$.

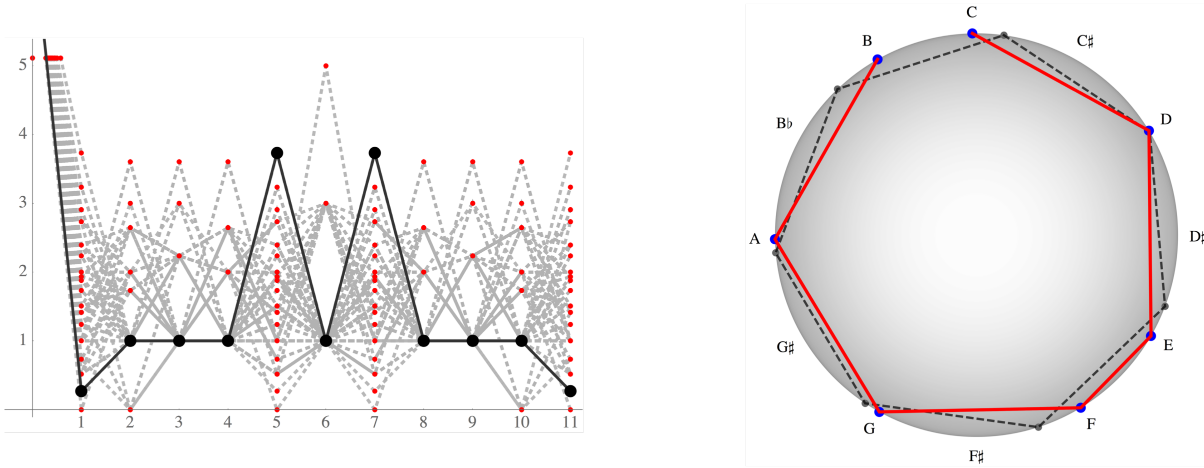

FIGURE 2 – Left : Diatonic scales (black) have maximal fifth coefficient among all 7-note scales (gray, dotted), and are closest to regular heptagons (Right).

appreciate in Caribbean and South-American musics. 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 is extremely convenient (especially when one does not have perfect pitch) and musically significant [10]. 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.

On the whole, many interesting interpretations arise when reading musical objects in the sine basis of Fourier decomposition. But still this does not prove that there does not exist some *other* basis, possibly even more appropriate for musical purposes.

1.3 Anatol Vieru, difference equations, and powers of eigenvalues

There are many situations where the iteration of a transformation asymptotically reduces to a geometric progression. For instance, the Lukas sequence (cousin to the famous Fibonacci sequence) :

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

has for nth element the *integer* closest to $\left(\dfrac{1+\sqrt{5}}{2}\right)^n$. For instance, $\left(\dfrac{1+\sqrt{5}}{2}\right)^7 = 29.0344\cdots$.

Generally speaking, this evidences 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 had found around 1200.

In the same vein, in the 1840's astronomer Urbain Leverrier determined the largest eigenvalue of a 4×4 matrix by studying the asymptotic behavior of its powers, a crucial step in computing the orbit of Neptune. Returning to music, eigenvalues asymptotically occur even more interestingly in several cases :

— 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

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

10. As will be argued below, these magnitudes encapsulate no more and no less than the intervallic information of a pc-set.

and one soon arrives at $(4,4,0,8,8,0,4,4,0,8,8,0)$ which the difference operator permutates circularly. See [9] for a seminal analysis.
- Composer Tom Johnson has explored *autosimilar melodies* : when extracting out of some periodic sequence one note out of (say) 3, sometimes one finds the initial melody [11]. In [5] the question is thoroughly explored mathematically, including 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* in [17], with similar results albeit the transformation is not linear.

These examples (approximately for the last one) mostly exemplify the general behaviour of all linear operators (on modules, not necessarily vector spaces), governed by **Fitting Lemma** :

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

When the space is finite, this means that when the nilpotent part has vanished under iteration, what remains is periodical.

The general idea underlying all these disparate cases is that **from a distance, iterating any linear transformation looks a lot like multiplying the initial vector by a geometric sequence** λ^n (where one may have to extend the initial domain of numbers to find λ, from integers to galoisian extensions [12], from reals to complex, from $\mathbb{Z}_{12}$ to a ring extension). Assume for simplicity's sake that $\lambda \in \mathbb{C}$, then

$$\lambda = |\lambda|\, e^{i\theta}, \lambda^n = |\lambda|^n e^{ni\theta},$$

and the exponential part $e^{ni\theta}$ (what remains after normalization) is a sum of cosine and sine yet again.

Very generally, when a constant appears in an iteration, there is some fixed point involved : for Vieru's persistant sequences we have fixed points of the *shift operator*, Lukas sequence is a fixed point of the operator $(u_n)_{n\in\mathbb{N}} \mapsto (u_{n-1} + u_{n-2})_{n\in\mathbb{N}}$ in sequences space, and in the next section we will point out a fixed point of a differential operator. In these situations, convergence is often geometric, which reminds of the last argument. But all of this is very general, mathematical and abstract. Let us refresh ourselves by turning to the physical nature of musical sound.

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

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

Here is y measures to 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). As we have just mentioned, 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 3.1) up to a constant.

11. Try (F **F** F D **F** F F **D** F F **F** D . . .) for a definitely Beethovenian example. There are even earlier examples.
12. For instance the Lukas sequence above, taken modulo 12, is periodic with period 24 and can be expressed with powers of numbers lying in a quadratic extension of the ambient ring, namely $\mathbb{Z}_{12}[X]/(X^2 - X - 1)$.

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

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

The simplicity of this equation, leading irrevocably 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 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 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 single points as involved in the pure harmonic oscillator). Could it be that the sine solutions are simply products of this oversimplification ?

2.2 D'Alembert's wave 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 modelized and proved with the following *wave equation*, which is derived in simplistic situations (infinitely thin strings, 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 (measured as a distance) and t is time, v is a constant comparable to a velocity ; $u(x,t)$ measures the swelling of the string at abscissa x and moment t, see Fig. 3 which should be understood as changing with time.

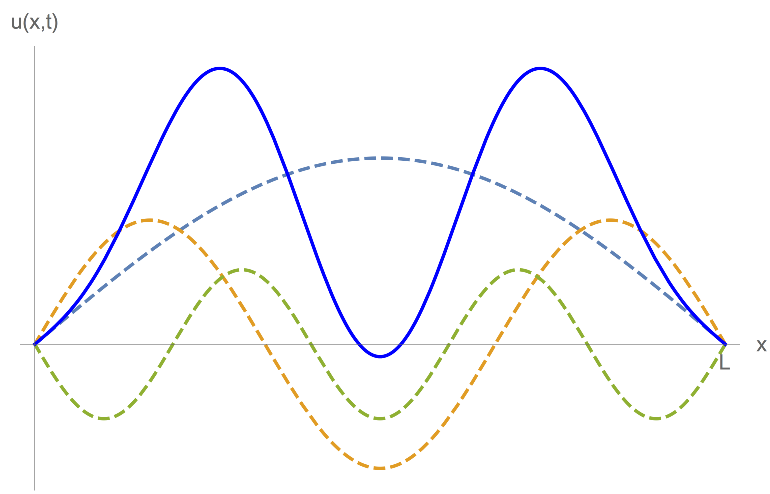

FIGURE 3 – A musical string's shape (strong line) follows a sum of sine vibrations (dotted).

This equation was stated by Jean le Rond d'Alembert in 1729 [13]. It is a *partial differential equation* (PDE), meaning it involves derivatives following time t or 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 [there is no essential difference for an air column like in the case of a wind instrument] attached at both ends, the value of u there is invariable in time, i.e. (for some adequate choice of constants)

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

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

Even in this seminal situation, a rigorous derivation of the most general solution was unattainable at the time [14]. A customary method for such (linear) PDE's 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 in the wave equation one finds

$$\frac{W''(x)}{W(x)} = \frac{1}{v^2}\frac{T''(t)}{T(t)}$$

and a function of x equal to a function of the independent variable t must be a constant, independent of both. For physical reasons [15] this constant must be negative. Let us call it $-\xi^2$, then firstly

$$W''(x) + \xi^2 W(x) = 0, W(0) = W(L) = 0$$

is a harmonic oscillator equation with boundary conditions which impose

$$\xi = \xi_n = \frac{n\pi}{L},\ W(x) = A\sin\frac{n\pi x}{L} \quad \text{for some integer } n.$$

(the cosine part is forbidden by the value in 0, and the integer n is imposed by $\sin(\xi L) = 0$.)
Knowing now the pulsation ξ_n one gets the other function T, solution of

$$T''(t) + \xi_n^2 v^2 T(t) = 0,$$

as a combination of sines and cosines :

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

(A, B, C, D are arbitrary constants than can be determined if initial conditions are provided)
Let us point out the most illuminating points of this resolution :

- Spatial and temporal frequencies are proportional – as Wagner's Gurnemanz shrewdly noticed, "Zum Raum wird hier die Zeit".
- There is discretization of the spectrum, which is *harmonic* : $f_n = nf_1$. 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. 3), but the corollary that the temporal spectrum is harmonic vindicates the pythagoreans' discovery two millenaries before.
- Paradoxically, uncoupling the space and time components forces a coupling of their respective frequencies. Also this reduces the resolution to the ODE case of the harmonic oscillator, with its plain sine solutions, fixed points of the squared differential operator.

Of course, we could again argue that such a reduction to a simplistic equation is a biased argument for the use of sine functions. While the more general solution rebuild from adding up elementary solutions for all values of n already covers a large [16] space of solutions :

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

(leaving aside problems of convergence for the infinite sum, mostly irrelevant in the physical world), we feel compelled to examine a less frugal model.

14. 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 fondation.
15. Else the solutions (hyperbolic sines and cosines) would be unbounded in time, and the instrument would break.
16. Counterintuitively, its dimension is non countable.

2.3 A more complicated case

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

Actually, it is not so. A very interesting illustration can be made with the analysis of [open-ended] tubular bells (also called chimes) published recently in [25] (among other recent papers on the topic). 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 published paper. In substance, it appears that the spectrum is not harmonic [17] (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 [18] when n grows to infinity, but these inaudible harmonics are physically irrelevant). The temporal frequencies are are no more proportional to the ξ_n, but instead to the ξ_n^2. [19]

This study indicates that even in complicated cases, when all else fails, the sines perdure : the superposition method will always yield sines, because these functions are the solutions of linear ODEs with constant coefficients, which generalize the harmonic oscillator to any conceivable order but whose solutions still are [complex] eigenvalues of differential operators, hence sines [20]. The MCM Occam argument actually applies to all solutions of all linear problems (roughly speaking, any oscillating system) in any dimension [21].

3 Musical operations : IFunc and convolution product

We now return to music theory and its abstract, cultural, and perhaps perceptual, concepts.

3.1 Some music : 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 some feature at least. This involves important cognitive capacities which are very useful in other situations, so much so that recent research encourages therapic use of music as intellectual stimulation, from kindergarten kids to Alzheimer patients [10].

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 [22].

The same operation was much discussed in the second half of the XXth century, with interval vectors, contents, and functions as described by Forte [14] or Lewin [20, 22] for instance.

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

18. Actually $\xi_n \approx (n + \frac{1}{2})\pi$.

19. Significantly, ξ_5^2/ξ_4^2 is very close to a tempered fifth, which is quite perceptible in the actual sound of a tubular bell and helps tune them together.

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

21. I leave aside here non-linear equations such as those modelizing drums, which also involve orthonormal families of solutions, like Bessel functions : this is relevant in signl processing, but so far not in music theory.

22. Classically the repetition is not exactly one fifth above, the tonic being substituted for the fifth above the dominant.

For simplicity we focus on the case of the internal IFunc, which is obtained by tallying every possible interval inside a pc-set $A \subset \mathbb{Z}_n$ [23] :

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

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

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

For instance, an octatonic scale satisfies $\text{IFunc}_{\text{Oct}}(3) = 8$, meaning that it is the same when transposed by a minor third ; and a modulation to a dominant turns C major, or $\{0, 2, 4, 5, 7, 9, 11\}$, into G major $= \{0, 2, 4, \mathbf{6}, 7, 9, 11\}$, with 6 notes in common. Hence $\text{IFunc}_{\text{CM}}(7) = 6$, which means that the scale is actually a sequence of fifths and also that a modulation to the dominant will change only one pitch class. Conversely, the notorious pc-set $\{0, 1, 4, 6\}$ shares exactly one pitch-class with any of its transpositions (2 with the tritone transposition). 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, as we will see when examining IFunc as a type of autocorrelation, that is to say measuring an object against its diverse translates.

3.2 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 seminal paper [21]. 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 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) : this quantity tallies the elements of $-B + k \cap A$. In other words [24],

$$\text{IFunc}(A, B) = \mathbf{1}_{-A} * \mathbf{1}_B.$$

Notice that the complexity of this computation is quadratic : one has to effect n times n comparisons. This operation $*$ is of course fundamental in signal processing and its properties are well known [25]. 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.

. *(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 elementary proof can be found in any textbook on the subject. Notice that the left-hand term

23. Or up to a constant the probability of occurrence of an interval when notes of A are played uniformly randomly.
24. $\text{IFunc}(A, B)$ tallies intervals from A to B.
25. It is a 'good' multiplication : Associative, commutative, with a neutral element $\delta = (1, 0, \ldots 0)$.

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.
From Thm. 1 we get immediately

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

This is capital, showing that the magnitude of the Fourier coefficients is equivalent to the intervallic information [26].
Another important formal information is that DFT itself – termwise product 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 maximally 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.

3.3 A last, improbable cosine

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

> Iterating and normalizing autocorrelation on **any** signal converges (quickly) to a cosine [27].

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 [28], after 5 or 6 iterations as can be seen on Figure 4 [online an animated gif can be provided]. 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.

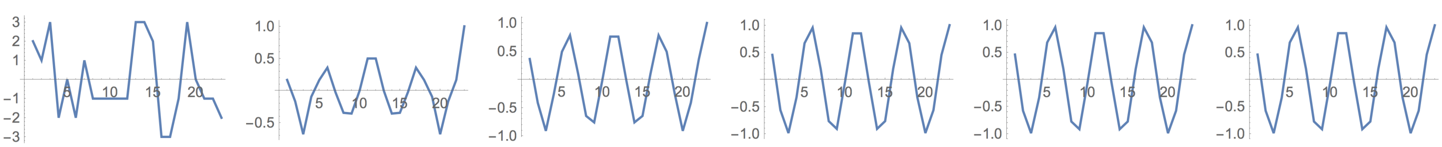

FIGURE 4 – Iterating autocorrelation and normalization always produces a cosine.

Assuming one of the Fourier coefficients is larger than the others, its $2^n th$ power will be infinitely larger than all others. Dividing by this largest value (normalizing), there remains a 1 at its position and vanishing values elsewhere [29]. By inverse Fourier transform one gets back a cosine, with the quick convergence explained by the exponential power 2^n. This portends a stupendous interpretation, that the wired-in comparison (correlation) operation in the brain spontaneously produces cosine waves !
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).

3.4 Mazzola's imaginary time vs. predictive power

Guerino Mazzola [24] has advocated the introduction of a second dimension to time, an imaginary axis purported to help parametrize 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 players.

26. 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.
27. After inquiry it seems that this was already noticed around the 50's by some engineers. It may have been known earlier still.
28. A fixed point of the operation, of course.
29. Actually, there is another 1 at the symmetrical position because the distribution is real-valued. This is but a technical detail.

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 coupling 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 [30], or diatonic and pentatonic (recall these last two have the same magnitudes for Fourier coefficients) or the different classes of pc-sets homometric to $\{0, 1, 4, 6\}$, as recently explored in [18].

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.

3.5 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 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, essentially : no, there is no other one :

. *Let* $E = \mathbb{C}^n$ *be the complex vector space of distributions endowed with convolution product* $*$. *Let* Φ *be some linear isomorphism of* E *such that* $*$ *is turned 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 set such that*

$$\Phi(s)_k = \widehat{s}_{\sigma(k)}.$$

For a detailed proof see [2, Thm. 1.11]. The gist of it is that by composing Φ with inverse Fourier transform, one creates a "licorn", a mathematical object with so many wondrous features that it collapses under its own constrains : 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 essentially 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 permutated. Notice that once again the notion of fixed point is instrumental in the proof.

Conclusion

As shown by the present (partial) survey, the range of Fourier transform in analyzing musical phenomena is bewildering. Although the amount of evidence carries some conviction in itself, it is not enough to prove that alternative bases of decomposition could not do better. However, DFT excels in two domains :

— Meaningfulness. The magnitudes of individual Fourier coefficients have direct musical significance (for instance their collection carries essentially all that information that is invariant under transposition or inversion) and mirrors perceptual information.

30. When considering the phase of $\hat{s}_2 \hat{s}_3 \hat{s}_7$.

— Simplicity. This feature is provably optimal in the sense of Thm. 2, for whatever manipulation of musical data reducible to a convolution operation, including all comparisons diagnosing repetitions, periodicities, similarities, predictability and other cognitive processes known to be involved in musical perception (and creation). On the other hand, modelization or problem solving is easier with economical tools, and sines (or complex exponential functions) are also minimal in this respect, being fixed points of differential operators and asymptotical limits of many natural processes, leading to incontrovertible usefulness in solving ordinary or partial differential equations and discrete difference equations. Indeed the basis of complex exponentials appears in renormalizing geometric sequences, generated by the second simplest operation, multiplication (whether involving numbers or more complex objects, such as matrices).

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

Références

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

[11] 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

[12] 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

[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] Yust, J., and 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)

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

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

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

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

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

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

[25] 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).

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

[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.121–81 (2015).

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