# Interval Content vs. DFT

Emmanuel Amiot(✉)

LAMPS, Perpignan, France
manu.amiot@free.fr

**Abstract.** Several ways to appreciate the diatonicity of a pc-set can be proposed: Anatol Vierù enumerates connected fifths (or semitones, as an indicator of chromaticity), Aline Honing similarly measures 'interval categories' against prototype pc-sets [8]; numerous generalizations of the diatonic scales have been advanced, for instance John Clough and Jack Douthett 'hyperdiatonic' [5] which supersedes Ethan Agmon's model [1] and the tetrachordal structure of the usual diatonic, and many others. The present paper purports to show that magnitudes of Fourier coefficients, or 'saliency' as introduced by Ian Quinn in [9], provide better measurements of diatonicity, chromaticity, octatonicity... The latter case may help solve the controversies about the octatonic character of slavic music in the beginning of the XX$^{th}$ century, and generally disambiguate appreciation of hitherto mostly subjective musical characteristics.

**Keywords:** Diatonic · Chromatic · Octatonic · Saliency
Fourier transform · Stravinsky

## 1 Introduction

Tautologically, the most diatonic seven-note scale is the diatonic scale, i.e. any collection/pc-set translated from $\{0, 2, 4, 5, 7, 9, 11\}$ in $\mathbf{Z}_{12}$. Slightly less obviously, the most diatonic collection in five notes is certainly the pentatonic scale $\{0, 2, 4, 7, 9\}$. But how is one to compare, say, $\{0, 2, 3, 5, 7, 8, 11\}$, $\{0, 2, 4, 5, 7, 9\}$ or $\{0, 2, 4, 6, 7, 11\}$? The question asked here is "how can one measure (with some precise, computable definition) the *diatonic character* of a pc-set?" While we are at it, it costs nothing to ask this question while replacing 'diatonic' with 'chromatic' or 'octatonic' (other adjectives will appear subsequently). Indeed it is a vexed issue (see [11]) whether Stravinsky's music is octatonic; alternatively, it would be nice to appreciate *objectively* the evolution of chromaticity throughout Wagner's Tetralogy (with Tristan in between) and what remains of it in Parsifal – similar questions abound.

Of course several answers have been advanced. We will present some of them through a few examples, and move on to argue why the most recent one, Ian Quinn's "saliency", is the best so far.

Some knowledge of pitch classes and pitch-class sets theory is assumed, alongside with basic music theory – common scales and chords, alongside with familiarity with Western Music. More elaborate machinery will be developed in Sect. 1.2 and later.


O. A. Agustín-Aquino et al. (Eds.): MCM 2017, LNAI 10527, pp. 151–166, 2017.
https://doi.org/10.1007/978-3-319-71827-9_12

### 1.1 Some Examples

Let us focus on four pc-sets occurring at the beginning of Stravinsky's *Rite of Spring.* The first two descending motives articulate C B G E B A i.e. the pc-set $X = \{0, 4, 7, 9, 11\}$. Then D and C♯ are added, making up $Y = \{0, 1, 2, 4, 7, 9, 11\}$; it turns into something messier with chromatic quarts in the bass, that cover the chromatic aggregate. I will complete the sample with the black-keyed motif in measures 9–12, playing C♯ F♯ D♯ with a G♯ thrown in at the end, i.e. $Z = \{1, 3, 6, 8\}$, and the new descending motif in measures 15–17 playing $T = \{0, 1, 3, 6, 7, 8, 9\}$.

Undoubtedly $X$ can be considered diatonic. After all, it is a subset of a major scale – better, *two* major scales. There is, or was, a large current in XX$^{th}$ century Music Theory that focuses on **inclusion** relationships – so-called set-complex theory in American Set Theory, but also the lesser known notion of 'poor' and 'rich' modes by Anatol Vierù [12][1], an independent and fairly well contrived alternative to the previous theory. However, numerous ambiguities arise:

1. How much, *exactly*, is $X$ diatonic? Can we *grade* it?
2. In particular, is it more or less diatonic than other 5-note pc-sets, like $\{0, 2, 4, 7, 9\}$ or $\{0, 2, 4, 5, 7\}$ which are also subsets of diatonic scales?
3. What of sets which are not *exactly included* in a diatonic mode (like $Y, Z$) but *almost*?

Possible answers, clinging to the set relationships of inclusion and intersection, take into account the (maximum) number of common notes between a pc-set and each and every diatonic collection; or the percentage of such common notes averaged over some common basis (the cardinality of the mode, or 7, for instance). In the chosen examples, $Y$ shares six notes $\{0, 2, 4, 7, 9, 11\}$ with C and G major, and six others $\{1, 2, 4, 7, 9, 11\}$ with D major. On the other hand, $Z$ is included in no less than four diatonic scales, (albeit far from the ones that 'neighbored' $X$ or $Y$), so $Z$ should be rated diatonic – but how much so, when we have so many diatonic contexts to choose from?[2] Meanwhile, $T$ intersects three diatonic collections in five notes, five others in four notes and the remaining ones in no less than three notes. How diatonic is that? Is it actually more chromatic? Or octatonic?

I will not waste time advocating *against* the set-theoretical approach, which fails because set-theory is too poor to take into account complex musical notions[3], but rather let the more elaborate models speak for themselves.

[1] In short, in his theory a poor mode is a subset of several rich modes.

[2] Going to extreme cases: is a single note diatonic? What about a minor third?

[3] Among other things, it does not integrate the group structure of intervals modulo octave, not to mention subtler features. As G. Mazzolla wryly observes in the preface of [10], it is hopeless to try and apprehend the huge complexity of music with only the simplest mathematical tools – though this complexity can be reconstructed from *all* its simplifications, if one construes 'simplification' as 'forgetful functor'.

The notion of interval vector (**iv**) is more precise, and provides several illuminating informations on a pc-set.[4] Simply put (following one of the latest of D. Lewin's illuminating comments), it is the probability[5] of hearing a given interval if two pcs are chosen at random in a given pc-set. Then

$$\mathbf{iv}_X(k) = \#\{(a,b) \in X^2 \mid b - a = k\} = \#\big(X \cap (X + k)\big)$$

i.e. the number of occurrences of interval $k$ between elements of $X$.[6]

Since a diatonic collection has maximal value for $\mathbf{iv}(5) = \mathbf{iv}(7) = 6$ (among 7-note scales), it is natural and (important in practice) fairly elementary[7] to compute $\mathbf{iv}_X(5)$ for any pc-set $X$ and compare it against that value.

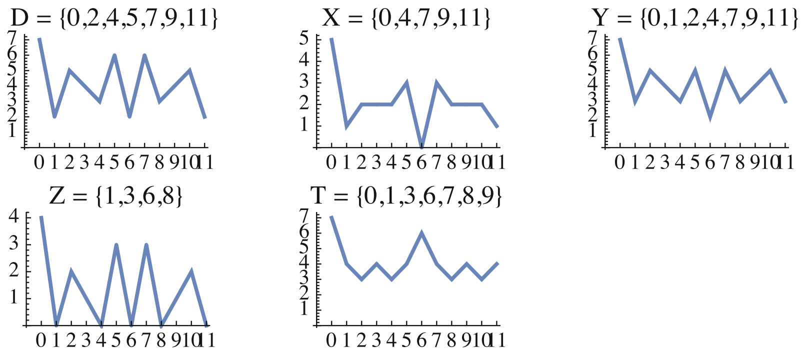


**Fig. 1.** **iv** for the diatonic $D, X, Y, Z$ and $T$

Already **iv** provides some satisfying information (see Fig. 1):

- For $X$, $\mathbf{iv}(5) = 3$ is indeed the maximal coefficient; but it is far below the value for the diatonic scale, which might express the contextual ambiguity (too many different diatonic scales include $X$). On the other hand, $\mathbf{iv}(1) = 1$, the chromatic value, is quite small with only one semitone.
- For $Y$, $\mathbf{iv}(5) = 5$ is almost as large as in the case of a diatonic collection. Notice however that $\mathbf{iv}(2)$ is just as large (many whole tones) and $\mathbf{iv}(1)$ is greater than it would be for a diatonic collection.
- For $Z$, $\mathbf{iv}(5) = 3$ is the largest coefficient and also the maximal possible value for a 4-note scale, confirming the diatonic character despite the contextual indetermination of its many diatonic neighbors.

---

[4] The machinery involved, as we will develop below, is actually an algebra structure (with a convolution product) on the vector space of distributions, i.e. vectors describing how much of C, C♯, D and so on, are featured in a much generalized pc-set.

[5] Up to a constant.

[6] For technical reasons that will be made clear below, we do not take into account the symmetries, e.g. $\mathbf{iv}(n - k) = \mathbf{iv}(k)$ and consider $\mathbf{iv}_X$ as a vector in $\mathbf{R}^n$.

[7] Just check the number of common tones between $X$ and $X + 5$, using the second formula in the definition above.

– Lastly, $T$ is much more contrasted, with $\mathbf{iv}(6)$ a clear maximum[8] and other coefficients between 3 and 4.

This looks fairly close to musical perception, at least as far as diatonicity and chromaticity are concerned. However, let us take a closer look at two hexachords which share the same value for $\mathbf{iv}(5)$ (see Fig. 2): $H = \{0, 2, 4, 5, 7, 11\}$ and $H' = \{0, 1, 5, 6, 7, 8\}$. The first one, $H$, is a subset of C major, the second $H'$ has only five pcs in common with C♯ and G♯ major and appears substantially more chromatic and less diatonic.[9]

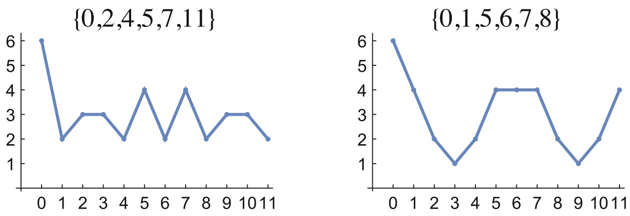


**Fig. 2.** **iv** for two hexachords

This provides evidence that, *at least in some cases*, the **iv** is not good enough to discriminate between different degrees of diatonicity. This requires both elucidation and improvement.

Anatol Vierù went deeper still in his analysis of diatonicity (or chromaticity), and understood the importance of *connectivity of fifths.* In a diatonic (or pentatonic) collection, we face an *uninterrupted* sequence of fifths, e.g. F C G D A E B. In $H, H'$, there are two *broken* fifth sequences, respectively (5, 0, 7, 2), (4, 11) and (5, 0, 7), (6, 1, 8): the first collection $H$ adheres more closely to the generating structure of the diatonic scale than $H'$. Hence Vierù's definition of diatonicity and chromaticity:[10]

**Definition 1.** *The diatonicity (resp. chromaticity) of a pc-set is the maximal number of* consecutive *fifths (resp. semitones) between elements of the pc-set.*

In the above example, $H$ gets 3 and $H'$ only 2, though the values of $\mathbf{iv}(5)$ are the same (4). Will the reader agree that the first is roughly 50% more diatonic than the second? Notice that this value is less obvious to compute than the **iv**, unless one skillfully multiplies[11] the pc-set by 5 and reads the sorted result

[8] Actually overrated since every tritone is tallied twice.

[9] Many other examples can be devised if this one does not sound convincing to you. A more blatant one would be $\{0, 2, 7, 9\}$ vs. $\{0, 1, 7, 8\}$, both with $\mathbf{iv}(5) = 2$.

[10] *"J'ai élaboré un procédé pour mesurer le degré de diatonisme et de chromatisme d'un mode, basé sur la comparaison de la suite des quintes parfaites connexes avec la suite des demi-tons connexes à l'intérieur du même mode."* [12]; Definition 1 is more or less a translation of this.

[11] Vierù had discerned that the two notions are interchanged by multiplication by 5 (or 7) modulo 12, the classical $M_5$ (or $M_7$) operator; and offered thoughtful insights on this dichotomy as expressed by the affine group on $\mathbf{Z}_{12}$.

for chromaticity, which is a way of reading visually the value on the chain of fifths (cf. right half of Fig. 3): the first pc-set turns into $\{10, 11, 0, 1, 7, 8\}$ and the second into $\{11, 0, 1, 4, 5, 6\}$.

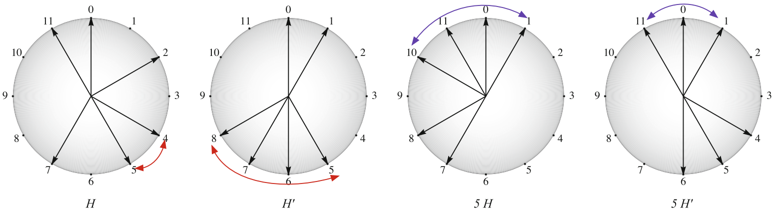


**Fig. 3.** Vierù's chromaticity is lesser in $H$ than $H'$ (left) but diatonicity stronger for $H$, as read on $5H$ and $5H'$ (right)

Let us cut this even finer. We would like to express that $H = \{0, 2, 4, 5, 7, 11\}$ is more diatonic than $H'' = \{0, 2, 4, 5, 7, 8\}$ (and $T = \{0, 1, 5, 6\}$ less than $T' = \{0, 3, 5, 8\}$) though the "Vierù indexes" are identical.

One possible, dual argument, would be that the *covering* chain of fifths is shorter in one case than the other: 5 0 7 2 (9) 4 11 vs 5 0 7 2 (9) 4 (11 6) 1 8 (Fig. 4). This compounds neatly the inclusion criterion, the first scale being a subset of a diatonic and not the second, but at the price of mixing two criterions and enhancing the computational complexity: should we then look up, first the lengths of connected by fifth-components, and then, in case of ex-aequo, the span of the including chain of fifths? This is getting excessively complicated.

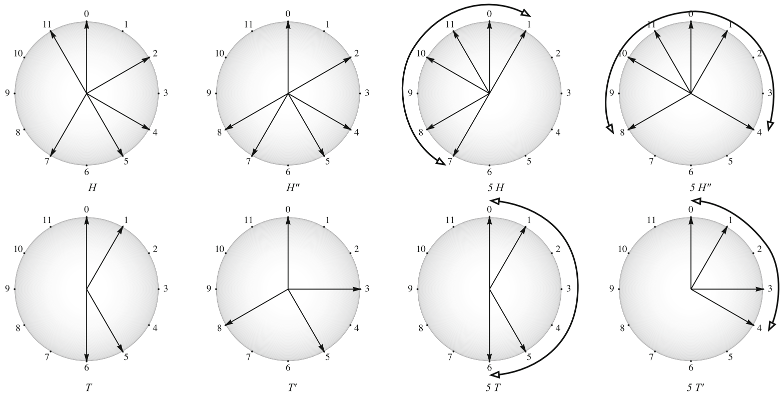


**Fig. 4.** Covering chain of fifths for $\{0, 2, 4, 5, 7, 11\}$, $\{0, 2, 4, 5, 7, 8\}$, $\{0, 1, 5, 6\}$ and $\{0, 3, 5, 8\}$

In [7,8], Aline Honingh endeavors to compare any pc-set with the appropriate 'prototype': for instance a hexachord will be measured against the Guidonian hexachord, a pentachord against the pentatonic, etc. For neatness, the pc-sets are first reduced to so-called 'basic-form'.[12] For instance, the two tetrachords in the last example would be compared with the prototype C D F G (numeric results depend on the choice of similarity measure), which may or may not favor 0 1 5 6 over 0 3 5 8. I will leave the reader to peruse further details in her papers, not because this measure lacks interest, but quite contrariwise (indeed it allows for instance to discriminate between Beethoven's compositions early, middle, and late periods): it gets extremely close to the last, simplest, and overall best candidate.

I present here without any technicity the values of saliency as defined in [9] and used in numerous analyses henceforth. *Saliency* is defined as the magnitude of one easily computed complex number, here (in the case of diatonicity) the fifth Fourier coefficient of a pc-set (formulas, references and properties will follow in the next section). For now, let us appreciate the values of this evaluation of diatonicity for all the above examples and some more. On Fig. 5, we can picture the magnitudes of all Fourier coefficients of the aforementioned heptachords, with the diatonic scale first. We focus on the fifth magnitude (equal to the seventh), highlighted by a dotted horizontal line, and notice that the ranking is: diatonic, Z, Y, X and T with little difference between Y and X, and a larger discrepancy with T.

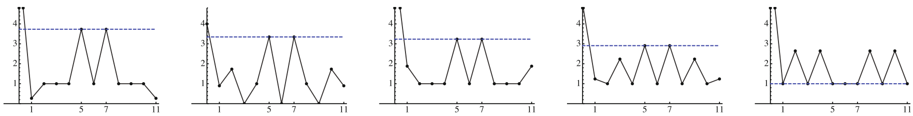

**Fig. 5.** Saliency for the diatonic, $Z, Y, X$, and $T$

A similarly satisfying result also arises with the hexachords on Fig. 6, with an unambiguous ordering of diatonicities: $\{0, 2, 4, 5, 7, 11\}$ followed by $\{0, 1, 5, 6, 7, 10\}$, and last $\{0, 1, 5, 6, 7, 8\}$.

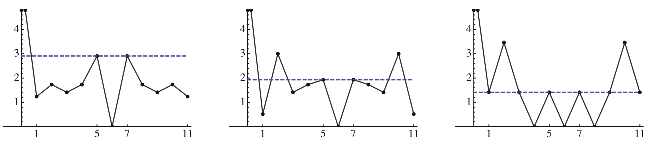

**Fig. 6.** Saliency for the hexachords $H, H', H''$ (horizontal line)

Others examples support unequivocaly this experimental evidence: that the fifth saliency corresponds very closely with the intuitive perception of diatonicity.

[12] In some cases this may not the best for coincidence measurements: the more compact form of a pc-set adresses its chromaticity, not its diatonicity – consider the preceding discussion where the pc-set is first transformed by $M_5$.

We must look into the mathematics to understand why this should be, and above all how this falls in with the competing measurements of diatonicity listed above.

### 1.2 Some Technical Definitions

I provide only a cursory outline; the reader of the present paper will only need to bear in mind that some easily computed[13] quantities, called Fourier coefficients, feature interesting characterizations of those pc-sets which divide the octave as evenly as possible.[14] For a very pedagogical introduction to Discrete Fourier Transform (DFT) of pc-sets, see [4]. For thorough discussion and details, see the recent reference [3] which purports to give the state of the art.

To each pc-set $A$ considered as a subset of $\mathbf{Z}_{12}$, is associated firstly its *characteristic function*
$\mathbf{1}_A : x \mapsto \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$ and second the Discrete Fourier Transform $\mathcal{F}_A = \widehat{\mathbf{1}_A}$ of this function, the DFT of the set:

$$\mathcal{F}_A : t \mapsto \sum_{x \in A} e^{-2i\pi x t/12}.$$

This function is a sum of complex numbers of the form $e^{i\theta}$ which can all be construed as vectors $(\cos\theta, \sin\theta)$ of length 1, whose direction is given by the phase $\theta$. The value $\mathcal{F}_A(k)$ is called the $k^{th}$ *Fourier coefficient*. We will mainly be concerned with its magnitude, i.e. the length of the sum of these vectors.[15]

Here is a list of elementary though useful results without proofs:

- The set $A$ can be reconstructed from the knowledge of the Fourier coefficients $\mathcal{F}_A(k)$.
- $\mathcal{F}_A(12-k) = \overline{\mathcal{F}_A(k)}$ (conjugate complex number).
- $\mathcal{F}_A(t) = -\mathcal{F}_{\overline{A}}(t)$ for $t \neq 0$ ($\overline{A}$ is the complement of $A$).
- $\mathcal{F}_A(0) = \#A$.
- $\sum |\mathcal{F}_A(k)|^2 = 12 \times \#A$.
- The Fourier transform of the (12-dimensional) interval vector $\mathbf{iv}_A$ is the square of the magnitude of $\mathcal{F}_A$:

$$\forall k \in \mathbf{Z}_{12} \quad \widehat{\mathbf{iv}_A}(k) = |\mathcal{F}_A(k)|^2. \qquad (\sharp)$$

Slightly more technical is the Huddling Lemma in [2]: in laymen's terms it states that, the closer the angles $\theta_k$, the larger the sum $\sum_k e^{i\theta_k}$ (the vectors pull roughly in the same direction, coordinating their efforts). We will only need a simple case:

[13] One can compute them online at http://canonsrythmiques.free.fr/MaRecherche/styled/.
[14] Originally discovered by Quinn [9] and formally proved in excruciating detail in [2].
[15] The length of a complex number $x + iy$ is $\|(x, y)\| = |x + iy| = \sqrt{x^2 + y^2}$.

**Proposition 1.** *When the cardinality of $A$ is fixed, $|\mathcal{F}_A(1)|$ reaches maximal value when the elements of $A$ are consecutive [i.e. when $A$ is a chromatic chunk].*

For us the most important result is

**Corollary 1.** *When the cardinality of $A$ is fixed, $|\mathcal{F}_A(5)|$ reaches maximal value when the elements of $A$ are consecutive* in the chain of fifths.

*Proof.* This follows from the relation $\mathcal{F}_A(5) = \mathcal{F}_{5A}(1)$, which results from $5 \times 5 = 1 \mod 12$: hence the elements of $5A$ must be consecutive, which is equivalent to the condition stated.

This is but a special case of Quinn's result:

> **Among all pc-sets with same cardinality $d$, the maximum magnitude for $\mathcal{F}_A(d)$ is obtained when $A$ is a Maximally Even Set (ME set).**

ME sets admit many equivalent definitions [2,5]. We will need only to remember the most important ME sets in $\mathbf{Z}_{12}$:

1. The octatonic scale for $d = 8$.
2. The diatonic scale for $d = 7$.
3. The whole-tone scale for $d = 6$.
4. The pentatonic scale for $d = 5$.

Quinn aimed at a landscape of chords (starting from experimental knowledge) and sketched first the highest peaks. From some kind of continuity principle, it was natural to infer that the height of a chord close to a summit would still be high. Hence the definition of *saliency*, as a quality of proximity to a ME-set (that Quinn called 'prototype'):

**Definition 2.** *The* d-saliency *of a chord $A$ is $|\mathcal{F}_A(d)|$.*

1. Among $d$-chords, saliency is maximal for $d$-ME sets.
2. Remember if convenient that $|\mathcal{F}_A(d)| = |\mathcal{F}_A(12-d)| = |\mathcal{F}_{\overline{A}}(t)|$, hence both diatonic and (non hemitonic) pentatonic scales have maximum saliency for index 5 (namely $2 + \sqrt{3} \approx 3.73$).
3. For any (reasonable) distance on the set of pc-sets, a pc-set close to a ME set has saliency close to maximal.
4. Any pc-set (with given cardinality) distributes its saliencies according to its geometry: the sum of the squares of all saliencies is a constant. This echoes the idea in [8] that the distribution of [IC] categories throughout a piece tells of its local character.

All this provides fairly good mathematical justification, corroborated by empirical knowledge, for defining

**Definition 3.** – *The chromaticity of a pc-set $A$ is $|\mathcal{F}_A(1)|$ (remembering Proposition 1).*

- *The diatonicity of a pc-set $A$ is $|\mathcal{F}_A(5)|$.*
- *The octatonicity of a pc-set $A$ is $|\mathcal{F}_A(4)|$.*

Some other values have actually been used for musical analysis: J. Yust calls 'quartal quality'[16] the magnitude $|\mathcal{F}_A(2)|$ which is, for instance, maximal among octachords for Tristan's motif pc-set $\{2, 3, 4, 5, 8, 9, 10, 11\}$; while the 'major-thirdishness' $|\mathcal{F}_A(3)|$, for want of a better term ('augmentedness'?) is maximal for an augmented triad, or for Schönberg's Napoleon hexachord $\{0, 1, 4, 5, 8, 9\}$.

Remembering the equation $\sum |\mathcal{F}_A(k)|^2 = 12\#A$, it could be argued that the proper measure should be the *squared* magnitude – perhaps averaged by the cardinality – since the sum of all these values is a constant. Also, it is the squared value that appears in the DFT of the intervallic function. I will keep to the original definition for the present paper, but would not be surprised if the squared value were to supersede it in the future (following [17]).

## 2 DFT vs. iv

### 2.1 Theoretical Advantage

DFT is a change of (orthogonal) basis among many (polynomials, wavelets...). The major advantage[17] of expressing a (musical: pc-set, rhythm...) phenomenon in a basis of exponential functions is in the following:

**Proposition 2.** *The DFT exchanges convolution product $*$ and termwise product $\times$. Namely, if $f, g$ are two maps from $\mathbf{Z}_{12}$ to $\mathbf{C}$ and $\widehat{f}, \widehat{g}$ their DFTs, then*

$$\widehat{f * g}(k) = \widehat{f}(k) \times \widehat{g}(k).$$

This is crucial because **iv** is a convolution product:

$$\mathbf{iv}_A(k) = \sum \mathbf{1}_A(t)\mathbf{1}_A(t-k) = \sum \mathbf{1}_A(t)\mathbf{1}_{-A}(k-t) = (\mathbf{1}_A * \mathbf{1}_{-A})(k)$$

and more generally, any coincidence measure or correlation (say, the number of elements of $A$ that lie in any diatonic scale i.e. any transposition $D + k$ of $D = \{0, 2, 4, 5, 7, 9, 11\}$) can also be read on a convolution product:[18]

$$\sum \mathbf{1}_A(t)\mathbf{1}_{D+k}(t) = \sum \mathbf{1}_A(t)\mathbf{1}_D(t-k) = (\mathbf{1}_A * \mathbf{1}_{-D})(k).$$

Now the convolution product is a... convoluted operation[19] while termwise product is straightforward. Cognitively speaking, this means that complicated operations become obvious in Fourier space (i.e. computing on Fourier coefficients) and perhaps suggests that the human mind processes some equivalent of Fourier coefficients.

16 In a convincing study of Ruth Crawford Seeger's *White Moon* [17].

17 This is characteristic of DFT up to permutations: see [3], Theorem 1.11.

18 Yust observed that conversely – by inverse DFT – the number of common tones between two pc-sets can be expressed as a sum of products of magnitudes of Fourier coefficients, pondered by cosines of the differences of phases.

19 It has quadratic complexity, while termwise product is linear.

### 2.2 Multiplying Saliencies

For the sake of simplicity I present computations for diatonicity only[20], i.e. comparing a pc-set $A$ with various transpositions of the Diatonic $D$ and considering the fifth saliency. This is the core of the present article, making sense in a unified way of all previous diatonicity measures. We analyse first the link between coincidence and saliency. Coincidence with a prototype is a variant of Honingh's measure: $\mathbf{1}_A * \mathbf{1}_B(k)$ is a high value when $A + k$ shares many common values with $B$. We are especially interested in the case when $B$ is a diatonic scale, $B = D$ or $-D$ or $k - D$ etc.

Applying Proposition 2 yields immediately

$$\mathcal{F}_A(5) \times \mathcal{F}_{-D}(5) = \widehat{\mathbf{1}_A * \mathbf{1}_{-D}}(5) : \qquad (\sharp)$$

the product of the (diatonic) saliencies of $A$ and $-D$ is a Fourier coefficient of the coincidence function of $A$ and the diatonic scale. Low values of the latter mean that bad correlation will limit the magnitude of $\mathcal{F}_A(5)$, i.e. the diatonicity of $A$. Conversely, when does this coincidence function $\mathbf{1}_A * \mathbf{1}_{-D}$ (replaced below by $\mathbf{1}_A * \mathbf{1}_D$ for simplicity's sake) exhibit a high diatonicity? On the left-hand side of equation ($\sharp$), it means simply that $A$ is highly diatonic (large value of $|\mathcal{F}_A(5)|$). On the right-hand side, it means that the coincidence function $\mathbf{1}_A * \mathbf{1}_D$

1. has at least some large values
2. **and** is 'diatonic' (large fifth Fourier coefficient).

In order to understand how the simple computation of saliency supersedes all previous notions, let us analyse this last feature, which means (in the case of diatonicity) being strongly 5-periodic: the prototype, the diatonic scale $D$, is a chain of fifths, meaning that $D + 5$ has $7 - 1 = 6$ common elements with $D$.[21] From this follows an automatic quasi-periodicity of $\mathbf{1}_A * \mathbf{1}_D$ (see Fig. 7):

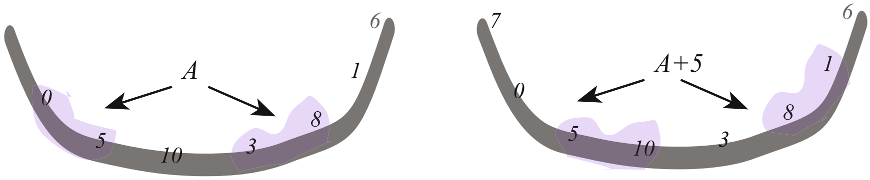


**Fig. 7.** Coincidence between $D$ and $A$ or $A + 5$ changes at most by 1

**Proposition 3.**

*The difference between the correlations* $|(\mathbf{1}_A * \mathbf{1}_D)(k + 5) - (\mathbf{1}_A * \mathbf{1}_D)(k)|$ *is either 0 or 1.*

[20] It would be even simpler for chromaticity (as suggested by a reviewer) but of less interest for actual analysis.

[21] One can use either 5 or 7 as generator of a chain of fifths.

*Proof.* These two convolution products expressed as sums share 6 common elements, plus another one than can be either 0 or 1. More precisely, setting $D = \{5m, m = 0 \dots 6\}$ for simplicity, we get

$$(\mathbf{1}_A * \mathbf{1}_D)(k) = \sum_{m=0}^{6} \mathbf{1}_A(k - 5m) = \mathbf{1}_A(k - 30) + \sum_{m=0}^{5} \mathbf{1}_A(k - 5m)$$

$$(\mathbf{1}_A * \mathbf{1}_D)(k + 5) = \sum_{m=0}^{6} \mathbf{1}_A(k + 5 - 5m) = \sum_{m=0}^{6} \mathbf{1}_A(k - 5(m - 1))$$

$$= \mathbf{1}_A(k + 5) + \sum_{m=0}^{5} \mathbf{1}_A(k - 5m),$$

hence the two values coincide when $\mathbf{1}_A(k + 5) = \mathbf{1}_A(k - 30)(= \mathbf{1}_A(k + 6)$ modulo 12), and differ by one if not.

How then can $\widehat{\mathbf{1}_A * \mathbf{1}_D}(5)$ be as large as possible? On the one hand, the geometry of the diatonic itself partly ensures some periodicity of $\mathbf{1}_A * \mathbf{1}_D$ (Proposition 3), which boosts its diatonicity. How can we further increase this periodicity?

Let for example $k = 0$ in the condition $\mathbf{1}_A(k + 5) = \mathbf{1}_A(k + 6)$ just derived: we will have $\mathbf{1}_A(5) = \mathbf{1}_A(6)$ when *neither F nor F♯* are elements of $A$ (or both), for instance when $A = \{0, 2, 4, 7, 9, 11\}$ (appropriately chiming the first notes of 'Do you know what if means'). But in order to *enlarge* the remaining sum $\sum_{m=0}^{5} \mathbf{1}_A(0 - 5m)$, we will need *as many elements of A as possible* in the partial chain of fifths C D E G A B (each adds 1 to the value of the convolution product). This will certainly be satisfied when $A$ features a long *connected* subsequence of the chain of fifths.[22] We have just understood, not only how the saliency notion includes Vierù's definition, but also why it is superior: Vierù's measure is identical for $H$ and $H''$ but in the latter case the elements of $H$ are better *huddled* in the chain of fifths, providing *a larger tally of large correlation values of the convolution product* $\mathbf{1}_H * \mathbf{1}_D$ (coincidence of $H$ with the prototypical diatonic scale). Let us check this by computing some numerical values. Listing the values of the convolution products from 0 to 11 yields

$$\mathbf{1}_H * \mathbf{1}_D = [\mathbf{6}, 2, 4, 3, 3, \mathbf{5}, 2, \mathbf{5}, 2, 4, 4, 2] \text{ and } \mathbf{1}_{H''} * \mathbf{1}_D = [3, 3, 3, 3, \mathbf{5}, 3, 4, 3, 3, 4, 3, \mathbf{5}].$$

For tetrachords $T = \{0, 1, 5, 6\}$ and $T' = \{0, 3, 5, 8\}$, it is perhaps even clearer:

$$\mathbf{1}_T * \mathbf{1}_D = [2, 2, 2, 2, 3, 2, 3, 2, 2, 2, 2, 4] \text{ and } \mathbf{1}_{T'} * \mathbf{1}_D = [2, 2, 3, 1, \mathbf{4}, 1, 3, 2, 2, \mathbf{4}, 0, \mathbf{4}].$$

Notice in the latter case how the value 4 occurs thrice in a row (in fifth order: at positions $11, 4, 9$), in agreement with the geometric constraint found above. Indeed the 5-saliency of $T'$ is greater than $T$'s. Similarly, $H$ is more diatonic than $H''$ because of the sequence of high values (in fifth order) $\dots 4, 5, 6, 5, 4 \dots$

Of course, computing these correlation vectors with the diatonic would provide an effective and convincing measurement of diatonicity[23]; but as we have

[22] But also *almost connected chains*, like F C G A E B.

[23] As a shrewd reviewer noticed, it would also be feasible to correlate interval profiles, but our aim is to find a recipe at once simple, general and efficient.

demonstrated, the lone and straightforward value of saliency neatly subsumes the whole vector.

### 2.3 Inclusion and *iv*

It is redundant but perhaps useful to synthesize briefly the case of the crude inclusion as compared to saliency in the light of the above calculations. Inclusion of a pc-set inside (say) a diatonic scale is indeed a coincidence measure that can be pinpointed as one large coefficient in $\mathbf{1}_A * \mathbf{1}_{-D}$ (at least one value equal to the cardinality of $A$, some other large values according to Proposition 3). This is but a special case of the preceding discussion, wherein it was shown that significant diatonicity depends not only on the number of coincidences but also on their grouping, or 'huddling'. The same goes for large values of $\mathbf{iv}_A(5)$ (many fifths), which are only indicative of diatonicity when most of the fifths are neighbors in the chain.[24] The extremities of the smallest chain of fifths containing a given pc-set are of course directly related to the number of overlapping diatonic scales – i.e. tally of maximum values of the convolution product –, as foretold in Vierù's notion of 'rich modes'.

### 2.4 Musical Examples

To gain perspective, let us vie away from diatonicity. D. Tymoczko's thoughtful analysis of Stravinsky in [11] draws interpretation of pc-sets towards specific *classes* of scales. To his credit, he acknowledges the numerous ambiguities, criticizes fuzziness in previous analyses and avoids dogmatic pronouncements. Still, dataless statistical sentences like '... [this] scale accounts for **virtually** all of the pitches present' leave room for contestation (I highlighted the adjective). On the other hand, exact measurements of diatonicity as magnitude of $\mathcal{F}_A(5)$ – and all other saliencies – can be compared both within Stravinsky's own music, as it varies within a single piece, and from one piece to another; furthermore, this objective indicator can be applied to other composers (notably Slavic) and provide objective comparisons of their relative degrees of diatonicity, chromaticity, or octatonicity.

The interest of such comparisons warrants general and systematic research that cannot be included in this short paper. Here is but a small sample.

(1) To assess the general appreciation allowed by measurement of saliencies, I have compared all six saliencies (from chromaticity to whole-toneness) on several pieces of *The Rite of Spring* and, as an external reference, the *Dance of the Firebird*. The pieces are imported as MIDI files and a time-window of fixed width moves over it for computation of the saliencies of its pc-sets. Figure 8 simply exhibits the mean values of these saliencies.[25]

[24] The converse is not true: consider CDE which is undoubtedly diatonic though $\mathbf{iv}(5) = 0$!

[25] It appears that there is little difference when the time-span of the window is expanded from 1 to 2 or even 3 s.

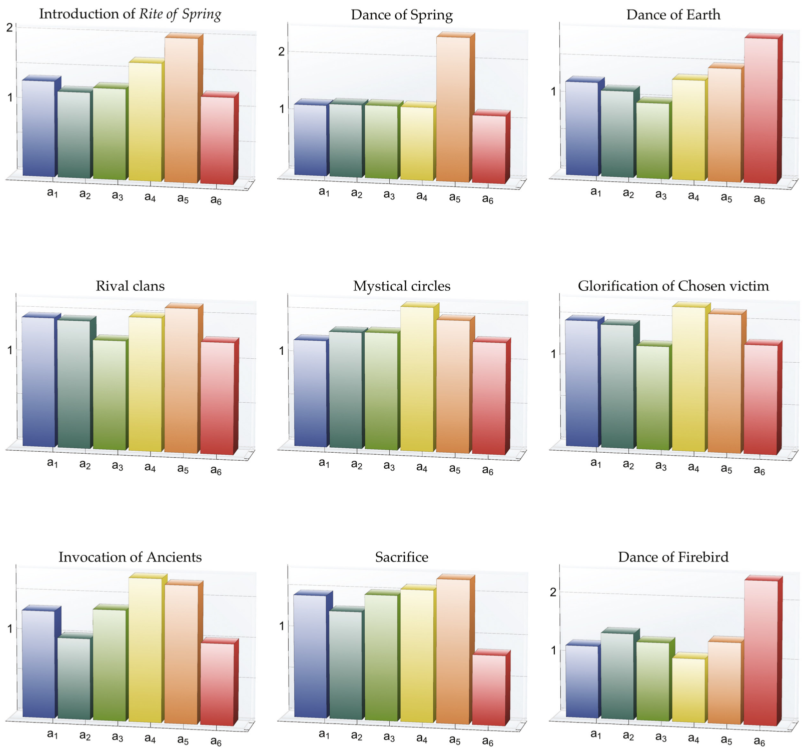


**Fig. 8.** Mean values of saliencies on some Stravinsky pieces

The figures show ambiguity in many pieces, which satisfyingly reflects the diversity of experts' interpretations! However, some clear-cut features do emerge:

1. Whole-tone character dominates *The Dance of the Firebird*.
2. The very first piece of *The Rite of Spring* is fairly diatonic.
3. The *Dance of Spring* is more clearly diatonic.
4. The *Dance of Earth* is mostly whole-tonish.
5. In other pieces, the balance (interplay?) between octatonic and diatonic is apparent – in line with Van der Toorn or Taruskin's analyses (as quoted in [11]).

(2) To give a feeling of the variety of these characters in the flow of the pieces, I provide some excerpts of saliencies as functions of time. On Fig. 9, following the first minute or so of the first movement of *The Rite of Spring*, the saliencies are squared (so that their sum is a constant[26]), and thus it is easily seen which character predominates in a given passage.

[26] Up to the cardinality of pc-sets. On these pictures, the dotted line shows the mean value of a saliency and the solid line a reference value – for $a_5$, say, it is the mean value found for a Mozart Sonata.

It best to look at Fig. 9 while listening to the *The Rite*'s beginning. One can practically see the indecisive first bars (motif $X$) flash a spurt of chromaticism (when the C♯ interferes ca. 6″) before settling for diatonicism (when the D is added to make up $Y = \{0, 1, 2, 4, 7, 9, 11\}$). Then the chromatic fourths around 15″ boost $a_1$; $Z = \{1, 3, 6, 8\}$ occurs between 36″ and 40″, flirting with a pentatonic i.e. largely diatonic character; finally, the last ambivalent motif $T$ is played after 1′, a short surge of chromaticism in a 'quartal' episode (large $a_2$).

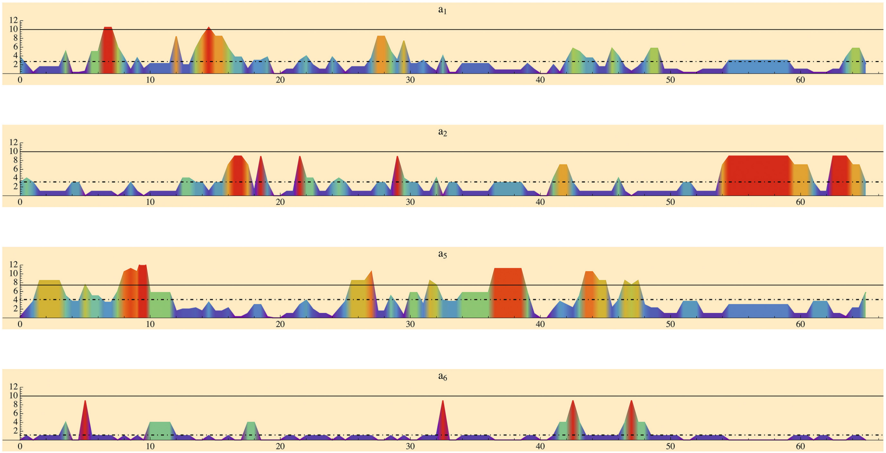


**Fig. 9.** Variations of saliencies in first minute of *The Rite of Spring*

This last moment exemplifies that other segmentations could, and should, be applied to music as it is perceived (as opposed to the music read on the score), for here $T$ is clearly perceptible against the bass, though the numerical computation mixed everything together. Indeed, analyzing separate instruments, or voices, or groups, if justified on perceptual grounds, can lead to finer analyses, see examples in [11,15], and would undoubtedly constitute an easy improvement of saliency analysis.[27]

### 2.5 Phase and Tonality

The (random) colors on these pictures could be adjusted to reflect the *phase* (direction of vectors) of the Fourier coefficient, which reflects a generalization of tonality (for $a_5$ it can be checked against the values for 12 major scales or triads, for $a_6$ it would be against the two whole-tone scales, etc...). Detection of the character of a passage (diatonic, octatonal etc.) can be compounded by pinpointing which (say) diatonic paradigm is involved, by computation of the phase.

27 Hopefully more exhaustive analyses of saliency of Slavic music of early XX$^{th}$ century will soon appear, and settle once and for all the question of their octatonicity.

This is a simple way to detect tonality, and its generalizations (which whole-tone, or octatonic, scale is prevalent, etc.). More about this in [3], Chap. 6.

### 2.6 Possible Applications to Dodecaphonic Music

A hasty reasoning might conclude that the calculations above are meaningless in dodecaphonic music, since the Fourier coefficients of the chromatic aggregate are nil. It is not so. It is certainly true of Nicolai Obouhow's "harmonie totale"[28], but usually false in classical serial music when an appropriate time-span is used for the window of analysis, because the tone-row is often stated horizontally, not vertically; furthermore, at least in the second Viennese school, composition using the two halves (tropes) of the row are frequent. Of course a trope can be any hexachord, with distinctive saliencies, however (essentially this is Babbitt's theorem) *the saliencies of both tropes of a row are identical*. For instance, analyzing both tropes in Alban Berg's *Lyrische Suite* op. 28 and Violin Concerto op. 34 shows very strong diatonic components, see [3], p. 122. I fancy that this is a general feature of Berg's serial music (as opposed to Webern or Schönberg, say) but my ongoing computations have been impeded by the lack of available Midi files for XX$^{th}$ century music.

## 3 Conclusion

From the perspective developed here, one gets a feeling that many worthy researchers have groped for years more or less in the same direction, feeling for the right definition of diatonicity without knowing exactly where it lay. Then came Ian Quinn, and lo! the Holy Grail was there for everyone to grasp.

Not only does saliency pinpoint the character (or lack thereof) of a piece of music, the other component of the Fourier coefficients (the phase) also points its precise direction (the tonality, in the diatonic case).

Precise measurements can, at long last, supersede empirical (at best, with bevies of bored and fallible test subjects) or completely subjective (at worst, and all the more virulent for it) evaluations.

Moreover, this kind of analysis is valid for a huge repertoire, since all that was said here mostly for the diatonic character stands just as well for the 5 other characters. It is hoped that saliency diagrams, pictures and movies will be developed for many pieces of music in the very near future. Indeed, it is only a slight exaggeration to fancy deaf people enabled at last to appreciate music, simply by looking at 'Fourier clocks' ticking as the Fourier coefficients vary throughout a piece![29] It is an urgent task to develop some appropriate software for this kind of streaming analysis, picturing the Fourier flow of music on the fly.

---

28 His chords systematically include all twelve pcs.

29 Technically this is true since the music can be retrieved from the data of all Fourier coefficients.

## References


1. Agmon, E.: A mathematical model of the diatonic system. J. Music Theor. **33**(1), 1–25 (1989)
2. Amiot, E.: David Lewin and maximally even sets. JMM **1**(3), 157–172 (2007)
3. Amiot, E.: Music Through Fourier Space. Springer, Cham (2016)
4. Callender, C.: Continuous harmonic spaces. J. Music Theor. **51**, 2 (2007)
5. Clough, J., Douthett, J.: Maximally even sets. J. Music Theor. **35**, 93–173 (1991)
6. Forte, A.: A theory of set-complexes for music. J. Music Theor. **8**, 136–184 (1964)
7. Honingh, A., Bod, R.: Clustering and classification of music by interval categories. In: Agon, C., Andreatta, M., Assayag, G., Amiot, E., Bresson, J., Mandereau, J. (eds.) MCM 2011. LNCS, vol. 6726, pp. 346–349. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21590-2_30
8. Honingh, A., Bod, R.: Pitch class set categories as analysis tools for degree of tonality. In: Proceedings of ISMIR, Utrecht, Netherlands
9. Quinn, I.: General equal-tempered harmony. Pers. New Music **44**(2), 114–118 (2006). **45**(1) (2007)
10. Mazzola, G.: Topos of Music. Birkhauser, Boston (2004)
11. Tymoczko, D.: Colloquy: Stravinsky and the octatonic: octatonicism reconsidered again. Music Theor. Spect. **25**(1), 185–202 (2003)
12. Vierù, A.: Un regard rétrospectif sur la théorie des modes. The Book of Modes. Editura Muzicala, Bucarest, pp. 48 sqq (1993)
13. Yust, J.: Schubert's harmonic language and Fourier phase space. J. Music Theor. **59**, 121–181 (2015)
14. Yust, J.: Restoring the structural status of keys through DFT phase space. In: Pareyon, G., Pina-Romero, S., Agustín-Aquino, O., Lluis-Puebla, E. (eds.) The Musical-Mathematical Mind. Computational Music Science. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-47337-6_32
15. Yust, J.: Applications of DFT to the theory of twentieth-century harmony. In: Collins, T., Meredith, D., Volk, A. (eds.) MCM 2015. LNCS, vol. 9110, pp. 207–218. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20603-5_22
16. Yust, J.: Analysis of twentieth-century music using the Fourier transform. Music Theory Society of New York State, Binghamton (2015)
17. Yust, J.: Special collections: renewing set theory. J. Music Theor. **60**(2), 213–262 (2016)