4

Saliency

Summary. In the seminal [72], Ian Quinn tries to define a 'landscape of chords' starting from cultural/intuitive knowledge of the most 'salient' chords, and from there infers in a prodigious leap of intuition the existence of a measurable 'chord quality', or saliency, maximal for the prototypical chords. Moreover, he notices that these chords are well known: they are the Maximally Even Sets, i.e. the most even divisions of the octave. In another brilliant intuition, he notices that such pc-sets are characterised by a maximal value of some Fourier coefficient. Thus his vision of a chord landscape is achieved by plotting the magnitude of this Fourier coefficient for all chords (with a given cardinality). Though other measures of chord quality have been devised (Douthett-Kranz, Junod), this notion of saliency will of course be the topic of this chapter.

It is important to mention that this notion applies equally well to periodic rhythms, or any (musical) phenomenon that can be modeled in a cyclic group; for instance, the *tresilo* which is prominent in much of Latin-American dance music will be mentioned below. But since the focus in correlated research has been on scales, I will stick mostly to pc-sets vocabulary and examples.

A selection of Fourier profiles (i.e. magnitudes of Fourier coefficients) of pc-sets is shown in Chapter 8. In this chapter, many references are made to these pictures and the reader is invited to browse the whole collection online at

`http://canonsrythmiques.free.fr/MaRecherche/photos-2/`

(pc-sets are considered up to transposition but not inversion for easier recognition).

Alternatively, the reader is invited to download some software for computing their own Fourier coefficients of any pc-set on

`http://canonsrythmiques.free.fr/MaRecherche/styled/.`

This requires Mathematica™ or the free CDF reader provided by Wolfram Research.

We will study three types of pc-sets with some overlapping between them: saturated scales, generated scales, and maximally even scales. All these highly polarised sets of notes have highly uneven magnitudes of Fourier coefficients; actually, all of them are characterised by some maximum Fourier coefficient. Once this classification is achieved, and some similar/close cases examined, we can move on to the opposite case, flat histogram of either intervals or magnitudes of Fourier coefficients, and prove that the one is flat if and only if the other is too. A seminal case of a flat profile is the aggregate minus one note, which is indeed often tiled by such subsets. Thus the landscape of chords/scales is well described by its peaks and valleys. For instance, the highest peaks in Fig. 4.1 for trichords are augmented triads.

E. Amiot, *Music Through Fourier Space*, Computational Music Science,
DOI 10.1007/978-3-319-45581-5_4

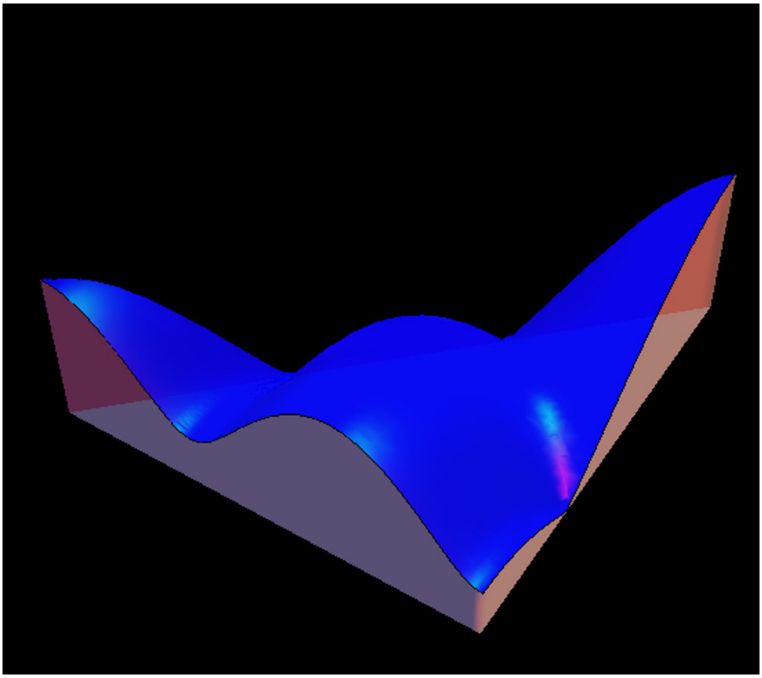

Fig. 4.1. The landscape of trichords

4.1 Generated scales

Much study has been devoted in music theory to the generation of musical scales, whether with just intervals (fifths, thirds) or otherwise. In this section we will consider the monogenous case in equal temperament, according to the following:

Definition 4.1. *A generated scale in* $\mathbb{Z}_n$ *is a subset*[1] *of* $\mathbb{Z}_n$ *generated by some arithmetic progression, i.e.* $A = \{a, a+f, a+2f, \dots a+(d-1)f\}$. *The generating interval*[2]*, or generator, or common difference, is* f*, the starting point is* a*.*

The most famous example is the diatonic scale, generated by fifths (or fourths). Other cases are the non-hemitonic pentatonic (‘Chinese’) scale and the whole-tone scale. These three are maximally even scales (see Section 4.2), which is not the case of the Guidonian hexachord $\{0,2,4,5,7,9\}$ though it is also generated.

[1] We require distinct elements, i.e. A is not a multiset. Of course A can be viewed as a periodic rhythm instead of a scale, but the historical context of study of these subsets being scale theory, the name stuck.

[2] The letter f is chosen as the initial of ‘fifth’, but of course it can take on any value.

4.1.1 Saturation in one interval

Since $a+kf$ can only be connected by an interval of f to $a+(k+1)f$ (upwards) or $a+(k-1)f$ (downwards), the number of occurrences of one given interval in a pc-set cannot exceed the set's cardinality. Conversely, we get the saturation characterisation:

Proposition 4.2. *If a scale A with d elements is generated by interval f, then the number of occurrences of f is $d-1$ or d. The latter case is that of a closed regular polygon. Conversely, a saturated scale is, in the latter case, a periodic subset or a reunion of periodic subsets with the same size (i.e. the orbit of a subgroup of $\mathbb{Z}_n$); and in the former, the same but with one incomplete subcycle.*

The more complicated case of several complete plus one incomplete cycles occurs fairly frequently in 19th century music, cf. the excerpt of Liszt's Piano Sonata in Fig. 4.2 featuring $\{2,5,8,11\}\cup 9$ and $\{1,4,7,10\}\cup 11$. Its Fourier profile appears in Fig. 8.21.

Fig. 4.2. Minor third with multiplicity 4 in 5 notes, in Liszt's Sonata in B.

We will find similar subsets when computing the maximal possible values of the magnitude of Fourier coefficients.

Proof. The number of occurrences of f in $\{a, a+f, a+2f, \dots a+(d-1)f\}$ is clearly at least $d-1$ and can only reach d if $a+df=a$ (in $\mathbb{Z}_n$), which means that $df=0 \mod n$; and hence the scale closes, i.e. A is a regular polygon. Conversely, the pairs $(x, x+f)$ cannot happen more than d times, in which case every single element $x \in A$ plays once the role of x in the pair and once the role of $x+f$, i.e. one has $x+f \in A$ and $x-f \in A$ (equivalently, the map $\tau_f : a \mapsto a+f$ is a permutation of the set A). This means that A is closed under translation by f, i.e. A is an orbit, or a reunion of orbits, of the group $f\mathbb{Z}_n$, i.e. a reunion of translates of $f\mathbb{Z}_n$. With a count of $d-1$ occurrences of interval f, the condition can and must be relaxed on one and only one x, which will satisfy x and $x-f \in A$ but $x+f \notin A$, so that by removing that element we get the same case with both $\#A$ and the number of occurrences of f decremented by one; so the proposition is proved by induction.

4.1.2 DFT of a generated scale

It is easy to compute the DFT of chromatic cluster $A = \{0,1,2,\dots d-1\}$, since all coefficients are sums of geometric series:

$$\mathscr{F}_A(t) = \sum_{k=0}^{d-1} \mathrm{e}^{-\frac{2\mathrm{i}kt\pi}{n}} = \frac{\mathrm{e}^{-\frac{2\mathrm{i}dt\pi}{n}} - 1}{\mathrm{e}^{-\frac{2\mathrm{i}t\pi}{n}} - 1} = \frac{\mathrm{e}^{-\frac{\mathrm{i}dt\pi}{n}}}{\mathrm{e}^{-\frac{\mathrm{i}t\pi}{n}}} \frac{\mathrm{e}^{-\frac{\mathrm{i}dt\pi}{n}} - \mathrm{e}^{\frac{\mathrm{i}dt\pi}{n}}}{\mathrm{e}^{-\frac{\mathrm{i}t\pi}{n}} - \mathrm{e}^{\frac{\mathrm{i}t\pi}{n}}} = \mathrm{e}^{\mathrm{i}(1-d)t\pi/n} \frac{\sin\frac{dt\pi}{n}}{\sin\frac{t\pi}{n}}.$$

Hence the magnitude of the DFT of any generated scale

$$B = fA + \tau = \{\tau, \tau+f, \tau+2f, \dots\}$$

(translation by τ does not change the magnitude, and multiplication by f multiplies the index of the coefficient):

Proposition 4.3. $|\mathscr{F}_B(t)| = \begin{cases} d & \text{if } \sin\frac{f\pi t}{n} = 0 \text{ (i.e. } n \mid ft) \\ (\pm)\dfrac{\sin\frac{fd\pi t}{n}}{\sin\frac{f\pi t}{n}} & \text{else} \end{cases}.$

For instance the value of $|\mathscr{F}_A(5)|$ when A is a diatonic scale is

$$-\frac{\sin\frac{5\times 7\pi 5}{12}}{\sin\frac{7\pi 5}{12}} = \frac{\sin\frac{7\pi}{12}}{\sin\frac{\pi}{12}} = \frac{1}{\tan\frac{\pi}{12}} = 2+\sqrt{3}.$$

It is obvious that the first case, d, is the maximum possible value, especially when one remembers that we just summed d complex numbers $\mathrm{e}^{-\frac{2\mathrm{i}kft\pi}{n}}$, all of them with magnitude 1. It is perhaps less obvious that the reciprocal is true (for the moment, we consider only generated scales): if any of the exponentials in the sum defining the Fourier coefficient do not have the exact same direction, then their sum has a smaller length than the sum of their lengths:

Lemma 4.4. *For $a,b \in \mathbb{C}$, $|a+b| = |a| + |b| \iff a,b$ have the same direction, i.e. $\exists\lambda \in \mathbb{R}_+, b = \lambda a$ (unless $a = 0$).*

So when the magnitude of the Fourier coefficient is maximum, all exponentials in it share the same direction. But equality of the phases of all $\mathrm{e}^{-2\mathrm{i}fkt\pi/n}$ means that $n \mid ft$, i.e. we are in the first case when $\sin\frac{f\pi t}{n} = 0$.

The other extreme case is $\mathscr{F}_B(t) = 0$, when bdt is a multiple of n but bt is not. Let us clarify the behavior of these values. Jason Yust noticed the periodicity of these coefficients:

Proposition 4.5. *Fix the generator f and the index of the Fourier coefficient, t. Then the magnitude[3] of this Fourier coefficient is periodic in the cardinality d of the generated scale: $d \mapsto |\mathscr{F}_B(t)|$ has period $\dfrac{n}{\gcd(n,ft)}$.*

For $n = 12$, this period boils down to:

[3] The complex Fourier coefficient itself is either periodic or anti-periodic.

- n/r, where r is the integer closest to 0 and congruent to $\pm ft$; and
- no period (i.e. period 12) when ft is coprime with 12 (for instance, $\mathscr{F}_B(1)$ for fifth-generated scales has no period).

A few examples will show how simple this is:

Example 4.6. Consider first chromatic clusters, like $\{0,1,2\}$, with generator 1 and let us look at $\mathscr{F}_B(4)$ as a function of the cardinality d: $|\mathscr{F}_B(4)| = \left|\frac{\sin(d\pi/3)}{\sin(\pi/3)}\right| = \psi(d)$ and ψ is 3-periodic ($|\sin|$ being π-periodic). Indeed the values taken for $d = 1,2,3\ldots$ are $1,1,0,1,1,0,1,1,0,1,1\ldots$.

For a less trivial case, take coefficient 5 and generator 2 (whole-tone scale chunks). Since $2\times 5 = 10 = -2 \mod 12$ we have $r = 2$, period 6, and indeed for $d = 1,2,3\ldots 11$ we compute $|\mathscr{F}_B(5)| = 1,\sqrt{3},2,\sqrt{3},1,0,1,\sqrt{3},2,\sqrt{3},1$. The associated pc-sets appear in the tables as Figs. 8.4, 8.8, 8.15, 8.20, and 8.23.

A more complicated case where Yust's rule of thumb does not apply: let $n = 24$ and $f\times t = 7\times 2 = 14$. Then the period is 12.

Lastly, a rhythm example: consider generator 3 in an eight beats bar; the *tresilo* (0, 3, 6) (modulo 8) is such a generated rhythm, with $d = 3$. The value of the Fourier coefficient $|a_3|$ takes on magnitudes $\frac{\sin 9d\pi/8}{\sin 9\pi/8}$, which is maximum when $d = 4$ for rhythm (0, 1, 3, 6). In general, $d \mapsto \frac{\sin\frac{fd\pi t}{n}}{\sin\frac{f\pi t}{n}}$ will be maximum when fdt is as close as possible to $n/2 \mod n$.

The proof of this periodicity lies in the formula in Proposition 4.3. Amusingly, Yust's shortcut for $n = 12$ works for the same reason that Lemma 4.20 below is true.

Another beautiful relationship between the chromatic case (generator 1) and the general case (generator f) is

Theorem 4.7 (P. Beauguitte, 2011). *Let $A_k = \{0,1,2\ldots k-1\} \subset \mathbb{Z}_n$. For k coprime with n, let $\ell = k^{-1}$ be the multiplicative inverse of k modulo n and $B = -kA_\ell = \{0,-k,-2k\cdots -k(\ell-1)\}$ the ℓ-scale generated by $-k$. Then $\mathscr{F}_B = 1/\overline{\mathscr{F}_A}$, i.e. the coefficients of one scale are the inverses of the coefficients of the other.*[4]

The choice of ℓ will be clarified below with the definition of ME sets. A common example with $k = 7, n = 12, \ell = 7$ yields the diatonic scale, but in general, the two scales have a different number of elements.

Proof. $\mathscr{F}_A(t) = 1 + e^{-2i\pi t/n} + \ldots e^{-2i\pi(k-1)t/n} = \dfrac{1-e^{-2i\pi kt/n}}{1-e^{-2i\pi t/n}}$ and

$\mathscr{F}_B(t) = 1 + e^{2i\pi kt/n} + \ldots e^{2i\pi(\ell-1)kt/n} = \dfrac{1-e^{2i\pi k\ell t/n}}{1-e^{2i\pi kt/n}} = \dfrac{1-e^{2i\pi t/n}}{1-e^{2i\pi kt/n}}$, hence the result by inverting the fraction and the phases.

This remarkable result shows that for many generated scales, the direction of the DFT is the same as for a chromatic sequence, whilst the magnitude is inversed. This appears clearly in Fig. 4.3, with $n = 10, k = 3, \ell = 7$:

[4] Except of course for index 0 which is the cardinality of the scale.

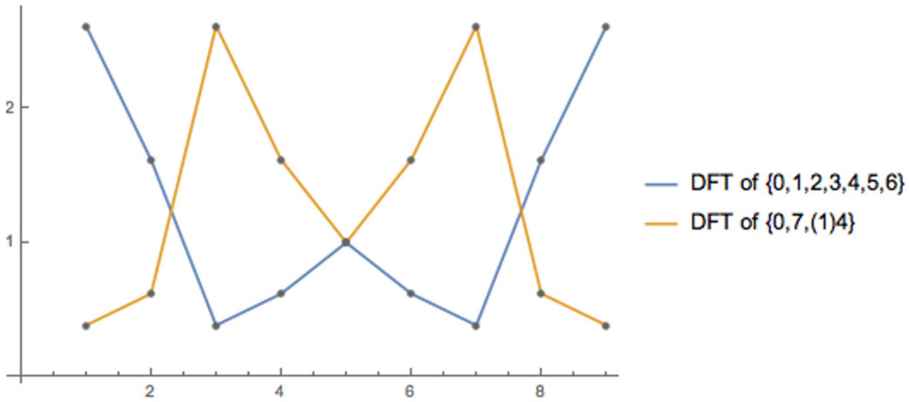

Fig. 4.3. Beauguitte's theorem: inverse magnitudes of two generated scales in $\mathbb{Z}_{10}$.

The saturation feature is linked with the *probability of occurrence* of intervals: in diatonic music, the fifth is more probable than other intervals (if the probability of any pitch-class is uniform, which admittedly is seldom the case except perhaps in strict dodecaphonic, non-serial music), as checked experimentally in [58] for instance. This suggests, in a broad sense, that generated scales are somewhat *periodic* and might be recognised by Fourier features. This is precisely the topic of the maximally even sets section below. For more about occurrences of intervals and their relationship with Fourier coefficients, see Section 4.3.

4.1.3 Alternative generators

Notice the extreme cases (first pointed out, to the best of my knowledge, by N. Carey in [28] wherein the first case of Theorem 4.8 is also proved) when f is a generator of $\mathbb{Z}_n$, and A is the whole aggregate, or $d = n-1$, i.e. A is the whole group $\mathbb{Z}_n$ minus one element. In this case, A has $\varphi(n)$ distinct generators[5] (and as many starting points), which is a somewhat unexpected behaviour for arithmetic sequences. For instance, the aggregate from C to $B\flat$, e.g. $\{0,1,2,3\dots 10\}$, can be written as four distinct arithmetic sequences:

$(0,1,2\dots 10),\ (4,9,2,7,0,5,10,3,8,1,6)$ and their reverses, with generators $11,7$.

This can be seen in Fig. 4.4 with 6 different generators for a 7-scale in $\mathbb{Z}_{21}$.

The converse is true:

Theorem 4.8. *[Amiot, 2011] The number of generators of a generated scale is always a totient number, i.e. $\varphi(n)$ for some n.*

More precisely:

[5] Remember φ is Euler's totient function, which gives precisely the number of generators of a cyclic group.

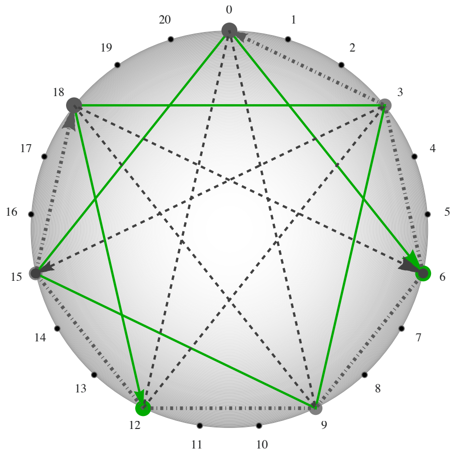

Fig. 4.4. Many generators for a regular polygon

◇ *if f is coprime with n then A has exactly two generators $\pm f$, unless A is the full aggregate ($A = \mathbb{Z}_n$) or the almost full ($\mathbb{Z}_n \setminus \{u\}$).*

◇ *if f is not coprime with n, the generated scale A with cardinality $d > 1$, has*

- *one generator when the scale is (a translate of) $\{0, n/2\}$ (a tritone);*
- *two generators (not coprime with n) when d is strictly between 1 and $n' - 1 = (n/m) - 1$ where $m = \gcd(n, f)$;*
- *$\varphi(d)$ generators when $d = n' = n/m$, i.e. when A is a regular polygon;*
- *$\varphi(d+1)$ generators when $d = n' - 1$, A is a regular polygon minus one vertex.*

The last two cases are those of a full or almost full regular polygon, whose picture is the same as the full or almost full aggregate but for a smaller cardinality $n' \mid n$. Moreover all generators share the same order in the group $(\mathbb{Z}_n, +)$.

Proof. First consider the case of a generator f coprime with n. Up to multiplication by the inverse f^{-1} of this generator modulo n and translation, we are dealing with the chromatic sequence $A = \{0, 1, \dots d-1\}$ and we are looking for an alternative generation to the obvious one (generator 1). So let us assume that A can also be generated as $A = \tau + b \times \{0, 1, 2 \dots d-1\} = bA + \tau$ and let us prove that $b = \pm 1$. My original proof made use of the interval vector of A, which is $(d, d-1, d-2 \dots d-2, d-1)$. An alternative one, more appropriate in the context of this book, uses the DFT:[6]

[6] Incredibly but appropriately, a recent formula [78] expresses the totient function as the DFT of the GCD: $\varphi(n) = \sum_{k=1}^{n} e^{\frac{2i\pi k}{n}} \gcd(n,k) = \sum_{k=1}^{n} \cos\left(\frac{2\pi k}{n}\right) \gcd(n,k)$.

$$\mathscr{F}_A(t) = \sum_{k=0}^{d-1} \mathrm{e}^{-2\mathrm{i}kt\pi/n} = \frac{\mathrm{e}^{-2\mathrm{i}dt\pi/n} - 1}{\mathrm{e}^{-2\mathrm{i}t\pi/n} - 1},$$

$$\mathscr{F}_{bA+\tau}(t) = \sum_{k=0}^{d-1} \mathrm{e}^{-2\mathrm{i}(bk+\tau)t\pi/n} = \mathrm{e}^{-2\mathrm{i}\tau t\pi/n} \frac{\mathrm{e}^{-2\mathrm{i}bdt\pi/n} - 1}{\mathrm{e}^{-2\mathrm{i}bt\pi/n} - 1}.$$

It is sufficient to focus on the magnitudes: since $|\mathrm{e}^{-2\mathrm{i}\varphi} - 1|$ is equal to $|2\sin\varphi|$, the respective magnitudes are

$$\frac{\sin(d\pi/n)}{\sin(\pi/n)} \text{ and } \frac{\sin(bd\pi/n)}{\sin(b\pi/n)} \qquad (0 < d < n).$$

(I removed the absolute values for readability). Replacing b if necessary by $n-b$ without changing the magnitude, one may assume without loss of generality that $b \in \{0, 1 \dots n/2\}$. A cursory study of next-to-maximum values[7] of function $f : b \mapsto \dfrac{\sin(db\pi/n)}{\sin(b\pi/n)}, 1 \leqslant b \leqslant n/2$ (see Fig. 4.5) proves that b must be equal to 1 for the respective magnitudes to coincide, hence $b = \pm a$. Let us now consider f non co-

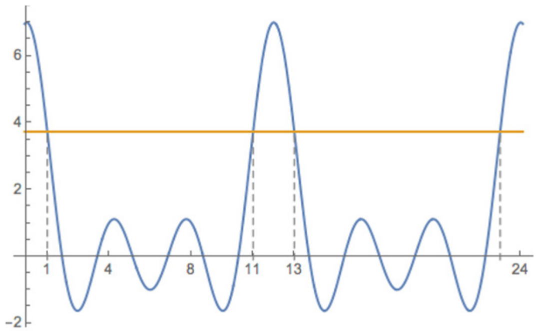

Fig. 4.5. Graph of $f : b \mapsto \dfrac{\sin(7b\pi/12)}{\sin(b\pi/12)}$

prime with n, i.e. $m = \gcd(n, f) > 1$. The cardinality of A is now less than n/m, since $\dfrac{n}{m} f = 0 \mod n$. The difficult question is: do we reach the same m if we start from another generator? But with a computation similar to the one above, if A is generated by f then

$$|\mathscr{F}_A(t)| = \begin{cases} \dfrac{|\sin(\pi d t f/n)|}{|\sin(\pi t f/n)|} & \text{or} \\ d & \text{when } \sin(\pi t f/n) = 0. \end{cases}$$

Moreover, $|\mathscr{F}_A(t)| \leqslant d$, and $|\mathscr{F}_A(t)| = d \iff \sin(\pi t f/n) = 0$. This entails the following:

[7] They occur for $b > \frac{2n}{d}$ and hence $f(x)$ does not exceed $\frac{1}{\sin(2\pi/d)}$, well under $\frac{\sin(d\pi/n)}{\sin(\pi/n)} = f(1)$.

Lemma 4.9. *If f, g are two generators of a same scale A, then*

$$m = \gcd(n, f) = \gcd(n, g).$$

NB: this lemma can also be reached algebraically, by considering the *group of differences*[8]

$$\Delta^\infty(A) = \lim_{n\to\infty} \Delta^n(A) = \bigcup_{n\geqslant 1} \Delta^n(A) \text{ where } \Delta(X) = X - X = \{x - y, (x, y) \in X^2\}.$$

This shorter but more abstract proof was used in [9].

Now the end is easy: up to translation, assume A contains 0. Then $A = mA'$ where the elements of A' are defined modulo $n' = n/m$, and we are back to the initial case $\gcd(n', f) = 1$ when we have only two generators, except if A' is an (almost) full aggregate. This yields the theorem.

Leaving aside the extreme cases of one-note scales and tritones, the geometry of generated scales comes in three types:

- The seminal case: ‘diatonic-like scales’, i.e. scales with only two (opposite) generators.
- Regular polygons.
- Regular polygons minus one note.

So this seminal case, with one beginning and one end, is by no means the only one. The three cases are summarised in Fig. 4.6.

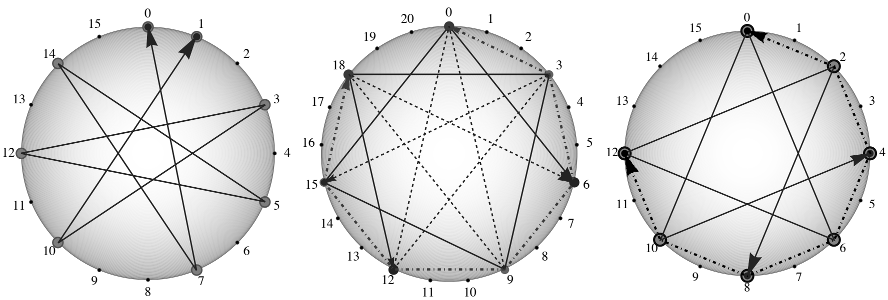

Fig. 4.6. The three cases: seminal, polygon and almost-whole polygon

4.2 Maximal evenness

Maximally even sets, or ME sets for short, were introduced in [31, 30] and developed by Jack Douthett and other co-authors. In the context of this book, his most

[8] [65], Section 7.26.

interesting paper is [33] wherein a ME set is described and defined as an equilibrium position for (say) electrons placed on several equally disposed sites on a circle; it is impressive that seven electrons on 12 sites will choose to settle as a diatonic scale!

There are many possible definitions of maximal evenness, an intuitive notion but a tricky one to nail down: see [31, 32, 38, 24]. The most practical appears in the seminal [30] as a consequence of more philosophical constraints:

Definition 4.10. *A maximally even set with cardinality d in* $\mathbb{Z}_n$ *is the set of values of one of the following J-functions:*

$$J^{\alpha}_{d,n}(k) = \lfloor \alpha + \frac{kn}{d} \rfloor \quad \mathrm{mod}\, n, k = 0, 1 \ldots d-1.$$

One can choose the round function instead of the floor function (or ceiling) with equivalent results. This formula approximates exact divisions of n into d parts, which is of course impossible to do exactly unless $d \mid n$.

Example 4.11. Depending on the offset α, the $J^{12}_{7,\alpha}$ generates the 12 major scales (in fifth order), for instance

$$J^{0}_{7,12}(\llbracket 0,6 \rrbracket) = \{0,1,3,5,6,8,10\}, \text{ i.e. D}\flat \text{ major,}$$

whereas C major is generated by $J^{5}_{7,12}(\llbracket 0,6 \rrbracket)$.

4.2.1 Some regularity features

It is possible to define the class $\mathrm{ME}_{n,d}$ as the generic ME set with d elements in $\mathbb{Z}_n$, because this class is invariant under the action of T/I: any ME set in the class is translated (and also inversed) from any other one.[9] It follows that the number of different ME sets with given (n,d) is a divisor of n, depending on inner periodicities in the set. We will see also that the complement set of a ME set is still a ME set.

An aesthetically remarkable feature of ME sets is the precise quantity of variants of intervals between consecutive elements, or more generally of typed subsets. This is better explained with an example: consider $\{0,2,4,7,9\} = \mathrm{ME}_{12,5}$. Consecutive intervals, or steps, come in exactly two sizes (2 or 3). The same is true for ‘thirds’, leaving every odd note out: they are $4-0=4, 7-2=5, 9-4=5, 0-7=5, 2-9=5$.[10] Similarly, consecutive triplets like $(0,2,4), (2,4,7), (4,7,9)$ come in three configurations, as do the ‘triads’ $(0,4,9), (2,7,0), (7,0,4)$ and so on. When this cardinality of a subset of the scale is always equal to the variety of different instances of the type of subset (‘Cardinality=Variety’), the scale is said to be Well-Formed, henceforth WF for short. See [28] for much more on this subject. ME sets are WF, or degenerate-WF; for instance the whole-tone scale $\mathrm{ME}_{12,6}$ has only one step size, not two.

One definite advantage of the definition of ME sets in terms of DFT below is that it makes obvious that the complement of a ME set is a ME set. Indeed, from the

9 This will be proved easily with the alternative DFT definition provided below.

10 Tymoczko points out these ‘thirds’ in pentatonic context in the last phrase of Debussy’s *La Fille aux cheveux de lin*.

typology below or the J-function definition one easily gets the following paradoxical statement:

Theorem 4.12. *Let $A \subset \mathbb{Z}_n$ be a ME set and $B = \mathbb{Z}_n \setminus A$ its complement. Then B is a ME set; moreover, some translate of B is included in A or the reverse.*

As I mentioned and proved in [9], this 'Chopin's theorem' holds *mutatis mutandis* for generated scales: when a scale and its complement are both generated, they share their set of generators. This is of course reminiscent of Babbitt's theorem. The reference to Chopin of course alludes to his *Etude op. 10, n° 5*, cf. Fig. 4.7, wherein the pentatonic played throughout the piece by the right hand is a subset of the major scales (mostly G♭ and B♭) played by the left hand.

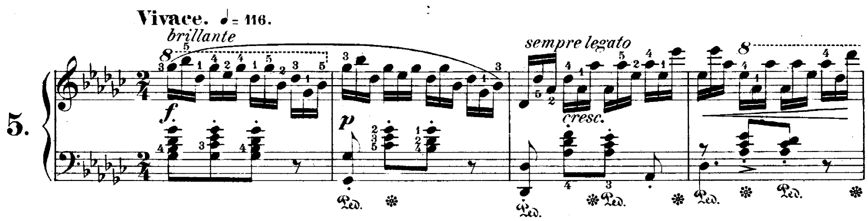

Fig. 4.7. Pentatonic vs. diatonic

A nice application to rhythms is Astor Piazzolla's use of the complement of tresilo $T = \{0,3,6\} \subset \mathbb{Z}_8$: he uses the pattern $C = \{1,2,4,5,7\}$, not only in its function of complement of T, which is fairly common in post 1950s-tango, but also as a basis for a secondary theme in *La Milonga del Angel*. As discussed in Theorem 4.12, the ternary pulsation is present also in this complement rhythm, see Fig. 4.8 which shows how 'the silence in tango is still tango'.

4.2.2 Three types of ME sets

A fine distinction

In [72], Quinn introduces a typology of ME sets, depending on $m = \gcd(d,n)$. We reproduce this classification here, since it is closely related to questions of inner periodicities and complementarity, qualities that can actually be diagnosed at a glance on the DFT. The seminal case is

Definition 4.13. *A type I ME set happens when $m = 1$. The scale is generated (and WF).*

It is generated by the multiplicative inverse f of d in $\mathbb{Z}_n$, or by $-f$ (these are the only two generators). Typical examples are the diatonic and pentatonic scales in $\mathbb{Z}_{12}$. *All its Fourier coefficients are non zero* (a trivial consequence of Theorem 4.7).

Fig. 4.8. Tresilo and its complement in Piazzolla's *Milonga del Angel*

Definition 4.14. *A type II_a ME set happens when $m = d$, i.e. $d \mid n$. The scale is generated, but it is degenerate WF, dividing $\mathbb{Z}_n$ into a regular polygon.*

Typical examples are the diminished seventh $D7 = \{0,3,6,9\}$ (Fig. 8.10) and whole-tone scale $WT = \{0,2,4,6,8,10\}$ (Fig. 8.23). The DFT is quite characteristic: coefficients are 0 except those whose index is a multiple of n/d, which are all equal to d. For instance, for a diminished seventh it is $(4,0,0,0,4,0,0,0,4,0,0,0)$.

Definition 4.15. *A type II_b ME set happens when $1 < m = n - d < d$. It is the complement of a type II_a ME set.*

Since the complement has cardinality m, which is a divisor of n, it is ME because the complement of a ME set is a ME set (proved below). The prototype is the octatonic collection $\{0,1,3,4,6,7,9,10\}$ (Fig. 8.31). Its DFT is the same as type II_a (except of course the 0^{th} coefficient).

Definition 4.16. *Type III ME sets gather the remaining cases: $1 < m < d, m \neq n - d$.*

The DFT is a compound of the two other types: the varied values of the DFT are the same as in type I, with 0's interspersed because of its periodicity (remember the formula for oversampling, cf. Fig. 1.5). For instance $\{0,2,4,6,9,11,13,15\} = \mathrm{ME}_{18,8}$ (Fig. 4.9) yields coefficients (magnitudes)

$$(8,0,1.06,0,1.3,0,2,0,5.76,0,5.76,0,2,0,1.3,0,1.06,0).$$

This classification in three types is stable by complementation.

The last two classes are ME sets with a smaller period, i.e. what Messiaen called Limited Transposition Modes. They are all concatenated from smaller ME sets.

Remark 4.17. Clampitt *et alii* [28] argue that type I is fundamental, inasmuch as this type generates all others: type III is obtained by slicing n into m equal parts and filling each part with the same type I ME set with d' notes among $n' = n/m$, see Fig. 4.9.

Remark 4.18. Types II and III are 'perfectly balanced' in the sense of [67], i.e. $a_1 = 0$ (they are unions of regular polygons). Note that this perfect balancing, a pure Fourier quality, fails to characterise ME sets: for $\mathrm{ME}_{(12,7)}$,

$$|a_1| = \frac{\sin(\pi/12)}{\sin(7\pi/12)} = 2 - \sqrt{3} \approx 0.26795$$

is not the smallest value for seven-note scales, superseded by $\{0,1,2,5,6,8,9\}$ for which $a_1 = 0$, cf. Fig. 8.28.[11]

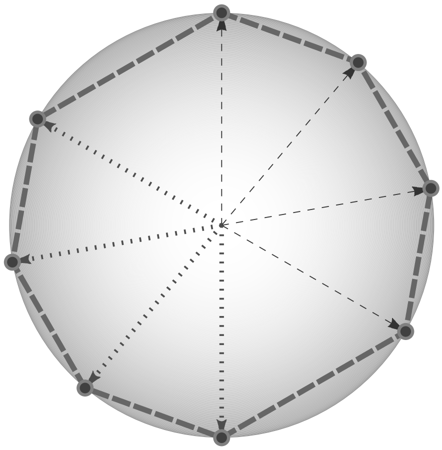

Fig. 4.9. A type III ME set : $\{0,2,4,6\} \oplus \{0,9\} \subset \mathbb{Z}_{18}$

Existence of type III ME sets

Quinn ([72]) was remarkably astute in this taxonomy, since as he himself pointed out there are no type III ME sets when $n = 12$, a rather prominent case for West-European music at least. This type exists though: for instance, when $n = 18$, consider $\mathrm{ME}_{(18,8)} = \{0,2,4,6,9,11,13,15\} = \{0,2,4,6\} \oplus \{0,9\} = \mathrm{ME}_{(9,4)}$ redoubled, shown in Fig. 4.9. Incidentally, its DFT can be computed easily from this decomposition, since the DFT of $\{0,2,4,6\}$ in $\mathbb{Z}_9$ is (in magnitudes)

[11] The fifth coefficient is also nil, since this balanced scale type is invariant by affine transformations: $5 \times \{0,1,2,5,6,8,9\} = \{0,1,2,5,6,8,9\} + 4 \mod 12$.

$$(4,0.53,0.65,1.,2.88,2.88,1,0.65,0.53)$$

and it only remains to intersperse zeroes and multiply by 2 to retrieve the DFT of $\mathrm{ME}_{(18,8)}$ already given above.

Of course type III is impossible when n is prime, since in this case only type I happens (barring the full aggregate or the empty set). But for large composite n, type III is always possible:

Theorem 4.19 (Amiot, 2005).
For composite $n > 12$, there exists d such that $\mathrm{ME}_{(n,d)}$ has type III.

The proof hinges on a technical

Lemma 4.20. *For composite $n > 12$, there exists $k \mid n$ and a prime number $p < k-1$ such that p is not a divisor of k.*

Proof. Notice that for $n = 12$ the lemma fails, since at most $k = 6$ and all prime numbers $p < 5$ divide 6.

Consider a composite $n \geqslant 25$ – lower values are checked by hand or computer. The general idea is to have k be the largest strict divisor of n. It can be written either $k = 2m+1$ or $k = 2m+2$. Since n/k is a smaller divisor of n, $k \geqslant n/k$, i.e. $k \geqslant \sqrt{n}$, hence $k \geqslant 5$ and $m \geqslant 2$.

- First case: $n = 2^r$. Let $k = n/2, p = 3$. Works whenever $n \geqslant 8$.
- Second case: n admits an odd divisor $k \geqslant 5$, not necessarily prime. Select this value for k, and let $p = 2$. This works for $n = 10, 14, 15\cdots$.
- Last case: $n = 2^a 3^b, a \geqslant 1, b \geqslant 1$. This is the trickier case, since it is for $n = 2 \times 2 \times 3$ that the lemma fails. It is not really difficult though, since whenever $n \geqslant 24$, setting $k = n/2$ and $p = 5$ satisfy the lemma conditions.

The theorem follows now from the construction

$$j \mapsto \lfloor \frac{kj}{p} \rfloor = \lfloor \frac{jn}{np/k} \rfloor$$

yielding a type III ME set, concatenated from $\mathrm{ME}_{k,p}$ which is a type I in $\mathbb{Z}_k$ since p does not divide k.

4.2.3 DFT definition of ME sets

This definition is our principal aim in this section: Quinn discovered that ME sets can be characterised by a high value of some Fourier coefficient. To quote [72]:

> We note that generic prototypicality may be interpreted as maximal imbalance on the associated Fourier balance – at least to the extent that a generic prototype tips its associated Fourier balance more than any other chord of the same cardinality possibly can.[12]

[12] Quinn was originally interested in what he calls ‘prototypical chords’, defined by cultural consensus, and which happen to be ME sets.

More precisely, as proved rigorously in [10] with excruciating detail, one can adopt the following definition as equivalent to the other ones (say Def. 4.10):

Definition 4.21. *The pc-set $A \subset \mathbb{Z}_n$, with cardinality d, is a ME set if the number $|\mathscr{F}_A(d)|$ is maximal among the values $|\mathscr{F}_X(d)|$ for all pc-sets X with cardinality d:*

$$\forall X \subset \mathbb{Z}_n, \quad \#X = d \quad \Rightarrow \quad |\mathscr{F}_A(d)| \geqslant |\mathscr{F}_X(d)|.$$

From the formulas already derived for DFT, it follows without further ado

Proposition 4.22. *Transposition, inversion* and complementation *of a ME set still yield a ME set: any pc-set homometric to a ME set is a ME set.*

This is obvious since all these operations preserve the magnitude of Fourier coefficients, which is a definite advantage over alternative definitions. It also hints that the magnitude of Fourier coefficients might be a perceptible quality – at least it is one commonly recognised.

We will show that the DFT definition is equivalent to the definition pinpointing a generated scale, in the spirit of Rem. 4.17. Reduction to the J-function definition has been carried in [10] and would be redundant here, since the equivalence of all previously known definitions had been already proved in seminal works on ME sets.

Proof. Quinn provided a simple argument which is fairly convincing for the type I case when $\gcd(d,n) = 1$, and even more in the degenerate case – but insufficient for the remaining cases. Remember

$$\mathscr{F}_A(d) = \sum_{k \in A} \mathrm{e}^{-2\mathrm{i}dk\pi/n} = \mathscr{F}_{dA}(1)$$

where dA may be a multiset.

When $d \mid n$, one easily gets $\mathscr{F}_A(d) = d$ for $A = \{0, n/d, 2n/d \dots\}$, a regular subdivision of $\mathbb{Z}_n$. Conversely, one has $|\mathscr{F}_A(d)| \leqslant 1 + 1 + \dots 1 = d$ by triangular inequality, and the equality (for a Euclidean norm) may only happen when the complex exponentials involved all point to the same direction, since $|z + z'| < |z| + |z'|$ for non-colinear z, z'. But this happens if and only if

$$\forall k, k' \in A \quad 2dk\pi/n = 2dk'\pi/n \mod 2\pi \iff k = k' \mod (n/d);$$

hence (since $\#A = d$) A is a whole arithmetic sequence with common difference n/d.[13] In this case, $A' = dA$ is a multiset with exactly one element repeated d times.

When $\gcd(d,n) = 1$, multiplication by d is bijective and $A' = dA$ is a genuine set with the same cardinality as A. All the exponentials must then be distinct, so the argumentation above does not work. Quinn argues that these exponentials should be as close as possible one to another[14], meaning that A' is a chromatic cluster $\{1, 2 \dots d\}$. This can (and should) be formally proved using

[13] The same argument proves that for $\#A < d$, $|\mathscr{F}_A(d)|$ will be maximal when A is a subset of such a sequence, see Section 4.3.

[14] 'The best the chord can do is to have pcs gathered in adjacent pans, so that the arrows point in approximately the same direction', ibid.

Lemma 4.23 (Huddling together).

Have d points $a_1,\dots a_d$ *on the unit circle* S^1*, and move* a_1 towards *the sum* $s = \sum_{k=1}^{d} a_k$*, meaning* a_1 *is replaced by* a'_1 *whose argument (or phase) is between the phases of* a_1 *and s.*

Then $|a'_1 + a_2 + \dots a_d| \geqslant |a_1 + a_2 + \dots a_d|$.

Proof. In a nutshell, the sum increases because the angle between $s = \sum_{k=1}^{d} a_k$ and $a'_1 - a_1$ is acute. Let us provide a comprehensive computation: up to rotation and symmetry, one can assume without loss of generality that $\arg(s) = 0$ and $\varphi_1 = \arg(a_1) \in]0,\pi]$; then $\varphi'_1 = \arg(a'_1) \in [0, \arg(a_1)] \subset [0,\pi]$ so a_1 and a'_1 are both 'above', see Fig. 4.10.

Since cos is decreasing on $[0,\pi]$ we have

$$\cos\varphi_1 + \cos\varphi_2 + \dots \cos\varphi_n \geqslant \cos\varphi'_1 + \cos\varphi_2 + \dots \cos\varphi_n.$$

These sums are the projections of s and $s' = a'_1 + a_2 + \dots a_d$ on the real axis. But s is assumed to be real, and $|s'|$ is greater than its projection. Hence $|s'| \geqslant s$ and more precisely $|s'| > s$ unless $\varphi_1 = \varphi'_1$.

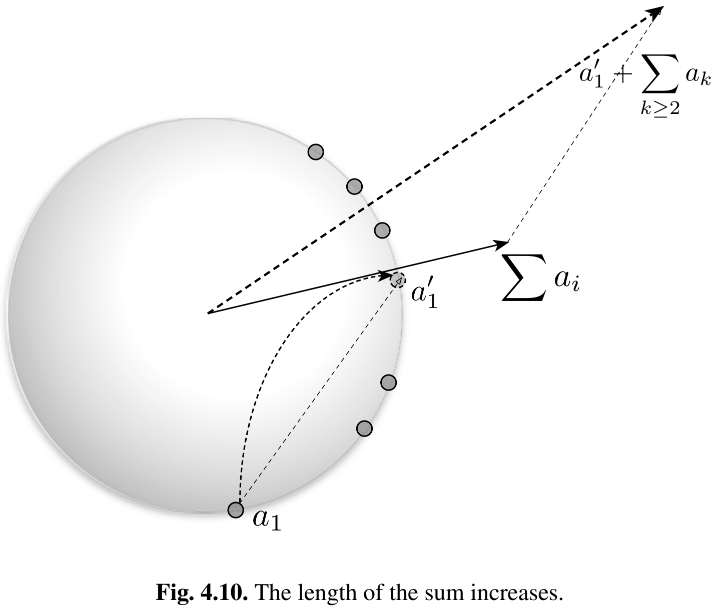

Fig. 4.10. The length of the sum increases.

The fact that A' must be a chromatic cluster follows: else, A' would feature holes in the sequence between its elements[15], and one extremal point could be moved to

[15] Writing A' in a 'basic form' such as $A' = \{0, \alpha, \beta \dots \omega\}$ with $0 < \alpha < \beta < \dots < \omega < n$ and $n - \omega$ maximal, for instance.

one such hole, increasing $|\mathscr{F}_{A'}(1)|$ in the process. This can be iterated until we get a chromatic cluster and no more.

Since $dA = A' = \{1, 2, \dots d\}$ or some translate thereof, we find $A = fA' = \{f, 2f, \dots df\}$ where f is the multiplicative inverse of d in $\mathbb{Z}_n$. In the seminal example, the diatonic collection with 7 elements is generated by fifths since $7^{-1} = 7 \mod 12$. The previous discussion on the number of generators of a generated sequence modulo n shows that in this case there are only the two generators f and $-f$.

The remaining case $\gcd(d,n) > 1$, with d not a divisor of n, is slightly more complicated. Let $m = \gcd(d,n), n' = n/m, d' = d/m$: then n' and d' are coprime and we aim at reducing the study to the preceding case. *For instance, consider the case of* $A = \{0,1,3,4,6,7,9,10\}$ *(the octatonic collection) with* $d = 8, m = 4, n' = 3, d' = 2$. Indeed $\pi_d : x \mapsto dx$ now maps $\mathbb{Z}_n$ to $\mathbb{Z}_{n'}$, each fiber (pre-image) having m elements. Assume $|\mathscr{F}_A(d)|$ is maximal and let $A' = \pi_d(A)$ (here we consider A' as a set, not a multiset. See [10] for a proof in the context of multisets). Then

$$\mathscr{F}_A(d) = \sum_{k \in A} \mathrm{e}^{-2\mathrm{i}kd\pi/n} = \sum_{k' \in A'} m(k')\mathrm{e}^{-2\mathrm{i}k'd'\pi/n'} = \sum_{k'' \in A'' = d'A'} m(k')\mathrm{e}^{-2\mathrm{i}k''\pi/n'}$$

where $m(k') = \#(\pi_d^{-1}(\{k'\}))$ denotes the cardinal of the fiber, i.e. the number of times k' is hit as an image of an element of A. Lemma 4.23 can be used here since it does not assume the points to be distinct. We can huddle the elements of $A'' = d'A'$ up to m times each, since $m(k') \leqslant m$. Hence in the maximal case, A' has d' elements, each fiber contains m antecedents, i.e. A is periodic since for any $a \in A$ we must have all the l different $a + k\frac{n}{m} \in A$ *(for the octatonic example,* A'' *is* $\{0,4\} \subset 4\mathbb{Z}_{12} = \{0,4,8\}$ *with each element repeated four times)*; hence

$$|\mathscr{F}_A(d)| \leqslant m|\mathscr{F}_{A'}(d')| \leqslant m \max_{B \subset \mathbb{Z}_{n'}, \#B = d'} |\mathscr{F}_B(d')|.$$

For the maximal value to be reached, A' must be maximally even (i.e. the elements of A'' form a chromatic cluster) and each fiber must be full (i.e. each $m(k')$ is equal to m, meaning A is the whole of $\pi_d^{-1}(A')$). This means

Proposition 4.24. *In the case* $m = \gcd(d,n) > 1$*, d not a divisor of n, a set* $A \subset \mathbb{Z}_n, \#A = d$ *is maximally even iff* $A' = dA$ *is maximally even in* $\mathbb{Z}_{n/m}$ *and A is m-periodic. In other words, A must be concatenated from A'.*

In the example proposed, $A' = \mathrm{ME}_{3,2}$ *– for instance* $A' = \{0,1\} \in \mathbb{Z}_3$ *– and hence* $A = \pi_4^{-1}(A') = \{0,1,3,4,6,7,9,10\} = A' \oplus 3\mathbb{Z}_{12}$ *with a slight abuse of notation.*

This description of the last case exemplifies the transfer of the DFT from A to its projection on an appropriate subgroup of $\mathbb{Z}_n$, cf. Proposition 3.36 above. It is illuminating to compare the DFTs of A and A' in Fig. 4.11, where a simple scale change allows us to superimpose both graphics.

To sum it up, the Fourier definition of ME sets pinpoints the quality of being as close as possible to a regular subdivision of the circle – etymologically, a cyclotomy.

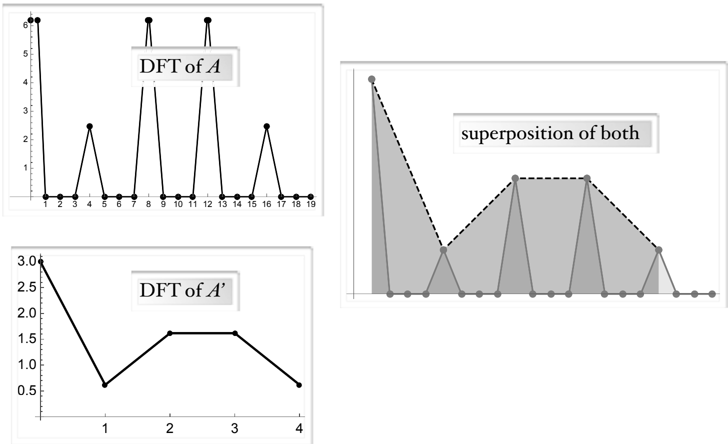

Fig. 4.11. Fourier magnitudes of a periodic ME set and its type I projection

4.3 Pc-sets with large Fourier coefficients

4.3.1 Maximal values

We have just seen that $|\mathscr{F}_A(d)|$ is maximal for $\mathrm{ME}_{n,d}$, among all d-subsets. One may well ask what are the maximal cases for other coefficients. For instance, when one keeps the cardinality d fixed, the pc-sets which maximise $|a_1|$ are the chromatic clusters, e.g. $\{0,1,2\ldots d-1\}$ as we have established during the proof of the type I-ME set case.

An extension of this result yields the maximum case for $|\mathscr{F}_A(k)|$ when k is co-prime with n: in this case kA is a set, not a multiset and

$$\mathscr{F}_A(k) = \mathscr{F}_{kA}(1)$$

is maximal when kA is a chromatic cluster, meaning that A is generated, with generating interval k^{-1}, the inverse of k in $\mathbb{Z}_n$. As a corollary, all maximum values of $\mathscr{F}_A(k)$ are identical for $k \in \mathbb{Z}_n^*$. For fixed d, this maximum is $\frac{\sin(d\pi/n)}{\sin(\pi/n)}$ (which gets close to d when n is large).

Remember that k, k' are *associated* if there exists $\lambda \in \mathbb{Z}_n^*$ such that $k' = \lambda k$. Then we can generalise slightly the above computation:

Proposition 4.25. *The maximum of $|\mathscr{F}_A(k)|$ on d-subsets is the same as the maximum of $|\mathscr{F}_A(k')|$ for k' associated to k.*

However, it is a completely different case when $\gcd(k,n) > 1$, because kA can then be a multiset, not a set, as we have seen for type II and III ME sets. It may even be possible to reach $|a_k| = d$, for type II ME sets or their subsets. This happens whenever $d \leqslant \frac{n}{\gcd(k,n)}$.

Example 4.26. Any subset of a whole-tone scale has maximum $\mathscr{F}_A(6)$: for instance for A =CDF♯G♯ $= \{0,2,6,8\}$, $\mathscr{F}_A(6) = 4 = \#A$, cf. Fig. 8.11.

The most complicated cases are reminiscent of the study of saturation in one interval: sometimes d is larger than all strict divisors of n. Of course, if $d > n/2$ we already know that the Fourier coefficients are the same as those of the complement subset, so let us assume $d < n/2$ (the case $d = n/2$ yields a maximum $\mathscr{F}_A(d) = d$ for $A = 2\mathbb{Z}_n$). Following the general idea of the proof of the DFT definition of ME sets, we want the multiset kA to be as huddled as possible: if repetition of a single value is not available, then we aim for repeating several huddled values. This happens when kA is a repetition of a subset of a regular polygon, with the eventual added points all situated on the same location, see Fig. 4.12.

Example 4.27. Consider $n = 75, d = 27 > n/3$. We can construct a perfect ME set with 25 elements, $A = \{0,3,6\ldots72\}$. Then for $k = 3$ one gets $A'_{mult} = 25A = \{0^{\#25}\}$, i.e. 0 repeated 25 times. Since there is no way[16] to enlarge A without adding new elements to A'_{mult}, the best one can do is to have these extraneous elements in A'_{mult} stay as close as possible to 0. For instance, one can add 4 and 31 to A, which turns A'_{mult} to $\{0^{\#25}, 25^{\#2}\}$, i.e. 0 25 times and 25 twice. The resulting set yields the maximum possible value of $|\mathscr{F}_A(25)|$ for 27-subsets of $\mathbb{Z}_{75}$.

It is not clear that this value is the greatest possible of $|\mathscr{F}_A(k)|$ for 27-subsets and any k. Indeed one has to check for other divisors of 75. In Fig. 4.12, I tried also B, saturated in interval 5, made of a 15-polygon and another, incomplete one as close as possible; and C, saturated in interval 15, union of five pentagons and two points on a sixth; and checked the values of the corresponding Fourier coefficients. In this case, $|\mathscr{F}_A(25)| = 24.062, |\mathscr{F}_B(15)| = 22.506$ and $|\mathscr{F}_C(15)| = 21.206$; hence A achieves the highest possible maximal value of a Fourier coefficient among all 27-subsets of $\mathbb{Z}_{75}$. For the record, $\mathscr{F}_M(27) = 21.658$ for $M = \text{ME}_{75,27}$, i.e. the ME set only beats C.

The general question now arises: for a given pair (n,d), what are the subsets $A \subset \mathbb{Z}_n$ with cardinality d that yield the maximal value of their largest $|\mathscr{F}_A(k)|$? There are three cases, summed up by the following:

Theorem 4.28. *Among d-subsets of $\mathbb{Z}_n$ (with $d < n/2$), the sets with the largest Fourier coefficients are*

1. *Subsets of regular polygons (when d is smaller than some divisor of n).*
2. *Maximally even sets.*
3. *The kind of saturated/huddled subsets shown by the example above.*

Notice that even in the last case, some solutions can be generated by J functions. For instance $(0,6,12,1,7,13,2,8)$ in Fig. 4.13 is the sequence of values of $\lfloor 6.34k \rfloor$ mod 18 for $k = 0\ldots7$; indeed even the tango/habanera pattern $\{0,3,4,6\}$ can be achieved as values of $\lfloor 2+2.5k \rfloor, k \in [\![1,4]\!]$.[17]

[16] If A is a true set, not a multiset.

[17] Keep in mind however that some pc-sets cannot be generated in this way, for instance $\{0,1,4\}$ when $n \geqslant 10$.

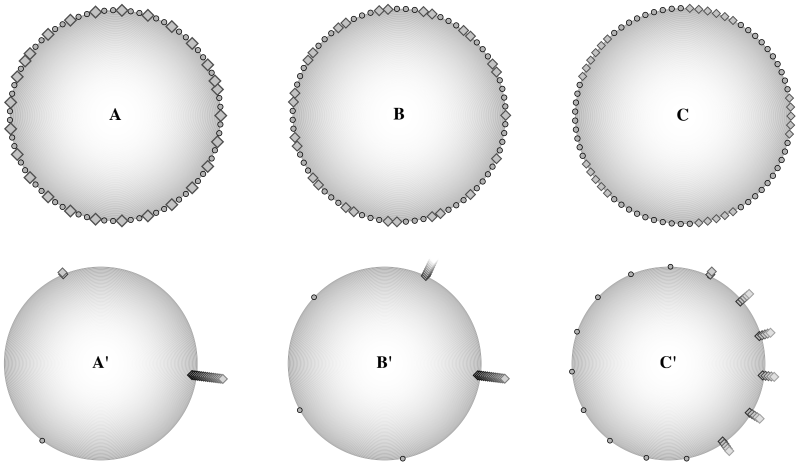

Fig. 4.12. Three *candidates* for maximum max $|\mathscr{F}_A|$ for 27-subsets of $\mathbb{Z}_{75}$

This was first analysed in the third online supplementary of [10]. The last case is somewhat messy: there is no simple formula (one has to check for k being any divisor of n, because the largest divisor does not always yield the highest Fourier coefficient) and the result is not unique up to isometry, in contrast to the ME set cases. The three different cases are exemplified in Fig. 4.13 with $n = 18$ and $d = 7, 5, 8$. The corresponding multisets are shown underneath.

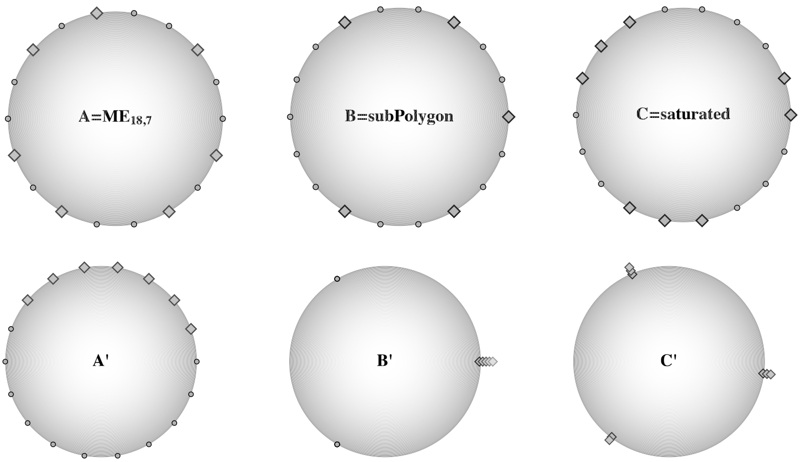

Fig. 4.13. Three *types* for maximum max $|\mathscr{F}_A|$.

When do large values occur?

All these results vindicate Quinn's notion of saliency, i.e. large a_5 show a large *fifthishness* (which we will rename *diatonicity* in the musical examples below) while large a_6 exemplifies *whole-tonedness*, etc. We have already explored the maximal cases, in the end of the discussion let us relax the condition to 'relatively large' with, of necessity, fuzzier assertions.

Example 4.29. For instance, for a short excursion in the rhythmic domain we can assert that the tresilo $(0,3,6)$ in $\mathbb{Z}_8$ has maximal 'ternariness', i.e. largest a_3 among all 3-sets ($|a_3| = 2.41$). But the standard tango pattern $(0,3,4,6)$ in the same $\mathbb{Z}_8$ has some ternary saliency too ($|a_3| = 1.85$), though its largest Fourier coefficient is the fourth ($|a_4| = 2$), asserting that tango music is binary though with a strongish ternary intent.[18] The four-note rhythm with best ternary saliency is $(0,1,3,6)$, a generated set generalizing the ME-sets construction:

$$\{0,1,3,6\} = \{0,3,6,9\} = 3 \times \{0,1,2,3\} \mod 8 \quad (|a_3| = 2.61).$$

It has been observed [25, 98] that frequent occurrences of some intervals between pc-sets (measured on a time span of one to five bars of the score, for instance) are correlated with large values of some Fourier coefficients – the fifth interval with the fifth coefficient, or minor thirds with the fourth coefficient, for instance. This is well in line with what we discussed in Section 4.1.1, and easier to adapt than the notion of maximal evenness. Is it a really reliable guideline though?

Example 4.30. Since it is a periodic ME set, $\mathscr{O} = \{0,1,3,4,6,7,9,10\}$ (the octatonic collection) has clear-cut Fourier coefficient magnitudes: $|\mathscr{F}_{\mathscr{O}}| = (8,0,0,0,4,0,0,\dots)$. The zeroes reflect the periodicity of this pc-set (the coefficients from 7^{th} to 11^{th} have been omitted since their values are reversed from the first ones).

Subsets of this collection still preserve the *saliency* of the fourth coefficient: for $A = \{0,1,3,4,7,9\}$, one finds $(6,1,1,2,\mathbf{3},1,2,\dots)$ and for $A' = \{0,1,3,4,6,7\}$, $|\mathscr{F}_{A'}| = (6,1.93,1.73,1.41,\mathbf{3},0.52,0,\dots)$.[19]

The last two examples both display four minor thirds, and though the fourth Fourier coefficient has the same magnitude, the other coefficients do not. The more we stir away from the regular subsets studied before, the less exact the correlation between saturation and saliency becomes, cf. 5.4 below.

For generated sequences whose generator is not a divisor of n, or bouts of such sequences which are not ME sets, first remember that a generated sequence features more occurrences of the generating interval than several juxtaposed partial sequences: there are six second intervals in a whole-tone scale $WT = \{0,2,4,6,8,10\}$, but only four in the Guidonian hexachord $GH = \{0,2,4,\mathbf{5},\mathbf{7},\mathbf{9}\}$ which is a reunion

[18] Indeed a kind of walz, *El vals criollo* is among the three principal styles of music played and danced in tango balls.

[19] This somewhat informal remark is very important, as it will lead us to replace advantageously the 'complex' manipulations in Forte's 'Set Theory' (i.e. subset relationships) by consideration of saliency. This is a *forte* of DFT theory, noticed by Yust.

of two three-note whole-tone sequences (i.e. a convolution product of $\{0,2,4\}$ by $\{0,5\}$, see Figs. 8.8 and 8.5 respectively). On the other hand, this last pc-set is a full-fledged fifth sequence $(5,0,7,2,9,4)$. All this appears clearly on the Fourier magnitudes, see also Fig. 8.26 and 8.23:

$$|\mathscr{F}_{WT}| = (6,0,0,0,0,0,6,\dots) \qquad |\mathscr{F}_{GH}| = (6,1.035,0,1.414,0,3.864,0,\dots).$$

Notice that the sixth coefficient, maximal for WT, altogether vanishes in GH despite the four whole tones in it[20] which shows crudely that the magnitude of a Fourier coefficient is not completely equivalent to the frequency of occurrence of a corresponding interval. However, in tonal music where a diatonic universe is often prevalent, the organisation of fifths often adheres to the generating sequence of the diatonic, which is maximal in number of fifths, and the 5^{th} Fourier coefficient is accordingly large – as we have seen in Section 4.2, the diatonic collection has maximum magnitude $(1+\sqrt{3} \approx 2.73)$ among all other seven-notes pc-sets for the fifth coefficient. Its most frequent subsets, the simple and popular boogie/rock bass sequence CFG (057) and the pentatonic collection, reach exactly the same value. In the former case (CFG) this is not far from the absolute maximum possible for the DFT of a 3-pc-set. In the latter we have the absolute maximum.

So when can we rely on the informal remark above, since it is not always true?

The Fourier transform being continuous, slight modifications of a pc-set entail slight modifications of the Fourier coefficients. Hence the somewhat vague, but informative, assertion:

Proposition 4.31. *Usually, pc-(multi)sets with a high frequency of occurrence of interval d are close to (subsets of) arithmetic sequences with generator d and yield a high value of their k^{th} Fourier coefficient, where k is*

- *n/d when d divides n, or*
- *$d^{-1} \in \mathbb{Z}_n$ when n,d are coprime.*[21]

This lacks a precise definition of ‘closeness’ to a given pc-set, a notion that is open to interpretation, and leaves aside the case of a loose relationship between d and n (neither divisor nor coprime). It is also debatable for small d and especially $d=1$, though there is some correlation in this case with the number of *successive* semitones but their overall distribution could ruin this character, see Fig. 8.28 where a scale with four semitones has $a_1 = 0$.[22]

We will discuss in Section 5.4 a relationship between size of DFT coefficient and voice-leading distance to a (usually virtual) chord with maximum value, first estimated by Tymoczko and improved for the present publication.

[20] This is because there are as many odd pcs as even. Another way to look at it is that this coefficient is nil already for the *factor* $\{0,5\}$.

[21] See Section 4.2 for an explanation of this value of k.

[22] The only such seven-note scale.

4.3.2 Musical meaning

A word of caution is in order: when considering the character of a pc-set (diatonic, whole-tonic, etc. . .) we usually compare the respective magnitudes of appropriate Fourier coefficients. But it could well be argued that these magnitudes should be weighted*: for instance, coefficient a_2 can be as large as 6 (for a whole-tone scale) but a_5 (or a_1) is never more than $\sqrt{2}+\sqrt{6}$ (Guidonian hexachord). However, these limitations fall when one drops genuine pc-sets and considers continuous DFT, even if the musical notions underlying, say, a regular division in seven of an octave, are more virtual than real. In balancing these arguments, I prudently chose not to choose and left the comparison of magnitudes of Fourier coefficients as is, though perhaps with a modicum of salt. For instance, the jingle for* la Société Nationale des Chemins de Fer *created by Michael Boumendil (which I quote because David Gilmour, Pink Floyd's lead guitarist, fell in love with it and used it as a leitmotif in his song* Rattle That Lock*: see* `https://www.youtube.com/watch?v=L1v7hXEQhsQ`*) arpeggiates a seventh chord CGA♭E♭; the corresponding profile in Fig. 8.12 shows a large a_3, i.e. 'major thirdishness' or 'augmentedness', which indeed correlates with the presence of three thirds (two major, one minor). But the value of $|a_5|$, though only 2/3 of $|a_3|$, is comparatively large because it is closer to the maximum theoretical value for a_5 (indeed, the pc-set is almost saturated in fifths), and hence the pc-set is also fairly diatonic, which is good for rock music.*

The six characters

We may as well begin with clarifying the meaning of saliency for coefficients 1,2,3,4,5,6 in $\mathbb{Z}_{12}$. I take them from the easiest to the less obvious. Examples are provided on Fig. 4.14.

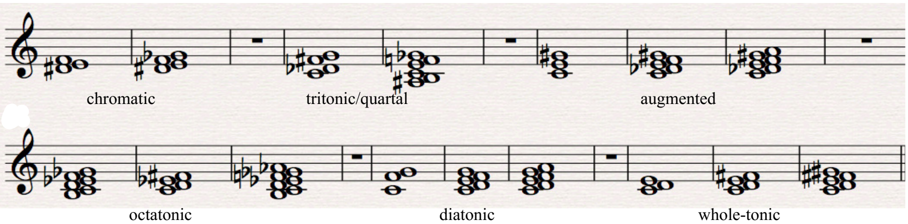

Fig. 4.14. Examples of the six characters

- The sixth is easiest to understand, especially using Quinn's (weighing) 'scales': this coefficient is greater when its pcs concentrate in one of the two whole-tone scales. It is uncontrovertibly the *whole-toneness*. Clear-cut examples can be found on Figs. 8.23, 8.15, 8.23; a more ambivalent case is the Guidonian hexachord in Fig. 8.23, a reunion of two whole-tone tetrachords CDE - FGA, but with opposite polarities resulting in zero 'whole-toneness'.

- As we have already discussed at length, the fifth coefficient can well and truly be called the *diatonicity* of a pc-set: it has everything to do with the tonal character (or alternatively the generatedness by fifths) which marks pentatonic and diatonic scales among other prominent specimens, see Fig. 8.18 and 8.29 or even 8.7 (CFG) and 8.8 (CDE).[23] Notice that the rather large index 5 discriminates dramatically between just and diminished fifths, since a tritone has nil a_5 but a fifth is the maximal dyad for this saliency, cf. Example 4.32 below.
- Third and fourth mark on the one hand generatedness (or saturation) in major and minor thirds respectively, but *minor(major)-thirdishness* is a somewhat ambiguous notion: among subsets with similar cardinality, any subset of a diminished seventh features a maximal magnitude for a_4, but so does the octatonic scale (among eight-notes pc-sets); and I can agree with J. Yust who dubs *octatonic* the pc-sets with large a_4 – they are usually subsets of some octatonic scale. As for *major* thirdishness, I like to think of it as 'augmentedness', good prototypes being the augmented triad or the 'magic' hexachord $\{0,1,4,5,8,9\}$ (also called 'ode to Napoleon'), cf. Fig. 8.24.
- From the discussion above, one could wonder whether a large a_1 corresponds to many semitones or many (major) sevenths, but the issue is not large and we will call *chromatic* any pc-set with a comparatively large a_1. However it should be noticed that too many notes will perforce diminish this coefficient. For instance, the scale B C D♭ E F G A♭ or $\{0,1,4,5,7,8,11\}$ has $a_1 = 0$ (see [67] and Fig. 8.28) though it features many semitones. Notwithstanding, decent prototypes are chromatic chunks of lengths 4 to 6, i.e. Figs. 8.14, 8.22, 8.27 with chromaticities equal respectively to $3.35, 3.73, 3.86$. These coefficients are less sensitive than a_5: a major triad is generated neither by major nor minor thirds but both coefficients a_3, a_4 (and of course a_5 too) are fairly large.
- The more troublesome coefficient is a_2. Yust uses Messiaen's Limited Transposition Mode $M_5 = \{0,1,2,6,7,8\}$ as a prototype (Fig. 8.25), together with $M_4 = \{0,1,2,3,6,7,8,9\}$ (Fig. 8.30) which sports almost the same value (and is more frequently used, if only in R. Wagner's *Tristan*, cf. [5]). I like the neologism *tritonic* to qualify pc-sets with large a_2, though Yust's *quartal quality* is convincingly expostulated in his example of Ruth Crawford Seeger's 'White Moon' [98, 100].[24] It is perhaps an artifact of working modulo 12, but as he points out, this quality quite often goes with a lack of thirds and sixths, which is a hallmark of some early 20^{th} century music: for instance, the prominence of this coefficient in B. Bartok's Fourth Quartet can be arguably correlated to its acknowledged 'modernism' [98].[25]

This classification makes it really easy to appreciate the character of any given pc-set:

[23] Actually the Guidonian hexachord does slightly better than all other pc-sets, with 3.86 instead of 3.73 for the diatonic, a minor triumph for archeo-musicologists perhaps.

[24] Sandburg Songs, n° 2.

[25] And of course, if one Fourier coefficient is large, then the others are left less room, since the sum of their squares is fixed.

Example 4.32. Consider for instance the three aggressive fifths initiating C. Debussy's *La Puerta del Vino* (Préludes, II): A♭-D♭, E-A, A-D, constituting the pc-set $\{1,2,4,8,9\}$. Its Fourier profile can be found online or computed with the software I provided with this book, or even roughly estimated: to begin with, the tritone $(2,8)$ can be cancelled out for any odd-indexed coefficient, leaving $\{1,4,9\}$ to be exponentiated and summed with diverse coefficients. Since this is very close to an equilateral triangle ($\{1,5,9\}$) the coefficient a_1 must be quite small, i.e. the pc-set is not chromatic (character 1). On the other hand, multiplying by 5 yields $\{5,8,9\}$, whereas a maximum would be reached for $\{7,8,9\}$; hence our pc-set is somewhat diatonic (bearing in mind though that only three of its five notes bear their weight on this character). Most of the bulk is carried by the quartal quality: multiplying by 2 yields the multiset $\{2,4,4,6,8\}$ (after reordering) whose vector sum has magnitude close to 3 after cancellation of the tritone $(2,8)$. The remainder is on the 'augmented' quality, i.e. a_3, which can be computed from multiset $\{3,0,3\}$.

All in all, this describes a non-chromatic, still diatonic but fairly modern pc-set, which I would say is an accurate description of a listener's intuitive perception. Check its profile in Fig. 8.19.

Examples in modal music

Fig. 4.15. *Voiles*, Preludes vol I, C. Debussy

In *Voiles*, C. Debussy opposes quite stringently two of those pure archetypes: the whole-tone scale, which is used for most of the piece, and the pentatonic (black keys) which occurs during the climax just before the last page, back in whole-tone (Fig. 4.15). A deaf scientist, riveted to the meters of Fourier coefficients during the piece, could not miss the exchange of high values between a_6 (from concentration in 6 down to 1) and a_5 (from $2+\sqrt{3} \approx 3.7$ down to 0) coefficients, even without any knowledge of scales and music theory, cf. Fig. 4.16. The Fourier profiles are provided separately in the tables, Figs. 8.23 and 8.18.

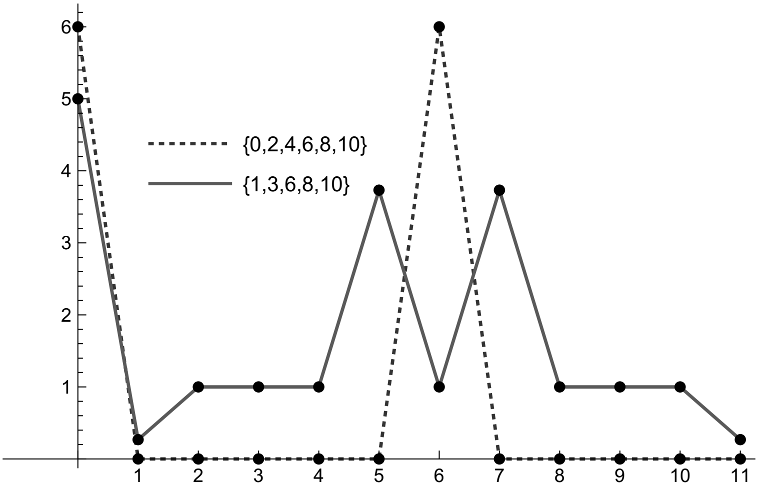

Fig. 4.16. a_5 and a_6 in Debussy's *Voiles*

This is caricatural of course, since traditional analysis of scale content gives the same result quite easily. The point of Fourier analysis of saliency is that it can help decide the character of (a passage of) a piece in less clear-cut cases. Less caricatural perhaps is the pivotal oscillation between the first and middle section, playing on the intersection of the two pc-sets $\{6,8,10\}$, i.e. G♭A♭B♭: again set-theoretic consideration provides an adequate explanation of this move (around A♭ which happens to be a common center of symmetry of both scales), but it does not hurt to recall that the pc-set $\{6,8,10\}$ is both whole-tonic and diatonic, i.e. with both a_5, a_6 large, see the central peak in Fig. 8.8.

A more striking example of the efficiency of DFT magnitude is Yust's analysis in [98] of the beginning of Bartok's Fourth String Quarter IV's movement iv, wherein the melody plays an acoustic scale opposing the accompaniment on *DE♭GA♭*, see fig. 4.17.

A much more detailed analysis is to be found in the reference given. Here we will simply observe that classical comparison of these two pc-sets is difficult, and

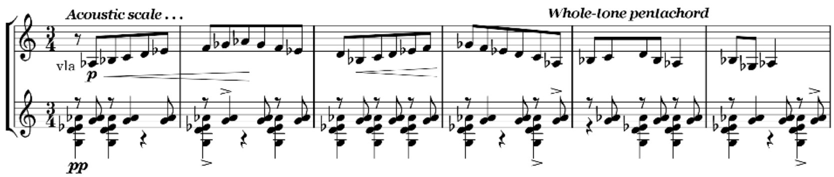

Fig. 4.17. Bartok's String Quartet 4, iv, mm 6-12

that the analyst is tempted to resort to subjective qualities of the scales involved, while confrontation of the DFT's magnitudes is illuminating, see Fig. 4.18.

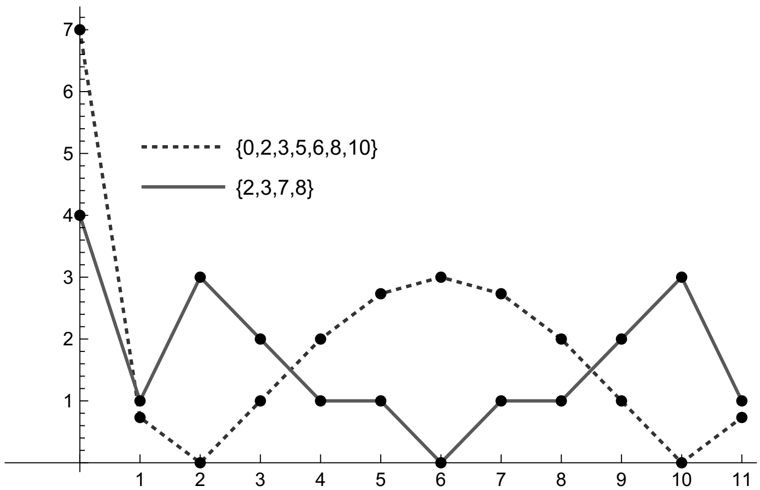

Fig. 4.18. DFT magnitudes of melody and accompaniment in Bartok's String Quartet 4, iv, mm 6-12

The second component, the quartal or tritonic quality, is nil for the acoustic scale in the melody, but large for the bass. Indeed, the latter shows a decidedly atonal quality. Conversely, the whole-toneness of the acoustic scale is large ($|a_6| = 3$) while the accompaniment's is nil, its four notes being equally distributed in the two whole-tone scales, i.e. Quinn's two 'pans'. The values of the first (chromaticity) and fifth (diatonicity), while not as contrasted as second and sixth, are also very revealing of the opposite characters of melody and accompaniment.

Magnitude of Fourier coefficients can help resolve old conflicts. In [100], Yust observes three clearly diatonic voices in Stravinsky's *Three Pieces for String Quartet*, first movement, namely GABC, C♯D♯EF♯ and CD♭E♭ (cf. Fig. 4.19 and their Fourier profiles respectively on Figs. 8.17, 8.16 and **??**), whose large coefficents a_5 more or less cancel each other out when the pc-sets are reunited, according to their

Fig. 4.19. Pc-content of three instruments in Stravinsky's *Three Pieces for String Quartet*, first movement

balanced phases around the circle[26], as formulated in Proposition 6.2; the union of these three pc-sets is CC♯D♯EF♯GAB, close to an octatonic collection (see profile in Fig. 8.32 with its spikes on a_4, even more pronounced if one unites the three voices in a multiset, not a set). The octatonic character of this movement is confirmed by the magnitude of a_4 for the second violin and cello.

Traditional set-theoretic analysis (using subset relationships or 'historical' arguments) of this passage and numerous others had so far spectacularly failed to achieve unanimity, see the fur fly between [92], [93] and more recently [91], [84] (I borrow these references from Yust). This kind of issue can now very easily be resolved, simply by measuring $|a_4|$, or looking it up on the Fourier profiles of pc-sets in Stravinsky's scores.

Tropes

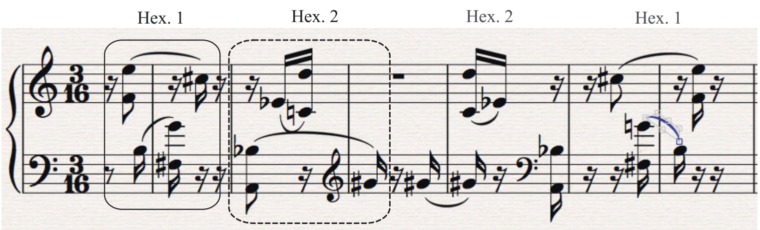

Fig. 4.20. First phrase of Webern op. 27, n° 1 (retranscribed)

[26] We will see that the phase of a_5 locates a pc-set on the circle of fifths.

A very common occurrence in dodecaphonic music is the division of the twelve tones in two hexachords, or 'tropes'. It is a golden opportunity to use Babbitt's theorem, either from the intervallic point of view, or using DFT. This is effective on almost any example, such as the first bars of Webern's first movement of Variationen op. 27 (Fig. 4.20).[27] The first hexachord is divided between the two hands[28] in two trichords, FEC♯ and BF♯G. The DFT magnitudes of both trichords and their reunion (dotted line) are shown in Fig. 4.21.

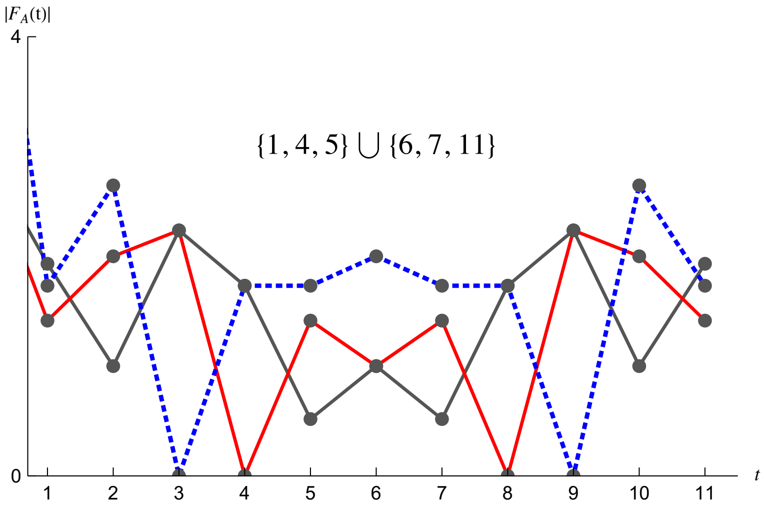

Fig. 4.21. Decomposition of a hexachord in Fourier space

The most interesting coefficient is the third, which is the highest valued for both trichords though it is nil for the hexachord. This brings us a taste of the analysis of directions, or phases, of Fourier coefficients, that will be developed in Chapter 6.

Remember that a_3 is about the 'major-thirdishness' (or 'augmentedness'): specifically, FEC♯ alias $\{1,4,5\}$ has a large $a_3 = 1-2i$ (with magnitude $\sqrt{5}$) that points towards the closest pc-*multi*set with maximal third coefficient on the continuous pitch circle, i.e. $\{0.5, 4.5, 4.5\}$, subset of an augmented triad. The other trichord has opposite $a_3 = 2i - 1$, because[29] $\{6,7,11\}$ is closest to $\{6.5, 6.5, 11.5\}$ which is in perfect opposition with the other augmented triad position. To sum it up, both trichords have

[27] Actually this part is often analysed as a superposition of the tone-row and its retrograde. At first hearing however what is perceived is what I develop here.

[28] It is known that this division was very important for the composer, who strongly opposed easier fingerings proposed by his pianist Peter Stadlen. However, no less a pianist than Glenn Gould suppressed the high-risk hand-crossing at the beginning of the second half of variation 2.

[29] One could also use a symmetry argument.

a strong *augmented flavour*, but live in opposite directions of the harmonic spectrum in that respect, so that they neutralise each other, cf. Fig. 4.22.

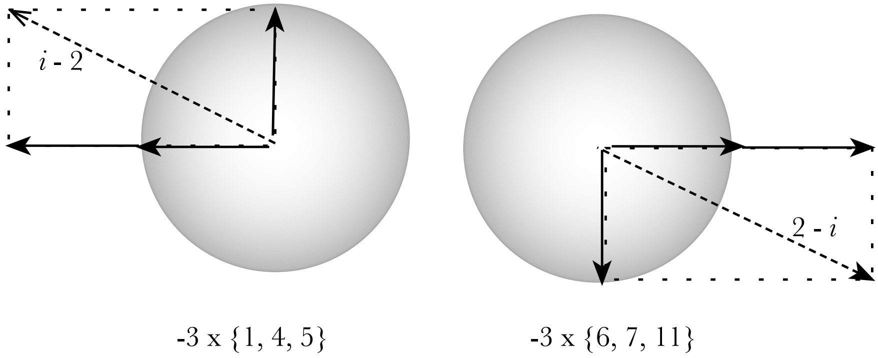

Fig. 4.22. Coefficients a_3 for both trichords are opposite

The fourth coefficient is a simpler situation: $\{6,7,11\}$ is devoid of any minor-thirdish/octatonic flavour (it touches all three diminished sevenths) and the whole minor-thirdishness of the hexachord is supported by the first trichord, $\{1,4,5\}$. For all other coefficients, the trichords more or less combine their strengths into the hexachord's. The overall picture of this hexachord is highly chromatic, and somewhat whole-tonic (high values of first and sixth).

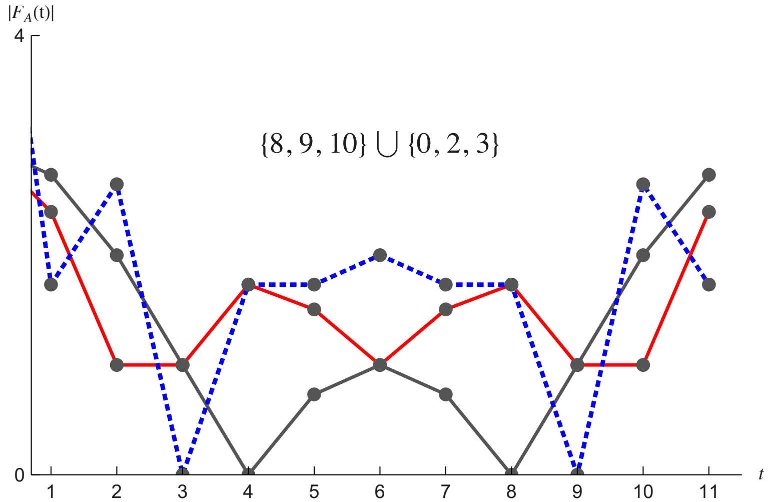

Fig. 4.23. Decomposition of the complement hexachord

It should come as no surprise that the second, complementary hexachord has the same Fourier distribution! But the decomposition into trichords introduces a slightly different fragrance: as we can see in Fig. 4.23, the (still opposite) third coefficients are much smaller ($\pm$i). Here it is a very chromatic trichord $\{8,9,10\}$ that is devoid of any specific diminished character (fourth coefficient nil), and the whole of this dimension in the hexachord is carried by the other trichord $\{0,2,3\}$.

All in all, this shows that despite the absence of isometry, the choice of contrasted constituent trichords enhances the balance between the hexachords, which goes well with Webern's use of symmetries in all three variations.

Another famous and much analyzed dodecaphonic example is the initial tone-row of Berg's *Lyrische Suite* op. 28 (Fig. 4.24). The following analysis adds a new perspective to the traditional analysis of hidden (fourth/fifth) cycles, like [71].

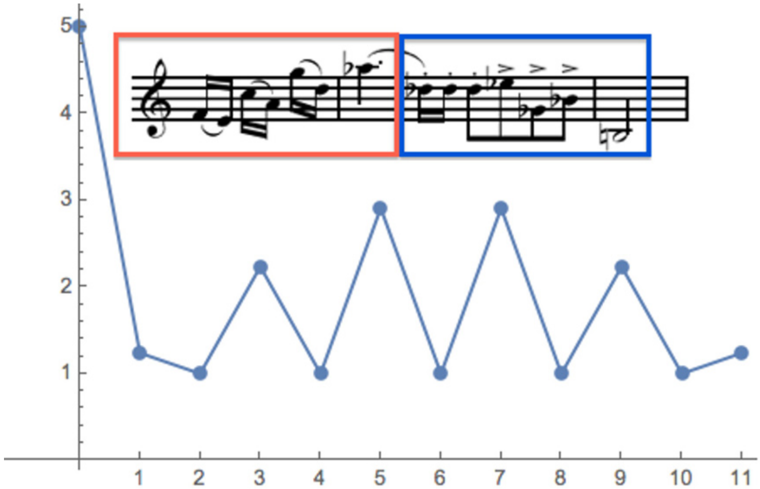

Fig. 4.24. DFT profile of the initial seven notes (or last five).

This time, there are conjoint rhythmic, melodic and dynamic reasons to segment this series into 7+5 notes, cutting between the high A♭ and the sequel. Of course, the DFT of FECAGDA♭ is identical (except the 0^{th} coefficient) with the DFT of the remainder D♭E♭G♭B♭B, by Babbitt's *generalised* theorem. However the shape of Fig. 4.24 deserves commentary.

The large fifth (or seventh) coefficient is known as an indicator of diatonicity. Indeed, both parts are close to diatonic and pentatonic respectively (rotating the last B to the beginning of the series would allow a perfect decomposition of this sort). If we remove the A♭ (equivalently, we may decide to segment the phrase *before* the A♭, the two hexachords are isometric a tritone away), it yields the Guidonian hexachord, which has a neat DFT profile (see Fig. 8.26) with maximal diatonicity (it is saturated in fifths), and nil even coefficients, enhancing the contrast with the fifth content –

which is rhythmically obvious when one picks every other note, getting fifths FC, EA, CG, AD...

Another well-known feature of this tone-row which deserves further comment is the structure of *consecutive* intervals: though we know (from Babbitt's theorem again) that the overall intervallic structures of both parts are equal (up to a constant since their cardinalities differ here), the composer manages to pick up different consecutive intervals. Specifically, if interval $\delta = b - a$ appears between two consecutive notes in the first seven, the opposite interval appears in the sequel (e.g. FE vs. B♭B). This is a delicate construct to achieve by hand, and I leave it to the reader to construct the 48 'all-interval series' beginning with the seven white keys in some order (the last being precisely Berg's tone-row, up to a cyclic permutation).

I think that it is not so far-fetched to infer from this example one reason why Berg seems more amenable to untrained ears (in 20^{th} century music): even in dodecaphonic music, he manages to keep a significant diatonic character. This idea is not original, but it can now be checked scientifically by using DFT. For instance, segmenting his sonata op. 1 every two seconds, the value of $|a_5|$ on each segment averages 1.57, a significantly large value. This should be researched more intensively of course[30], studying motives and especially hexachords throughout his work vs. Webern's and Schönberg's. I will venture just another (well-known) example of clear diatonicity in Berg, the initial and last bars of his Violin Concerto arpeggiating fifth cycles (as a four-note cycle and then a diatonic F major scale), and the main tone-row featuring remarkably diatonic hexachords, see Fig. 4.25.

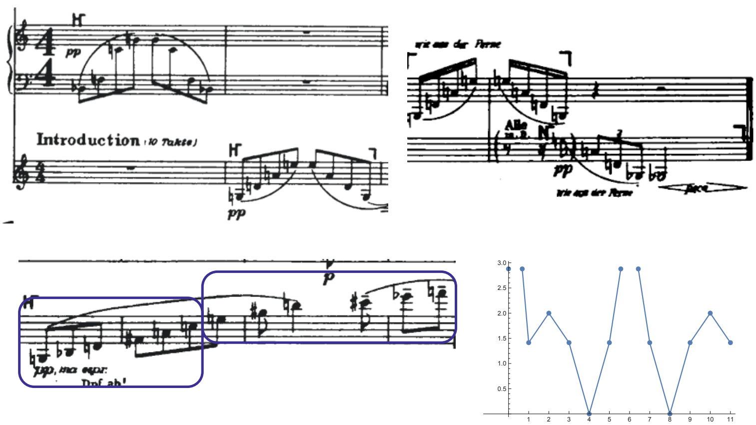

Fig. 4.25. First bars and tone-row in Berg's Violin Concerto, with its hexachords' clearly diatonic Fourier profile

[30] I carried out a cursory analysis of Berg's *Vier Stücke* op. 3, an 'intermediate' piece, atonal but not yet dodecaphonic; interestingly, it seems to exhibit much lower values of a_5.

In conclusion, DFT now provides precise, objective, quantitative measurements of diatonicity (or octatonicity, or whole-toneness, etc.) for almost any given piece of music.

4.3.3 Flat distributions

FLIDs

In his talk at the first MCM convention (Berlin 2007), Canadian theorist Jonathan Wild introduced FLIDs – Flat Interval Distribution Sets. The idea was a generalisation of the famous case of 'all-interval pc-set', e.g. for $A = \{0,1,3,7\} \subset \mathbb{Z}_{12}$, every interval occurs exactly once (except the tritone, because $7-1$ is the same as $1-7=6$ modulo 12).

Definition 4.33. *$A \subset \mathbb{Z}_n$ is a FLID if $\mathrm{IC}(A)(k)$ is constant for $k = 1,2\dots n/2$.*

Wild allowed the tritone interval $n/2$ when n is even, or else there are no possible FLIDs since a tritone must occur twice or not at all ($n/2 = -n/2$)[31]; we cannot take this view here because of Theorem 4.36 below and hence restrict FLIDs to odd values of n. One index which cannot be taken into account is 0, because $\mathrm{IC}(A)(0)$ is always the cardinality of A, larger than all other possible values of $\mathrm{IC}(k)$.

Actually the notion has been well studied in combinatorics under the name of 'difference sets'. There is a nice relationship with block designs[32]: if $D = \{d_1, \dots d_k\} \subset \mathbb{Z}_n$ is such that any $b \neq 0$ in $\mathbb{Z}_n$ can be expressed in λ different ways as $d_i - d_j$, then the $D + \tau, \tau \in \mathbb{Z}_n$, form a (n,k,λ) block design.

Example 4.34. Let $n = 11$ and consider the *quadratic residues*, i.e. all squares reduced modulo n (for instance $5^2 \equiv 3 \mod 11$). Their set, $D = \{0,1,3,4,5,9\}$ is a 3-FLID: in $D - D$ all possible values (except 0) occur thrice, see Fig. 4.26 (this construction, known as *Hadamard difference sets*, works for prime powers $n \equiv 3$ mod 4). The associated block design is $(11,6,3)$: any pair of translates of D, e.g. D and $D+3 = \{\mathbf{1},\mathbf{3},\mathbf{4},6,7,8\}$, intersects in exactly three points.

The last example is invariant under multiplication (squares of multiples are squares). More generally, since affine maps *permute* the values of the IC[33], we can state that

Proposition 4.35. *Any affine transform of a FLID $A \subset \mathbb{Z}_n$ (i.e. any $aA + b$ for a coprime with n) is also a FLID.*

---

[31] If the tritone is counted only once, then $\{0,1,3,7\}$ mod 12 or $\{0,2,3,5\}$ mod 6 (i.e. the French augmented sixth $A\flat CDF\sharp$ as a subset of a whole-tone scale) are FLIDs. A variant of Theorem 4.36 below could be established for this generalised definition, with the DFT's magnitude oscillating between two close values.

[32] A block design (n,k,s) is a collection of k-subsets of a n-set such that any pair of subsets shares s elements. When $s = 1, A$ is called a *projective plane*, like the famous Fano plane which is the reunion of the seven 'lines' $\{0,1,3\} + \tau$ in $\mathbb{Z}_7$ which intersect one another in one point exactly.

[33] Under the bijection $x \mapsto ax + b$, any interval δ is mapped to $a\delta$.

-	0	1	3	4	5	9
0	0	10	8	7	6	2
1	1	0	9	8	7	3
3	3	2	0	10	9	5
4	4	3	1	0	10	6
5	5	4	2	1	0	7
9	9	8	6	5	4	0

Fig. 4.26. Differences mod 11 of $D = \{0,1,3,4,5,9\}$.

Hence two such sets are usually considered equivalent if one is the affine image of the other.

Since affine maps also permute Fourier coefficients, this yields a neat proof of the easy implication of the following theorem, which links intervals and Fourier coefficient distributions:

Theorem 4.36. *A is a FLID iff its Fourier transform is flat. More precisely,*

$$\mathrm{IC}(A) = (d,m,m,m\dots) \iff |\mathscr{F}_A|^2 = \big(d+(n-1)m, d-m, d-m, \dots d-m\big).$$

Remark 4.37. By a continuity argument, this means that the dispersion of values of the DFT (the 0^{th} coefficient excepted) is correlated to the dispersion of the intervallic distribution: both are nil for FLIDs. We have studied the opposite case before: maximum values for one Fourier coefficient coincide with maximum occurrences for a given interval. Explicit but messy formulas for these dispersions can be computed.

Proof. The direct implication is straightforward, since $\widehat{\mathrm{IC}(A)} = |\mathscr{F}_A|^2$: for $\mathrm{IC}(A) = (d,m,m,m\dots)$ one computes its Fourier transform,

$$|\mathscr{F}_A|^2(k) = d + \sum_{t=1}^{n-1} m\mathrm{e}^{-2\mathrm{i}\pi kt/n} = (d-m) + \sum_{t=0}^{n-1} m\mathrm{e}^{-2\mathrm{i}\pi kt/n} = d-m \text{ for } k \geqslant 1.$$

The value in 0 is the ‘cardinality’ of $\mathrm{IC}(A)$, i.e. the sum of its elements $d+(n-1)m$.

The reverse implication is trickier. My original proof in [13] uses the algebra of circulating matrixes isomorphic with Fourier space. Here is a shorter one with DFT only, but it is not constructive.

Assume that $|\mathscr{F}_A|^2$ is flat, i.e. $|\mathscr{F}_A|^2 = (k,\ell,\ell,\ell\dots)$ for some $k,\ell \in \mathbb{R}_+$. Define $d,m \in \mathbb{R}$ such that $d-m = \ell, d+(n-1)m = k$; then by the direct computation, the Fourier transform of the distribution $f = (d,m,\dots m)$ is $|\mathscr{F}_A|^2$. Since DFT is bijective, and $\widehat{\mathrm{IC}(A)} = |\mathscr{F}_A|^2 = \widehat{f}$, we have $\mathrm{IC}(A) = f = (d, m, \dots m)$, i.e. $\mathrm{IC}(A)$ is flat.

Large determinants

An equivalent characterisation stems from the following remark. The determinant of the circulating matrix associated with A (see Section 1.2.3) is simply the product of its Fourier coefficients: $\det(\mathscr{A}) = \prod_k \mathscr{F}_A(k)$. Consider $|\det(\mathscr{A})|^2$, which is the product of the Fourier coefficients of $\mathrm{IC}(A)$. From Parseval-Plancherel's identity (Theorem 1.8),

$$\sum_{k \in \mathbb{Z}_n} |\mathscr{F}_A(k)|^2 = n\,d \text{ where } d = \#A.$$

As we have stated again, $\mathscr{F}_A(0) = \#A$ cannot vary. But in order to maximise the product of the other Fourier coefficients $\prod_{k=1}^{n-1} |\mathscr{F}_A(k)|^2$ under the condition $\sum_{k=1}^{n-1} |\mathscr{F}_A(k)|^2 = (n-d)d$, one must have them all equal.[34] Hence

Proposition 4.38. *Among all d-subsets $A \in \mathbb{Z}_n$, the maximal possible value of $|\det(\mathscr{A})|$ is reached when A is a FLID.*

Geometrically, this means that the columns of $\mathscr{A}$ are the least colinear as possible, i.e. that the translates $A, A+1, A+2\dots$ are as much apart (in $\mathbb{R}^n$) as possible i.e. that their mutual angles are as close to a square angle as possible.

FLIDs do not exist for any pair (n,d)[35], but this yields an explicit universal majoration:

$$\#A = d \Rightarrow |\det(\mathscr{A})| \leqslant d\left(d\,\frac{n-d}{n-1}\right)^{\frac{n-1}{2}}.$$

For instance, for 4-subsets of $\mathbb{Z}_{12}$, the maximum determinant is reached for the all-intervals tetrachords $\{0,1,3,7\}$ (or $\{0,1,4,6\}$) and is equal to 1,024, though the formula's upper bound yields about 1,421; there are no genuine FLIDs in $\mathbb{Z}_{12}$ because of the tritone doubling.

Perhaps this notion of the size of the determinant should warrant additional research. Obviously it is

- nil for subsets which tile;
- small for subsets with irregular interval distribution, like ME sets;
- and maximal for FLIDs.

FLIDs which tile

When the multiplicity m of all intervals in a FLID A is equal to 1, we reach a very interesting situation, because A tiles almost all of $\mathbb{Z}_n$:

[34] This is well known and can be proved for instance with convexity arguments.

[35] At least if one insists on actual pc-sets, i.e. distributions with values in 0-1.

Definition 4.39. *A* Golomb ruler *is a set A such that all difference values occur exactly once:*

$$a_i - a_j = a_k - a_l \iff (i, j) = (k, l)$$

It is perfect *if all possible values (except 0) are obtained once, i.e.*

$$\mathrm{IC}(A) = (d, 1, 1, 1 \dots).$$

A Sidon set *is a set A such that all sum values occur exactly once:*

$$a_i + a_j = a_k + a_l \iff (i, j) = (k, l)$$

It is complete *if all possible values (except 0) are obtained once, i.e. A tiles a subset of $\mathbb{Z}_n$.*

Hence a perfect Golomb ruler in $\mathbb{Z}_n$ is a 1-FLID.

Proposition 4.40. *Sidon sets = Golomb rulers.*

Proof. $a_i - a_j = a_k - a_l$ has a unique solution $\iff$ $a_i + a_l = a_k + a_j$ has a unique solution.

This trivial proposition yields a very nice link between intervallic studies and tilings; unfortunately there is no way these sets can provide true tilings of the whole of $\mathbb{Z}_n$. For instance $\{0, 1, 3\}$ only tiles $\{0, 1, 2, 3, 4, 6\}$ in $\mathbb{Z}_7$. Even almost FLIDs like $\{0, 1, 4, 6\}$ in $\mathbb{Z}_{12}$ cannot tile without overlapping[36] since

Proposition 4.41. *An all-interval set intersects any of its translates. The cardinality of the intersection $A \bigcap (A + t)$ is* $\mathrm{IC}(A)(t)$.

Though difference sets are mostly studied in $\mathbb{Z}$ (or even larger structures) they deserve a mention in this book. [37] For one thing, Sidon originally created the eponym sets during his investigation of Fourier series.[38] Some very specific constructions are known which yield spectacular results.

For instance, in [80] Singer inadvertently constructed a superb 1-FLID[39], alias Sidon set:

Theorem 4.42. *For any prime p there exists a subset A of $\mathbb{Z}_n$ with $p + 1$ elements, where $n = p^2 + p + 1$, such that the intervallic distribution is uniform:* $\mathrm{IC}(A)(k) = 1$ *for all k (except $k = 0$ of course).*

[36] Composer Tom Johnson has practiced with graphs between pc-sets with the relationship ‘not intersect’, see for instance [52].

[37] See also the notion of spectral set which can be expressed in terms of differences, cf. Proposition 3.58.

[38] Sidon sets are still instrumental in the study of lacunar and/or random Fourier series in Harmonic Analysis.

[39] The construction also yields $\left(\frac{p^{n+2} - 1}{p - 1}, \frac{p^{n+2} - 1}{p - 1}, \frac{p^{n+2} - 1}{p - 1}\right)$ difference sets, cf. [29].

The construction is non-trivial, making use of cubic extensions of finite fields (which appear to crop up often, quite unexpectedly, in tiling theory, see [6] for instance). Examples are $\{0,1,4,6\}$ for $p = 3, n = 13$ or $\{0,1,4,6,13,21\}$ for $p = 5, n = 31$ or $\{0,1,6,15,22,26,45,55\}$ for $p = 7, n = 57$.

It is easy to check [29] that for these distributions

Proposition 4.43.

$$\forall k \neq 0 \quad \mathscr{F}_A(k) = \sqrt{p}.$$

Proof. This a special case of Theorem 4.36 which also yields the reciprocal. In this case, one can compute directly

$$\begin{aligned}|\mathscr{F}_A(k)|^2 = \mathscr{F}_A(k)\overline{\mathscr{F}_A(k)} &= \sum_{x,y\in A} \mathrm{e}^{-2\mathrm{i}\pi k(x-y)/n} \\ &= \sum_{x,y\in A, x\neq y} \mathrm{e}^{-2\mathrm{i}\pi k(x-y)/n} + \sum_{x\in A} \mathrm{e}^{-2\mathrm{i}\pi k(x-x)/n} \\ &= \sum_{z=1}^{n-1} \mathrm{e}^{-2\mathrm{i}\pi kz/n} + \sum_{x\in A} \mathrm{e}^{0} = -1 + (p+1) = p.\end{aligned}$$

For practical purposes, it is often convenient to assume that the Singer set begins with (0,1) (up to affine transform). A feature these sets share with FLIDs is the stability of their class under affine transformations, since these transformations only permute the interval distribution. Jon Wild sent me the following collection of FLIDs/Singer sets in $\mathbb{Z}_{31}$:

$$(0,1,3,8,12,18), (0,1,4,10,12,17), (0,1,16,18,22,29),$$
$$(0,1,11,19,26,28), (0,1,15,19,21,24)$$

which are all affine images one of another[40] and can be arranged to tile $\mathbb{Z}_{31}^*$ with appropriate translations.[41] This is an instance of different but homometric tiles which have perfectly balanced saliency for all coefficients. It might seem strange that the tiles have no nil Fourier coefficients in this situation. But it could be surmised from the fact that they tile $\mathbb{Z}_n \setminus \{0\}$, complement of the Dirac distribution (neutral element for $*$), whose DFT is non singular (it is $(n-1,-1,-1,-1,-1,\dots)$).

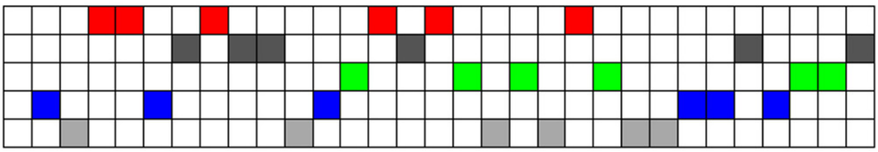

Fig. 4.27. Tiling with different Singer sets modulo 31

[40] See in exercises.

[41] He also found tilings with two or four tiles out of these five.

The interplay with the affine group suggests looking for stability features. Quite often, a Singer set A (or a FLID, actually) is invariant under an affine map or simply multiplication by a constant, i.e. $pA = A$.

Definition 4.44. *Such a p is called a multiplier of A. The set of all multipliers of A is a subgroup of* $\mathbb{Z}_n^*$.

For instance (1,2,4) in $\mathbb{Z}_7$ has multiplier 2: it is actually the orbit of 1 under multiplication by 2. In the above example, if $A = (0,1,4,6,13,21)$ we can see that $5A = (0,5,20,30,65 = 3,105 = 12)$, i.e. $5A+1 = (0,1,4,6,13,21) = A$ or $5(A+8) = A+8$, i.e. 5 is a multiplier of $A+8 = (8,9,12,14,21,29) = (1,5,25) \otimes (8,12)$. It is conjectured that in general, some *translate of* a FLID has multipliers.[42] This might be interesting for composers who play with affine transforms, and should perhaps warrant exploration with non-commutative Fourier transform in the affine group.

To round up this discussion and generalise the last example, let us mention that other tilings of $\mathbb{Z}_n \setminus \{0\}$ by augmentation have been discovered in investigating composer Tom Johnson's autosimilar melodies [54]. To quote him, the absence of 0 is a welcome respite – he devotes a whole chapter to 'punching some holes in the melody,' because:

> The musical interest can be quite a bit greater after punching some holes, however. The different durations define themes with more character, that can be more easily recognised, and this is a great advantage when we are trying to hear the theme in two or three different tempos.

Example 4.45. Consider motif $A = \{1,2,4\}$. It is an orbit of $x \mapsto 2x$ in $\mathbb{Z}_7$. The other orbit is $\{3,6,12 = 5\}$. Thus $A \bigcup 3A = \{1,2,3,4,5,6\}$. From there, one can associate one note to each orbit and thus reach a melody with ratio 2 autosimilarity (picking every other note yields the same melody, twice slower); or tile $\mathbb{Z}_7^*$ with a cross-section of the orbits – say $S = \{2,3\}$ – and its augmentations $\{4,6\}$ and $\{1,5\}$, see Fig. 4.28 with the autosimilar melody first, then the tiling by augmentation.

Fig. 4.28. Autosimilar melody and dual tiling by augmentation

For a larger example, take $n = 31, A = \{1,2,4,8,16\}$.
The other orbits, $3A = \{3,6,12,17,24\}, 5A = \{5,9,10,18,20\}, 7A = \{7,14,19,25,28\}$,

[42] At least when p divides $n-m$. This is equivalent to invariance under (some kind of) affine transformation.

$11A = \{11, 13, 21, 22, 26\}$ and $15A = \{15, 23, 27, 29, 30\}$, partition $\mathbb{Z}^*_{31}$. Now choose any cross-section of these orbits, say $S = \{1, 3, 5, 7, 11, 15\}$; another partition uses $S, 2S, 4S, 8S, 16S$. This enables one to construct a tiling by binary augmentation, see Fig. 4.29, reminding of Fig. 4.27.

Fig. 4.29. Tiling by augmentation modulo 31

Finally, there are tilings of the *complement of a subgroup* using all different ratios of augmentation exactly once: in [53], T. Johnson cites a tiling with augmentations of $\{0, 1, 3\}$ with ratios 1,2,3,4,5,6,7,8 which leave aside all beats congruent with 2 mod. 3, see Fig. 4.30. More complicated examples are also mentioned but we are straying away from the topic of this book. See [4], online supplementary, for details and generalisations.

An open question is the characterisation of these objects by way of using Fourier transform of the multiplicative abelian group $\mathbb{Z}^*_n$, whose structure varies a lot according to the value of n and boils down to multidimensional DFT because of the decomposition of any abelian finite group into a product of cyclic groups.

for Jon Wild

Extra Perfect

Tom Johnson

17/09/2008

Fig. 4.30. Tiling with 013 and augmentations, leaving holes every third eighth-note

Heisenberg's uncertainty principle

We have now seen enough varied material, from both ends of the spectrum so to speak, that we can perhaps address some broader issues. I will broach the question of the cardinality of A (i.e. pcs, for scales or chords, or beats, for periodic rhythms) vs. the zeroes of $\mathscr{F}_A$ which have been so important in different situations.

Much of the material in this book addresses questions of retrieval (phase retrieval for homometry, support retrieval for a complement of a tiling motif, and so on). There is a definite advantage then, when the number of coefficients to retrieve is small. This can be foretold in some measure. It is well known, at least informally, that the DFT spreads when the information (size of time window, sampling, smaller periodicity...) decreases. A few examples:

- If A has a period $d \mid n$, A has at least n/d elements and $\mathscr{F}_A$ is nil except on the subgroup with d elements.
- If A is a FLID (see 4.3.3) then $\#A$ may be small but $\mathscr{F}_A$ never vanishes.
- If A tiles $\mathbb{Z}_n$, then $\#A$ divides n, i.e. is usually comparatively small; $\mathscr{F}_A$ vanishes on $Z(A)$, the union of all elements whose order belongs to $R(A)$ (a reunion of orbits of the action of $\mathbb{Z}_n^*$, cf. Theorem 3.11) and hence a sizeable subset of $\mathbb{Z}_n$. Besides, if A tiles with B, then $\#Z(A) + \#Z(B) \geqslant n-1$ while $\#A \times \#B = n$.

These relationships between the zeroes of $\mathscr{F}_A$ and those of $\mathbf{1}_A$ can be quantised by the following result, commonly used by researchers in various fields but not really pointed out in textbooks, and reminiscent of Heisenberg's famous inequality in quantum physics:

Theorem 4.46 (Discrete Uncertainty Principle).
Let f be a distribution on $\mathbb{C}^n$, $\widehat{f}$ its DFT, and let $\operatorname{Supp}(f)$ *stand for* $\{x \in \mathbb{Z}_n \mid f(x) \neq 0\}$. *Then*

$$\#\operatorname{Supp}(f) \times \#\operatorname{Supp}(\widehat{f}) \geqslant n$$

This means that if $\widehat{f}$ has few zeroes then f has many, and conversely. Notice that $\operatorname{Supp}(\widehat{f})$ is just the complement of the zero set of the Fourier transform.

Proof. Recall Parseval-Plancherel equality (Theorem 1.8):

$$\sum |\widehat{f}(k)|^2 = n \sum |f(k)|^2,$$

and plug in the following elementary inequalities:

$$\sup |f(x)| = \sup_x \left| \frac{1}{n} \sum_k \widehat{f}(k) e^{2i\pi kx/n} \right| \leqslant \frac{1}{n} \sum_k |\widehat{f}(k)| \quad \text{(inverse DFT)}$$

$$\sum |f(k)|^2 = \sum_{k \in \operatorname{Supp}(f)} |f(k)|^2 \leqslant \#\operatorname{Supp}(f) \times \sup |f(x)|^2$$

$$\sum |\widehat{f}(k)| = \sum_{k \in \operatorname{Supp}(\widehat{f})} 1 \times |\widehat{f}(k)|$$

$$\leqslant \sqrt{\sum_{k \in \operatorname{Supp}(\widehat{f})} 1^2} \sqrt{\sum_{k \in \operatorname{Supp}(\widehat{f})} |\widehat{f}(k)|^2} = \sqrt{\#\operatorname{Supp}(\widehat{f})} \sqrt{\sum |\widehat{f}(k)|^2},$$

this last one being Cauchy-Schwarz inequality. Combining all this,

$$\sum |\widehat{f}(k)|^2 = n\sum |f(k)|^2 \leqslant n\#\operatorname{Supp}(f) \sup |f(x)|^2 \leqslant n\#\operatorname{Supp}(f) \left(\frac{1}{n} \sum_k |\widehat{f}(k)| \right)^2$$
$$\leqslant \frac{1}{n}\#\operatorname{Supp}(f)\#\operatorname{Supp}(\widehat{f})\sum |\widehat{f}(k)|^2,$$

hence the result.

The inequality is sharp: if n is a square $n = d^2$ then for $A = \{0, d, 2d, \dots d^2 - d\} = d\,\mathbb{Z}_n$, the DFT $\mathscr{F}_A$ is proportional to $\mathbf{1}_A$ itself and both supports have d elements.

Improvements of this lower bound are known in the case of very simple cyclic groups (for instance when n is prime) – see [83] from which I borrowed much of this section – but are so far of little interest to musicians.

It should be noted that the extreme cases of maximal vs. nil Fourier coefficients are by no means contradictory. It could even be argued that an ubiquitous motif like CDE ("Brother John"...) owes much of its versatility to the dual facts that on the one hand it tiles, having several nil Fourier coefficients ($a_2, a_4 \dots$), but on the other hand it exhibits strong characters: a_6 is maximal since CDE is a chunk of whole-tone scale, cf. Fig. 8.8.

Exercises

Exercise 4.47. Generated scales: peruse the online catalog for Fourier profiles of generated scales in $\mathbb{Z}_{12}$, on
`http://canonsrythmiques.free.fr/MaRecherche/photos-2/`.

Exercise 4.48. Saturation: find musical instances of pc-sets saturated in major thirds, like $\{0,1,4,8\}$ or $\{0,1,4,5,8\}$.

Exercise 4.49. Generated scales: create a scale with 20 generators in some $\mathbb{Z}_n$.

Exercise 4.50. Generated scales: find other occurrences of the complement set of the tresilo rhythm in tango or elsewhere.

Exercise 4.51. ME sets: compute instances of $\mathrm{ME}_{(11,7)}, \mathrm{ME}_{(19,7)}, \mathrm{ME}_{(24,7)}$ ('diatonic scales' in other divisions of the octave).

Exercise 4.52. ME sets: find some other type III ME sets.

Exercise 4.53. Saliency: using the online catalog of Fourier profiles, study the saliencies of some pc-sets in a musical piece of your own choosing. Early 20^{th} century is a good starting point.

Exercise 4.54. FLID: find some FLID, for instance the Hadamard kind for $n = 23$ or $n = 43$, and check its interval content.

Exercise 4.55. FLID: prove Proposition 4.41.

Exercise 4.56. Singer sets: find the affine maps transforming the first motif in Fig. 4.27 in each of the others. Check that they are indeed Singer sets.

Exercise 4.57. Pick up a tiling $A \oplus B = \mathbb{Z}_n$ in the examples given or otherwise. Check Heisenberg's uncertainty principle on each factor.