An Algebra for Periodic Rhythms and Scales

Emmanuel Amiot[*] and William A. Sethares[†]

May 13, 2009

Abstract

This paper shows how scale vectors (which can represent either pitch or rhythmic patterns) can be written as a linear combination of columns of scale matrices, thus decomposing the scale into musically relevant intervals. When the scales or rhythms have different cardinality, they can be compared using a canonical form closely related to Lyndon words. The eigenvalues of the scale matrix are equal to the Fourier coefficients, which leads to a number of relationships between the scale vectors and the decompositions. Overcomplete dictionaries of frame elements can be used for more convincing representations by finding sparse decompositions, a technique that can also be applied to tiling problems. Scale matrices are related to familiar theoretical properties such as the interval function, Z-relation or homometry, all of which can be efficiently studied within this framework. In many cases, the determinant of the scale matrix is key: singular scale matrices correspond to Lewin's special cases; regular matricesallow a simple method of recovering the argument of an interval function and elicit unique decompositions; large determinant values correspond to flat interval distributions.

1 Introduction

A common way to represent musical scales, chords, and tunings that repeat at some specified interval (such as the octave) is to encode into a characteristic function: a *scale vector* of ones (indicating the presence of the note) and zeroes (representing the absence of a note). For instance, the interval of a minor third in 12-tone equal temperament corresponds to the scale vector

$$\{1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0\}', \tag{1}$$

[*]manu.amiot@free.fr
[†]sethares@ece.wisc.edu

1

where the prime $'$ indicates that we conventionally use column vectors. The familiar diatonic scale in C is

$$\{1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1\}' \tag{2}$$

and a whole tone scale is

$$\{1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0\}'. \tag{3}$$

Similarly, the major scale in 19-tone equal temperament is

$$\{1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0\}'. \tag{4}$$

Musical rhythms are also commonly represented analogously [19], with 'one' representing the time of onset of a sound and 'zero' representing times at which no onset occurs. For example, the classic tango 3-3-2 rhythm is

$$\{1, 0, 0, 1, 0, 0, 1, 0\}', \tag{5}$$

King's "standard pattern" in Yoruba music [11] is notated

$$\{1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0\}' \tag{6}$$

and the son clave rhythm is

$$\{1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0\}'. \tag{7}$$

A fundamental similarity between these two representations (of pitch and rhythm) is that the vectors are a shorthand representation of a single period of a periodic element.

When two scale vectors have the same dimension, it is natural to wish to combine them. For instance, adding the whole tone scale to a shifted version of itself yields a complete 12-TET chromatic scale. Adding a regular pattern of two claps per period $\{1, 0, 0, 1, 0, 0\}'$ to a regular pattern of three claps per period $\{1, 0, 1, 0, 1, 0\}'$ yields the 3-2 polyrhythm $\{2, 0, 1, 1, 1, 0\}'$, where the "2" indicates that two notes coincide at the first timepoint.

This paper addresses several issues that arise with scale vectors. We show three techniques for decomposing scale vectors, thus expressing one scale (or chord or tuning or rhythm) as a linear combination of others. In Section 2, the set of all transpositions of a scale is gathered into a circulant matrix. When this matrix is invertible, it is possible to express any scale (or chord) as a linear combination of these transpositions in a unique way. This provides an analog of the Fourier transform, but with basis functions defined by musical scales rather than by sinusoids. Section 3 explains how vectors of different lengths can nevertheless be compared.

There is a striking relationship with the Fourier coefficients, which is explored in Section 4, including discussion of extreme values of the determinant of this matrix

in connection with Lewin's 'special cases' [13]. This relationship leads to interesting techniques related to the interval function and the Z-relation that bypass the computation of Fourier coefficients.

When the matrix is not invertible, an exact decomposition is generically impossible. The pseudo-inverse gives the best L^2 match, while a linear program can be used to find the best L^1 solution. This last is particularly useful in supplying sparse answers, which tend to give more musically sensible results. Alternatively, a dictionary of scales can be used to form a frame (an over-complete spanning set) which can also be used to decompose a given scale; there are many solutions to choose from. Again, the linear programming L^1 solutions tend to be easier to interpret than the solutions given by the pseudo-inverse. Also, there is some measure of control over the result, as a preference can be expressed for some pitches by choosing the specific linear form to be minimized by the algorithm. This opens promising alleys into the computation of rhythmic canons and several examples are shown in Section 5. A *Mathematica* notebook is available on the website [21] which provides function definitions that allow the reader to easily duplicate and extend the examples given throughout the paper.

2 Scale Transforms

2.1 Definitions and Examples

Consider scales in a n-TET or rhythmic patterns with n timepoints. When not otherwise specified in the examples, $n = 12$. Let $s = \{s_0, s_1, ...s_{n-1}\}'$ be a *scale vector* of dimension n, and let $s(i)$ be the same vector after a circular rotation of i steps. The *scale matrix*

$$S = [s(0)|s(1)|...|s(n-1)]$$

is a circulant matrix formed using the scale vector and all its rotations as columns. To be specific, the $(i, k)^{th}$ element of S is $s_{(i-k) \mod n}$.

If S is invertible, the columns can be viewed as forming a basis of $\mathbb{R}^n$ and any element $r \in \mathbb{R}^n$ can be written uniquely as a linear combination of these basis elements:

$$r = x_0\, s(0) + x_1\, s(1) + \ldots + x_{n-1}\, s(n-1) = \sum_{i=0}^{N-1} x_i s(i) = Sx.$$

The transformation $T = S^{-1}$ maps from r to x while S gives the inverse transformation from x to r. The simplest case is the identity matrix, $S_I = I_n$ formed from the scale vector

$$e_1 = \{1, 0, 0, ..., 0\}'. \tag{8}$$

In this case, any scale vector x is expressed as itself: $r = S_I x = I_n x = x$.

Example 1 (A Basis of Major Thirds) *Let $s = \{1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0\}'$ be the dyad of major thirds. The corresponding circulant matrix S is invertible, with determinant $|S| = 16$. The major seventh chord $M_7 = \{1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1\}'$ can be expressed in the basis of major thirds by calculating*

$$TM_7 = \{1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0\}' = f \tag{9}$$

where $T = S^{-1}$, which shows how the major seventh chord is expressible directly as a sum of two major thirds a fifth apart.

Example 1 shows how one chord (or scale or tuning) can be expressed in terms of another, in this case, how the major seventh chord can be expressed in terms of major thirds. The basis can consist of any scale for which the S matrix is invertible. The next example shows how minor scales can be decomposed in terms of diatonic scales.

Example 2 (A Diatonic Basis) *Let s be the diatonic scale (2). The corresponding circulant matrix*

$$S = \begin{pmatrix}
1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\
1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\
1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1
\end{pmatrix} \tag{10}$$

is invertible, with determinant $|S| = -7$. Let

$$T = S^{-1} = \frac{1}{7} \begin{pmatrix}
3 & 3 & 3 & 3 & 3 & 3 & -4 & -4 & -4 & -4 & -4 & 3 \\
3 & 3 & 3 & 3 & 3 & 3 & 3 & -4 & -4 & -4 & -4 & -4 \\
-4 & 3 & 3 & 3 & 3 & 3 & 3 & -4 & -4 & -4 & -4 & -4 \\
-4 & -4 & 3 & 3 & 3 & 3 & 3 & 3 & -4 & -4 & -4 & -4 \\
-4 & -4 & -4 & 3 & 3 & 3 & 3 & 3 & 3 & -4 & -4 & -4 \\
-4 & -4 & -4 & -4 & 3 & 3 & 3 & 3 & 3 & 3 & -4 & -4 \\
-4 & -4 & -4 & -4 & -4 & 3 & 3 & 3 & 3 & 3 & 3 & -4 \\
-4 & -4 & -4 & -4 & -4 & -4 & 3 & 3 & 3 & 3 & 3 & 3 \\
3 & -4 & -4 & -4 & -4 & -4 & 3 & 3 & 3 & 3 & 3 & 3 \\
3 & 3 & -4 & -4 & -4 & -4 & 3 & 3 & 3 & 3 & 3 & 3 \\
3 & 3 & 3 & -4 & -4 & -4 & -4 & 3 & 3 & 3 & 3 & 3 \\
3 & 3 & 3 & 3 & -4 & -4 & -4 & -4 & -4 & 3 & 3 & 3 \\
3 & 3 & 3 & 3 & 3 & -4 & -4 & -4 & -4 & -4 & 3 & 3
\end{pmatrix}.$$

The melodic minor scale $m_1 = \{1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1\}'$ can be expressed in the basis of diatonic scales by calculating $Tm_1 = \{1, 0, 0, 0, 0, -1, 0, 0, 0, 0, 1, 0\}'$, showing how this minor scale can be expressed as a (signed) sum of three major scales. Similarly, the harmonic minor scale $m_2 = \{1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1\}'$ can be expressed as $Tm_2 = \{1, 0, 0, 1, 0, -1, 0, 0, 0, 0, 0, 0\}'$: the scale of C minor is thus the sum of C major and E♭ major, minus F major. Alternatively, this means that the tones common to C and E♭ major, together with the other tones in one of these two scales that lie outside of C minor, make up F major. This is shown graphically in Fig. 1.

The next two examples show the complementary relationship between the diatonic and pentatonic scales.
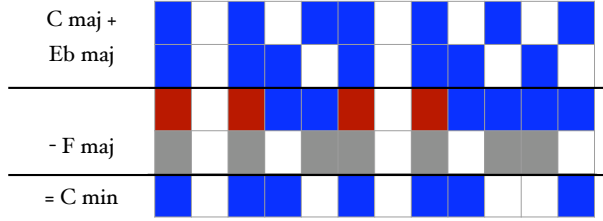
Figure 1: Sum and difference of scales

Example 3 (A Diatonic Basis) *Let s, S, and T be defined as in Example 2. The pentatonic scale*

$$p = \{1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0\}' \tag{11}$$

can be expressed in the basis of diatonic scales by calculating

$$T_p = \frac{1}{7}\{1, 1, 1, 1, 1, 1, -6, 1, 1, 1, 1, 1\}'.$$

Thus the pentatonic scale is a collection of 11 diatonic scales minus a weighted diatonic scale at the tritone. Alternatively, consider the aggregate: it is one seventh of the sum of all twelve major scales. Subtract F♯ major to get its complement set. Thus the pentatonic scale p can be expressed succinctly in the diatonic basis as

$$T_p = \frac{1}{7}\mathbf{1}_{12} - \{0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0\}'$$

where

$$\mathbf{1}_n = \{1, 1, ..., 1\}' \tag{12}$$

is the chromatic scale vector of size n.

Example 4 (A Pentatonic Basis) *Let s be the pentatonic scale (11). The corresponding circulant matrix S is invertible, with determinant $|S| = 5$. The diatonic scale d in (2) is expressible in the basis of pentatonic scales as $T_d = \frac{1}{5}\{1, 1, 1, 1, 1, 1, -4, 1, 1, 1, 1, 1\}'$. Thus the diatonic scale is a collection of 11 pentatonic scales minus a weighted pentatonic at the tritone. Or, using again the aggregate as 1/5 of the sum of all pentatonic scales:*

$$T_d = \frac{1}{5}\mathbf{1}_n - \{0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0\}'.$$

When using the familiar (discrete) Fourier Transform, a vector x is decomposed into a weighted set of sinusoidal basis elements. It is not clear, however, that a sinusoidal basis is necessarily the most instructive way to represent scale vectors. Viewing

the T matrix as a transform (with inverse transform S) suggests the interpretation that the scale vector x is decomposed into a weighted set of basis elements consisting of musically sensible terms: the set of all rotations (transpositions) of the scale s that define S. This is especially meaningful when the answer consists of a small number of simple coefficients, like 0, 1 or -1, where the minus sign implies a complementary relationship.

Another interesting use of these decompositions starts from the multiset m of the number of occurences of all pitches in a musical piece. Then Tm (with T defined as above) provides a picture of the distribution of all diatonic scales used in the piece.

2.2 Some General Features

2.2.1 Singular Scale Matrices

So far the construction of a scale matrix yields a square $n \times n$ matrix. Of course, this matrix S need not always be invertible. In some cases, the columns of S repeat themselves (consider the scale matrix of the tritone, or of the whole-tone scale for instance) which means that the model of a square matrix is inappropriate.

Straightforward calculation shows that any dyad in 12-TET (other than the major third from Example 1), has $|S| = 0$. The whole tone scale (3) has $|S| = 0$. In these cases, it is not possible to represent an arbitrary scale vector in the basis defined by the columns of S.[1] A simple characterization of these singular cases is

Proposition 1 *A scale yields a singular scale matrix iff the scale is a linear combination of its translates.*

For instance, the whole-tone scale is equal to every translate by an even number of semitones. Less trivially, a minor third is a combination of other minor thirds, as for instance

$$(C, E\flat) - (E\flat, F\sharp) + (F\sharp, A) = (A, C)$$

This might appear to be a consequence of the minor third dividing the octave equally, but this is wrong, since the scale matrix of the major third is invertible, and the scale matrix of the fifth is singular.[2] Letting $(0, d)$ represent the dyadic scale vector with a one in the first and d^{th} entries, the general result is:

Theorem 1 *The scale matrix of $(0, d)$ in n-TET is always invertible if n is odd. If $n = 2^v q$ with q odd, the scale matrix is invertible iff 2^v divides d, the span of the dyad.*

---

[1] In special cases, such a decomposition exists, but will not be unique: for instance the guidonian hexachord $\{1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 0\}'$ is a sum of three fifths, though the scale matrix of the fifth is singular.

[2] The sum of all fifths beginning on one whole-tone scale is equal to the whole, as is the similar sum starting on the other whole-tone scale. So one perfect fifth is a linear combination of all the others.

Otherwise, the rank of the scale matrix $(0, d)$ is equal to $n - \gcd(d, n)$; it is minimal for $d = n/2$, the equal division of the octave (the generalized tritone). In that case it is equal to $n/2$.

[All proofs may be found in the appendices.] For instance, the only dyads that allow expression of any scale in quarter-tone TET (i.e. $n = 24$) are for $d = 8$, i.e. the usual major third (and its inversion). It is only for $n = 2^v$ that all dyad matrices are singular.

2.2.2 Scale Matrix and Complementation

There are some close relationships between the representation of a scale vector x with respect to a scale matrix S and the representation of the complementary scale vector $\mathbf{1}_n - x$ with respect to S.

Theorem 2 *Let s be a scale vector with corresponding scale matrix S and let z be the representation of the n-dimensional scale vector x in the basis defined by the columns of S. Then the complementary scale $\mathbf{1}_n - x$ is represented in the basis defined by the columns of S by $\frac{1}{m}\mathbf{1}_n - z$, where m is the number of nonzero elements of x.*

It is also possible to characterize the representation of a scale vector directly in terms of the basis defined by its complement.

Theorem 3 *Let the n-dimensional scale vector x define the scale matrix S_x and let $\mathbf{1}_n - x$ define the scale matrix $S_{\mathbf{1}_n - x}$. Assuming that S_x is invertible, x can be represented in the basis of the columns of $S_{\mathbf{1}_n - x}$ by $z = \frac{1}{m}\mathbf{1}_n - e_1$ where m is the number of nonzero elements in x and e_1 is given by (8).*

Theorem 3 explains Examples 3 and 4, since the diatonic scale (2) is complementary to (a shifted version of) the pentatonic scale (11). Any pair of complementary scales can be decomposed similarly, without carrying out the calculations.

3 Canonical Forms

Each scale vector represents a single cycle of a periodic structure. When the vector refers to pitch, the period is (typically) the octave. When the vector refers to rhythm, the period refers to one temporal cycle. In either case, representations are not unique, and this section explores a canonical form that can represent any scale vector and allows easy comparison between different length vectors. There are two kinds of ambiguities.

Since a periodic sequence can be thought of as starting at any element, one ambiguity in representation is that two scale vectors of the same length may be related

by circular rotation. Let $q^\ell(x)$ be the operator that circularly rotates x by ℓ positions to the right. For instance, $x = \{1, 1, 0, 1\}'$ is equal to $y = \{1, 1, 1, 0\}'$ after a single rotation to the right, that is, $y = q^1(x)$. Similarly, the diatonic scale vector d in (2) is equal to the Yoruba rhythm scale vector y in (6) after seven rotations to the right, $y = q^7(d)$. In general, the scale matrices S_x and $S_{q^\ell(x)}$ corresponding to x and $q^\ell(x)$ have the same columns but in a different order. Hence the representations are the same up to multiplication by some permutation matrix P, that is, $S_x = PS_{q^\ell(x)}$.

The most interesting ambiguity is in length. For example, a short rhythm such as $x_1 = \{1, 1, 0, 1\}'$ represents three events that occur at one-quarter, one-half, and three-quarters of a period. If the time-length of the period is fixed, this is identical to $x_2 = \{1, 0, 1, 0, 0, 0, 1, 0\}'$, which also represents three events that occur at one-quarter, one-half, and three-quarters of the period. Thus the two rhythms x_1 and x_2 are related by an upsampling by 2 (insertion of zero between each element). Similarly, the whole tone scale (3) in 12-TET is the same as an upsampling of the chromatic scale in 6-TET ($= \mathbf{1}_6$).

More generally, let $\uparrow m$ (upsampling by m) define the operation that inserts $m - 1$ zeros between each element of a scale vector. Observe that rotation and upsampling can be applied in either order, specifically,

$$q^{m\ell}(x \uparrow m) = (q^\ell x) \uparrow m. \tag{13}$$

In words, upsampling by m and then rotating by $m\ell$ gives the same answer as rotating by ℓ and then upsampling by m.

Given two scale vectors x and y, if there are integers ℓ and m such that either

$$x = q^\ell(y \uparrow m) \text{ or } y = q^\ell(x \uparrow m)$$

then $x \sim y$ are said to be in the same equivalence class. In terms of pitch, $x \sim y$ means that the two scales contain the same intervals in the same order. In terms of rhythm, $x \sim y$ means that the time intervals between successive events in the two rhythms are identical. A *canonical form* y is defined to be the smallest (in length) representative of its equivalence class. Of course, any rotation of y has the same length. A unique shortest member can be chosen by treating the elements of the binary vector $q^\ell(y)$ as if they were the digits of a binary number, calculating

$$c_\ell = \sum_i 2^{q^\ell(y(i))} \tag{14}$$

for each rotation ℓ. The scale vector $q^\ell(y)$ corresponding to the smallest c_ℓ is the unique canonical form of this equivalence class. For example, $\{1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0\}'$ is the canonical form of both the diatonic scale vector (2) and the Yoruba rhythm vector

(6).[3]

One way to find canonical forms is to downsample: to remove elements from the vector so as to leave only every m^{th} element remaining.[4] Unlike standard downsampling (as in [20]) however, downsampling of scale vectors is only allowed if all elements removed are zero. For instance, $\{1, 0, 1, 0, 0, 0, 1, 0\}'$ can be downsampled by 2 to give $\{1, 1, 0, 1\}'$ but $\{1, 0, 1, 0, 0, 1, 1, 0\}'$ cannot be downsampled, because doing so would require removal of nonzero elements, thus changing the essential character of the scale or rhythm.

Downsampling can be stated formally as follows: let n be the number of elements in a scale vector x. Since x is binary, $|x| = \mathbf{1}'_n x$. Let m be a divisor of n (i.e., $d = \frac{n}{m}$ for some integer d) and let $\hat{x} = \{x(0), x(m), x(2m), ..., x((d-1)m)\}'$. Then downsampling by m is allowed if x and $\hat{x}$ have the same number of nonzero elements, that is, if $\mathbf{1}'_n x = \mathbf{1}'_d \hat{x}$. In this case, we write $\hat{x} = x \downarrow m$. Up and down sampling are inverses whenever downsampling is defined. Thus $(x \downarrow m) \uparrow m = x$ whenever the downsampling is allowed, while $(x \uparrow m) \downarrow m = x$ always holds.

A simple algorithm for finding the canonical form of a scale vector x is to attempt all possible downsamplings by all possible divisors of n. A more algebraic way is the computation of the gcd of the differences of elements of x, which generates a cyclic group. Whenever downsampling occurs, the scale vector decreases in length. The canonical form is found when no more downsamplings are possible and the result is rotated to smallest (binary) value as in (14).

It is possible to express the decomposition of a scale vector directly in terms of a decomposition of its canonical form.

Proposition 2 *Suppose x is a scale vector of size n with scale matrix S_x, and that y can be expressed in the basis defined by the columns of S_x as $S_x z = y$. Let $S_{x \uparrow m}$ be the scale matrix corresponding to the upsampled scale vector $x \uparrow m$. Then $y \uparrow m$ can be expressed in a basis defined by the columns of $S_{x \uparrow m}$ as $S_{x \uparrow m} z_m = y \uparrow m$, where $z_m = z \uparrow m$.*

Let $\psi_1, \psi_2, ..., \psi_n$ denote the eigenvalues of S_x. Then the upsampled scale matrix $S_{x \uparrow m}$ corresponding to $x \uparrow m$ has the same n eigenvalues, each with multiplicity m.[5] Hence the determinants are related by $|S_x|^m = |S_{x \uparrow m}|$. For instance, the determinant of the scale matrix generated by $\{1, 1, 0\}'$ in 3-TET is 2. Upsampling by $m = 2$ gives a scale matrix generated by $\{1, 0, 1, 0, 0, 0\}'$ in 6-TET, with determinant 4, while upsampling by $m = 4$ gives the major third dyad in 12-TET (which has determinant

---

[3]Word theorists call these *Lyndon words*, computer scientists talk of *circlists*. Forte's *basic form* is almost identical but its definition is needlessly complicated.

[4]For theoretical applications of downsampling in rhythmic canons see [2].

[5]It is the tensorial product of S_x and I_m.

16 as shown in Example 1). In particular, $|S_x| = 0$ exactly when $|S_{x \uparrow m}| = 0$. Thus upsampling (and downsampling, when possible) of a scale vector cannot change the invertibility of the corresponding scale matrices.

Since canonical forms preserve the representation and decomposition of scale vectors, they can be used even when two scale vectors have different lengths. Suppose that the canonical form x is of length n and the canonical form y is of length m. Let $k = \text{lcm}(n, m)$. Then $x_k = x \uparrow \frac{k}{n}$ is the same length as $y_k = y \uparrow \frac{k}{m}$, and any operations such as creation of scale matrices and the decompositions of scale vectors can be carried out on x_k and y_k and the results transfered immediately to all members of their equivalence classes as in Proposition 2.

Example 5 (Clave in basis of the Tango) *Let x be the 8-element tango rhythm (5), y the 16-element clave rhythm (7), and $S_{x \uparrow 2}$ be the scale matrix corresponding to the upsampled tango. Then the clave can be expressed in the basis of the (upsampled) tango as*

$$z = S_{x \uparrow 2}^{-1} y = \{1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0\}' - \frac{1}{3}\mathbf{1_{16}}.$$

4 Relationship with the Fourier Transform

Because of the circulant structure of the scale matrix S, there is a close relationship between the representation of a scale vector in a basis of scale elements and its representation in the Fourier basis. This section explores this relationship and some of its consequences.

4.1 DFT of a Chord or Scale

The DFT of a pc-set is identified with a subset of the cyclic group $\mathbb{Z}_n$ by $\mathcal{F}_A(t) = \sum_{A \in s} e^{-2i\pi kt/n}$. This is the Discrete Fourier Transform (DFT) of the characteristic function 1_A of the subset A, where A specifies the nonzero indices of a scale vector s (i.e., $\mathcal{F}_A(t) = \sum_{k=0}^{n-1} s(k)e^{-2i\pi kt/n}$) . For instance, $A = (0, 2, 4, 5, 7, 9, 11)$ specifies the diatonic scale (2) and the associated scale matrix S (10). Detailed discussion of special values and features of this DFT can be found in [1]. All information about A is present in its DFT, since by the inverse Fourier transform

$$s' = 1_A(k) = \frac{1}{n} \sum_{t \in \mathbb{Z}_n} \mathcal{F}_A(t)e^{2i\pi kt/n}.$$

This can be simplified using the present matrix notation, because, (as shown in [9]),

the eigenvalues of S (which are the same as the eigenvalues of S') are

$$\psi_m = \sum_{k=0}^{n-1} s_k e^{-2i\pi m\, k/n} \tag{15}$$

with corresponding normalized eigenvectors

$$y_m = \frac{1}{\sqrt{n}} \{1, e^{-2i\pi m/n}, ..., e^{-2i\pi(n-1)m/n}\}'. \tag{16}$$

Equation (15) shows that the eigenvalues of S are given by the DFT of the first column of S. Conversely, the first column of S is the inverse DFT of the eigenvalues.

The unitary matrix Ω whose columns are the eigenvectors y_m is called the DFT matrix. Every scale matrix S (indeed any circulant matrix with first column equal to s'), has the diagonalization

$$\Omega^{-1} S\, \Omega = \begin{pmatrix} \psi_0 & 0 & \cdots & 0 \\ 0 & \psi_1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \psi_{n-1} \end{pmatrix}.$$

This a bijective and linear transformation between three algebras:

1. the algebra of columns $\mathbb{C}^n$, identified with the maps from $\mathbb{Z}_n$ to $\mathbb{C}$,

2. the algebra of circulating matrices, and

3. the algebra of diagonal matrices.

The product in the first algebra is the convolution product. The last isomorphim stems from the fact that all circulant matrices are polynomials in

$$J = \begin{pmatrix} 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & 1 & 0 \end{pmatrix}.$$

As $J^n = I_n$ (J is a circular permutation of the canonical basis), its eigenvalues are the nth roots of unity that appear in the DFT formula, and the eigenvectors of J remain eigenvectors for every polynomial in J, e.g., for every circulating matrix $S_v = \sum_{k=0}^{n-1} v_k J^k$.

4.2 Singular Scale Matrices and Lewin's Special Cases

4.2.1 Interval function

Since the determinant of any matrix is the product of its eigenvalues,

$$|S| = \prod_{i=0}^{n-1} \psi_i.$$

11

in particular, the determinant of S is zero when at least one of its eigenvalues is zero.

In his first paper [13], David Lewin endeavoured to reconstruct a pc-set B from its intervallic relationship with a given pc-set A. Let a, b denote the scale vectors (columns) as elements of the first algebra. The intervallic function, counting the number of occurences of any possible interval between elements of A and B, is a convolution product:

$$\text{IFunc}\,(A, B)' = a \star (-b) \qquad (17)$$

This can be translated into matrix form using the product ST' of the scale matrices S and T corresponding to a and b[6]. Since S and T are circulant, so are T' and ST'; hence ST' is completely defined by a single column. For example, in $\mathbb{Z}_8$, IFunc $((0, 1, 3, 7), (0, 1, 4, 6))$ can be calculated as the first column of

$$
\begin{pmatrix}
1&1&0&0&0&1&0&1\\
1&1&1&0&0&0&1&0\\
0&1&1&1&0&0&0&1\\
1&0&1&1&1&0&0&0\\
0&1&0&1&1&1&0&0\\
0&0&1&0&1&1&1&0\\
0&0&0&1&0&1&1&1\\
1&0&0&0&1&0&1&1
\end{pmatrix}
\begin{pmatrix}
1&1&0&0&1&0&1&0\\
0&1&1&0&0&1&0&1\\
1&0&1&1&0&0&1&0\\
0&1&0&1&1&0&0&1\\
1&0&1&0&1&1&0&0\\
0&1&0&1&0&1&1&0\\
0&0&1&0&1&0&1&1\\
1&0&0&1&0&1&0&1
\end{pmatrix}
=
\begin{pmatrix}
2&3&1&2&1&3&2&2\\
2&2&3&1&2&1&3&2\\
2&2&2&3&1&2&1&3\\
3&2&2&2&3&1&2&1\\
1&3&2&2&2&3&1&2\\
2&1&3&2&2&2&3&1\\
1&2&1&3&2&2&2&3\\
3&1&2&1&3&2&2&2
\end{pmatrix}.
$$

It is in practice there is a shortcut: the multiplication of T by the vector s', leaving the answer in the form of the row $s'T$:

$$
\begin{pmatrix}1&1&0&1&0&0&0&1\end{pmatrix}
\begin{pmatrix}
1&0&1&0&1&0&0&1\\
1&1&0&1&0&1&0&0\\
0&1&1&0&1&0&1&0\\
0&0&1&1&0&1&0&1\\
1&0&0&1&1&0&1&0\\
0&1&0&0&1&1&0&1\\
1&0&1&0&0&1&1&0\\
0&1&0&1&0&0&1&1
\end{pmatrix}
=
\begin{pmatrix}2&2&2&3&1&2&1&3\end{pmatrix}.
$$

The Fourier transform of the interval function is $\mathcal{F}(\text{IFunc}\,(A, B)) = \overline{\mathcal{F}_A} \times \mathcal{F}_B$, meaning that the Fourier coefficients are multiplied one to one. This means that knowledge of IFunc (A, B) and B entails a complete knowledge of A, *except when $\mathcal{F}_B$ vanishes*. In such cases, the corresponding value of $\mathcal{F}_A$ cannot be recovered. Lewin called these the 'special cases.'

Example 6 *Let B be the whole-tone scale with b given by (3). Then any odd dyad A will have the same intervallic relation with B, that is,*

$$\text{IFunc}\,(A, B) = \mathbf{1}_{12}. \qquad (18)$$

Taking the Fourier Transform of both sides of (18) gives

$$\mathcal{F}(\text{IFunc}\,(A, B)) = \mathcal{F}(\mathbf{1}_{12}) = (12, 0, 0, 0, \ldots 0).$$

Since the Fourier transform of a convolution is the product of the transforms, (17) shows that

$$\mathcal{F}(\text{IFunc}\,(A, B)) = \mathcal{F}_A \times \mathcal{F}_B = \mathcal{F}_A \times (6, 0, 0, 0, 0, 0, 6, 0, 0, 0, 0, 0).$$

---

[6]The transposition takes into account the minus sign before b. In full generality, it is necessary to use the transpose conjugate of T, but the conjugacy is irrelevant in the present setting since T is real valued.

The only way this equality can hold is if both $\mathcal{F}_A(0) = 2$ and $\mathcal{F}_A(6) = 0$. Thus there are two constraints on A; the ten remaining coefficients may take on any values. This allows for instance both the fifth $\mathfrak{f}$ (9) and the minor third $\mathfrak{t}$ (1), with respective DFT's

$$\mathcal{F}_f = \left\{2, 0.13 + \frac{i}{2}, \frac{3}{2} - \frac{\sqrt{3}}{2}i, 1 + i, \frac{1}{2} - \frac{\sqrt{3}}{2}i, \frac{2 + \sqrt{3}}{2} + \frac{i}{2}, 0, \right.$$
$$\left. \frac{2 + \sqrt{3}}{2} - \frac{i}{2}, \frac{1}{2} + \frac{\sqrt{3}}{2}i, 1 - i, \frac{3}{2} + \frac{\sqrt{3}}{2}i, 0.13 - \frac{i}{2}\right\}$$
$$\mathcal{F}_t = \{2, 1 - i, 0, 1 + i, 2, 1 - i, 0, 1 + i, 2, 1 - i, 0, 1 + i\}$$

to share the same IFunc with B. In his dissertation [16], Ian Quinn summed up the different forms of these singular cases as Lewin exposed them, using a characterization of 'balances.' In the example above, $\mathfrak{t}$ and $\mathfrak{f}$ share the Fourier Property 6, also known as the whole-tone-scale property [13], e.g. they both "have the same number of notes in one whole-tone set, as they have in the other." This corresponds to the equality of Fourier coefficients $\mathcal{F}_f(6) = \mathcal{F}_t(6) = 0$.

Example 7 *A less obvious example is the ascending melodic minor scale $(0, 2, 3, 5, 7, 9, 11)$ with scale vector $m = \{1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1\}'$. As two of its Fourier coefficients ($\mathcal{F}_m(2) = \mathcal{F}_m(10) = 0$) are zero, the determinant of its scale matrix S is zero, meaning that it is not possible to express all scales as linear combinations of rotations of m[7]. The intervallic relation to $(0, 1, 6, 7)$ is the same as to $(3, 4, 9, 10)$, because only the second and tenth Fourier coefficients differ:*

$$\mathcal{F}_{(0,1,6,7)} = \left\{4, 0, 3 - i\sqrt{3}, 0, 1 - i\sqrt{3}, 0, 0, 0, 1 + i\sqrt{3}, 0, 3 + i\sqrt{3}, 0\right\}$$
$$\mathcal{F}_{(3,4,9,10)} = \left\{4, 0, -3 + i\sqrt{3}, 0, 1 - i\sqrt{3}, 0, 0, 0, 1 + i\sqrt{3}, 0, -3 - i\sqrt{3}, 0\right\}.$$

Thus scale matrices provide a simple algebraic characterization of Lewin's special cases, namely that the scale matrix S of B be singular, or $|S| = 0$. This is interesting because, as Quinn points out, the corresponding Fourier Property is fairly obscure[8] though the computation of the Fourier coefficient by a geometrical method called the *multiplication principle* is more comprehensible. The computation of a determinant, or rank of circulant matrix, is perhaps a more straightforward alternative.

The matrix scale representation also provides a neat way to recover a pc-set B from A and the combination IFunc (A, B).

---

[7]The sum of those scales beginning on C, F, F♯, B is equal to the sum of the same scales beginning on D, D♯, G♯, A. Hence by Prop. 1, the scale matrix is singular.

[8]For instance, the 'tritone property' states: "for any (0167)-set K, [it] has the same number of notes in common with T3(K), as it has in common with K."

Theorem 4 *Let A, B, and C be the circulating matrices associated with scales a, b and $c = $ IFunc (A, B). Then $AB' = C$, and hence whenever $|A| \neq 0$, $B' = A^{-1}C$, i.e. the scale vector b is defined by the first row of $A^{-1}C$, e.g. $b' = c'A^{-1}$.*

This has the additional advantage that A^{-1} need be computed only once if one tries the exercise for different targets b.

Example 8 *We have already computed the inverse of the scale matrix S of the diatonic scale s (2). Say a scale t has the IFunc (2, 3, 2, 2, 3, 1, 4, 1, 3, 3, 1, 3) with s; then we compute*

$$(2, 3, 2, 2, 3, 1, 4, 1, 3, 3, 1, 3)S^{-1} = (0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1)$$

and recognize Tristan's chord.

For singular scales (i.e. scales with a singular matrix) it might be interesting to try the techniques developed *infra*, by way of L^2 or L^1 closest solutions.

4.2.2 Homometry

The difficult question of homometry, or Z-relation as it is called traditionally in American musicology,[9] is nicely expressed in the matrix framework. The following definitions are equivalent (see for instance [17]):

Definition 1 *The scales defined by vectors a, b are* **homometric** *if any of the following (equivalent) conditions are true:*

1. *IFunc (a, a) = IFunc (b, b) (every interval occurs the same number of times inside a and b);*

2. *$AA' = BB'$ where A, B are the scale matrices of a, b.*

3. *$|\mathcal{F}_A| = |\mathcal{F}_B|$, meaning that the elements of the diagonal matrices $\Omega^{-1}A\Omega$ and $\Omega^{-1}B\Omega$ have one to one the same magnitude;*

4. *There exists a matrix U, which is both a circulating matrix like A, B, and unitary like $\Omega : U\overline{U}' = I$, such that $B = UA$.*

These last factors are called spectral units *in the crystallography literature. In terms of vectors, it is equivalent to $b = u \star a$. In terms of Fourier coefficients, the values of $\mathcal{F}_U$ are the quotients of the Fourier coefficients of a and b and thus lie on the unit circle.*

---

[9]The original definition precludes the easy case, when the scales are identical under action of the T/I group.

It is clear that any two isometric scales are homometric, but the reverse is not true (though counter-examples are uncommon). So far a general formula producing all possible Z-related pc-sets is unknown. A possible approach would be the computation of all possible spectral units U (bearing in mind that U multiplied with a scale matrix does not automatically yield a regular scale matrix, e.g. a matrix with 1's and 0's).

From the last relation it is clear that the coefficients of U are rational numbers (the denominator being the determinant of A), at least when A is invertible from Cramer's formulas.[10] Hence the following result that gives a (large) upper size for the group of spectral units:

Theorem 5 *The eigenvalues of U are roots of unity, with maximal possible order n if n is even or $2n$ if n is odd.*

This enables (in theory) a specification of all possible spectral units via the inverse Fourier transform.[11]

Example 9 *Well-known homometric scales in $\mathbb{Z}_{12}$ are $a = (0, 1, 4, 6)$ and $b = (0, 1, 3, 7)$. Computation of $U = BA^{-1}$ yields the spectral unit $u = \dfrac{1}{4}(1, 1, 0, 1, -1, -2, 1, 1, 0, 1, -1, 2)$. Its Fourier coefficients are roots of unity of order 2, 3, 4 or 6:*

$$\mathcal{F}_U = \left\{ 1, -\frac{i}{2} + \frac{\sqrt{3}}{2}, \frac{1}{2} + \frac{i\sqrt{3}}{2}, -1, i, -\frac{i}{2} - \frac{\sqrt{3}}{2}, -1, \frac{i}{2} - \frac{\sqrt{3}}{2}, 1, i, \frac{1}{2} - \frac{i\sqrt{3}}{2}, \frac{i}{2} + \frac{\sqrt{3}}{2} \right\}.$$

We leave it to the reader to check that the matrix $U = \frac{2}{n} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} - I_n$ is a spectral unit, and that for any hexachord (= pc-set with $n/2$ elements, for even cardinalities n) with scale matrix S, $U\,S$ is the scale matrix of its complement. This is a rather indirect proof of Babbit's hexachord theorem.

4.3   Largest Values of $|S|$

Parseval's identity implies that the sum of the squares of the eigenvalues is a constant:

$$\sum |\psi_m|^2 = n\,d.$$

---

[10]In his seminal paper, Rosenblatt [17] proves that this is true even for singular matrixes, though he mentions only once circulating matrixes and this in order to clarify the null space in such cases, wherein several different spectral units can satisfy the equation $b = u \star a$ for a given a, b pair.

[11]Not all combinations of roots of unity need be considered, as the condition that U is rational-valued excludes many of the possibilities (for instance, the final terms must be conjugate to the initial terms in reverse order).

Hence among d note scales, the largest determinant $|S|$, e.g. the largest value of $\prod |\psi_m|$, occurs when the eigenvalues, i.e., the Fourier coefficients, are most nearly equal.[12]

Jon Wild [22] has investigated special scales that provide such distributions of the Fourier coefficients. The FLID (FLat Interval Distribution) is defined as a pc-set A with constant interval vector: $iv(A) =$ IFunc $(A, A) = (d, m, m, m, \dots m)$. (The first value must be the cardinality of the pc-set) For obvious reasons, it is not always possible to get a FLID – for instance, the total number of intervals must be d^2 and hence $m \times (n-1) = d^2 - d$, meaning $n-1$ must divide $d(d-1)$. A simple example is $A = (1, 3, 4, 5, 9)$ in 11-TET, with $iv(A) = (5, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2)$: each interval (except the unison) appears exactly twice. It is straightforward to check that a FLID has a Flat Fourier coefficient distribution, as the DFT of $iv(A)$ is precisely $|\mathcal{F}_A|^2$. The computation when IFunc $(A, A) = (d, m, m, m, \dots m) = (d - m)\delta_0 + m\mathbf{1}_n$ yields

$$\mathcal{F}_{iv(A)} = |\mathcal{F}_A|^2 = (d + (n-1)m, d-m, d-m, \dots d-m) = (d^2, d-m, d-m, \dots d-m)$$

We have proved that the reverse is also true:

Theorem 6 *A is a FLID if and only if the magnitude of all its Fourier coefficients (except the zeroth) are equal:*

$$iv(A) = \textit{IFunc}\ (A, A) = (d, m, m, m, \dots m) \iff \forall m = 1 \dots n-1, |\psi_m|^2 = d - m$$

This case gives the maximum possible value of the determinant of the associated scale matrix. Since $m = \dfrac{d(d-1)}{n-1}$, $d - m = d\dfrac{n-d}{n-1}$ and hence:

Proposition 3 *For a d-note scale s in a n-TET, the determinant of the associated scale matrix S is bounded:*

$$|S| \le d\left(d\,\frac{n-d}{n-1}\right)^{\frac{n-1}{2}}$$

This bound is reached iff s is a FLID.

The Hadamard inequality, which is true for any matrix, would yield the larger bound $d^{n/2}$.

More generally, even when a FLID with d elements does not exist in $\mathbb{Z}_n$, the determinant of the scale matrix is maximum (for all d note scales) when the interval distribution is 'as flat as possible': the (absolute) value of the determinant can be chosen as a measure of the 'flatness' of the scale. For instance, for 5-note scales in 11-TET, the scale $\{1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0\}'$ gives the maximum determinant of 1215. For 7-note scales that are subsets of 12-TET, the upper theoretical bound is

---

[12]The product of positive real numbers whose sum is fixed is maximal when they are equal.

about 4072, but the actual maximum determinant is 2240, which occurs for scales such as $\{1, 1, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0\}'$, whose Fourier coefficient as close as possible to uniform[13] for binary-valued scale vectors. For 4-scale notes the flatest cases are the well known pc-sets $(0, 1, 4, 6)$ and $(0, 1, 3, 7)$ (though the tritone is in fact redoubled, hence these are not real FLIDs), with a determinant of 1024 for a theoretical maximum around 1421.

5 Non-invertible Representations

When the scale matrix S is singular, it is still possible to find combinations of elements in the range space of the columns that are close to a specified scale vector.

5.1 Rank Deficient Scale Matrices

One approach is to use the pseudo-inverse, which gives the closest L^2 approximation for underdetermined systems.

Example 10 (Minor Thirds) *Let s be the minor third scale vector (1). The corresponding circulant scale matrix S has determinant $|S| = 0$ and so is not invertible. Applying the pseudo-inverse T to the diatonic scale d in (2) expresses d as closely as possible in terms of minor thirds. The result is $Td = \frac{1}{8}\{2, 2, 5, -1, 3, 1, 2, 2, 1, 3, -1, 5\}$. As a check, one can calculate $STd = \{1, 0, \frac{5}{4}, 0, 1, \frac{3}{4}, 0, 1, \frac{1}{4}, 1, 0, \frac{3}{4}\}$. When rounded to the nearest integer, this is equal to d.*

This pseudo-inverse sometimes provides unsatisfying results – one might prefer sparser vectors for a more concise representation, as in Section 5.3.

5.2 Compound Spanning Sets

The representation of one scale x in the basis of another scale s can be complicated when there is no easy way to write x as a weighted sum of the columns of S. Thus it may be desirable to attempt to decompose x using a dictionary of different scale vectors. Formally, a *frame* is a set of *dictionary* elements that form an over-complete spanning set [3]. There are several methods for finding the "best" representation in this set including Huber's "projection pursuit" [10] and Mallot's "matching-pursuit" [14]. Perhaps the most straightforward method is Daubechie's "method of frames" [8] which calculates the pseudo-inverse of a (large rectangular) matrix composed of all the vectors in the dictionary.

---

[13]Namely $\{7, \sqrt{2}, 2, \sqrt{5}, 2, \sqrt{2}, 1, \sqrt{2}, 2, \sqrt{5}, 2, \sqrt{2}\}$.

For example, let s_1 be two scale vectors of length n with associated circulant scale matrices S_1 and S_2. The concatenated matrix $S = [S_1|S_2] \in \mathbb{R}^{n \times 2n}$ forms a dictionary of vectors from which the representation draws.

Example 11 (Major and Minor Thirds) *Let s_1 be the major third scale vector as in Example 1, s_2 the minor third scale vector (1), $S \in \mathbb{R}^{12 \times 24}$ the dictionary, T the pseudo-inverse of the S, and d diatonic scale vector (2). Then Td expresses d as closely as possible in terms of both minor and major thirds. The result is*

$$\frac{1}{720}\{371, 87, 67, -9, -13, 327, -29, 327, -13, -9, 67, 87, 61, 61, 293,$$
$$-139, 301, 13, 101, 101, 13, 301, -139, 293\}.$$

This is not a particularly simple or concise representation.

5.3 Linear programming

The problem with Example 11 is that the answer minimizes an L^2 error, distributing the energy among all the dictionary elements. The intent of a "simple" representation is better represented by a sparse solution; to achieve a decomposition of x that uses as few elements in the dictionary as possible. While finding the sparsest solution is difficult, it has recently been recognized that a linear programming L^1 solution can often give an answer that is close to sparse [5], a result sometimes called *compressive sensing*. Accordingly, it is worthwhile revisiting Example 11.

Example 12 (Major and Minor Thirds) *With s_1, s_2, d, and $S \in \mathbb{R}^{12 \times 24}$ as in Example 11, a linear programming problem can solve for the x that minimizes $\mathbf{1}'_{24}x$ subject to $Sx \geq d$ and $x \geq 0$. The answer is*

$$x = \frac{1}{2}\{1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1\},$$

which exactly represents the diatonic scale as the halved sum of three major thirds and four minor thirds. This is musically sensible, as these two kinds of thirds are precisely the only ones occuring in the diatonic.

This kind of optimization of a solution is also rewarding when the system is under-determined. Not only does it yield a solution (albeit an approximate one), but this solution tends to be sparse and can in some measure be controlled by choice of the linear form used in the minimization.

Example 13 (Perfect Fifths) *Let s be the perfect fifth scale vector, S the corresponding scale matrix, d the diatonic scale vector (2), and let $\mathbf{1}_n$ be the n-vector of all ones.*

The linear program seeks to find the x that minimizes $\mathbf{1}_n' x$ subject to $Sx \geq d$ and $x \geq 0$. The answer is $x = \{1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0\}'$, which approximates the diatonic scale as a sum of four perfect fifths. To see the quality of the approximation, calculate $Sx = \{2, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1\}'$, which is the same as d but for the first entry, but doubling the root note of the scale. Other choices of the direction of minimization (i.e. vectors other than $\mathbf{1}_n$) will give other solutions, for instance the other coverings of the diatonic with fifths. It is perhaps amusing that the natural choice of $\mathbf{1}_n$, privileging no particular pitch class, should give particular preeminence to the tonic of the Ionian mode.

6   Tilings

The special case of a decomposition of the chromatic aggregate $\mathbf{1}_n$ yields a tiling or mosaic when there is a solution with 0 or 1 coefficients. Each note in a tiling is covered by exactly one translate of the original scale. For instance, diminished seventh chords tile the aggregate:

$$(0, 3, 6, 9) \cup (1, 4, 7, 10) \cup (2, 5, 8, 11) = \mathbb{Z}_{12}. \tag{19}$$

which can be restated using scale vectors as

$$(1,0,0,1,0,0,1,0,0,1,0,0)' + (0,0,1,0,0,1,0,0,1,0,0,1)' + (0,1,0,0,1,0,0,1,0,0,1,0)' = \mathbf{1}_{12}.$$

or more traditionally

$$(0, 3, 6, 9) \oplus (0, 1, 2) = \mathbb{Z}_{12}.$$

Such mosaics have been of particular interest in the domain of rhythm, where rhythmic canons have been thoroughly investigated (see [1, 2]). It is already hard to know *whether* a particular 'tile,' i.e. subset of $\mathbb{Z}_n$, does indeed tile, though some results are known in that direction; but computationally, when n gets large it is extremely difficult to find *how* it tiles, i.e. which translates of the tile must be used in order to realize the tiling.

In terms of scale matrices, the question is to factor

$$Sx = \mathbf{1}_n \tag{20}$$

where x is the unknown disposition of the translates of tile s with associated matrix S.

In order to tile, S will never be invertible (indeed it is necessary that all columns of Ω be either in the null space of S or the nullspace of X, the scale matrix of x), so there is no unique solution to this problem.

However, linear programming can often provide a solution, and Mikhalis Kolountza-kis has conjectured [12] that this method always will.[14] For instance with

$$s = (0, 3, 9, 33, 40, 48, 75, 81, 88, 96, 105, 136)$$

the linear program provides in a fraction of a second

$$x = (0, 2, 4, 18, 20, 22, 36, 38, 54, 112, 128, 130)$$

for a tiling of $\mathbb{Z}_{144} = x \oplus s$. This type of tiling where neither of the factors exhibit any periodicity is called a Vuza Canon [2].

7 Discussion and Conclusion

This paper has investigated representations of musical scales and rhythms as binary vectors corresponding to periodic structures. Scale vectors can be decomposed in bases that represent musically sensible scales and chords (or simple rhythmic patterns) using scale matrices, and the resulting decompositions can be musically meaningful, as for instance the representation of a pentatonic scale in terms of the diatonic or a representation of the clave rhythm in a basis defined by the tango.

Several properties of scale matrices are studied including the relationship with the familiar DFT basis and the correspondence between the invertibility of the scale matrix and Lewin's special cases. The Z-relation (homometry) can be expressed directly in terms of scale matrices, and flat interval distributions (FLIDS) can be studied in terms of the determinant.

When the scale matrices are not invertible, sensible decompositions can be (sometimes) be made using a linear programming technique, which can also be applied when the matrix is non-square. In this case, the linear program solves an $L1$ minimization problem that is often associated with sparse (concise) solutions. This approach may also be useful in the study of rhythmic canons (tilings).

Acknowledgements

---

[14]Specifically, that minimizing the linear form $\mathbf{1}'_n x = \sum_{k=0}^{n-1} x_k$ on the polyhedron defined by $\forall k, x_k \geq 0$ and $\sum_{j=0}^{n-1} s_{k,j} x_j = 1$ by the simplex method yields a vector x with 0-1 coefficients satisfying eq. 20.

8 Proofs

8.1 Proof of Thm. 1

The scale matrix S of the dyad $(0, d)$ is equal to identity + the matrix D of the permutation $i \mapsto i + d \mod n$. Hence the kernel (or nullspace) of S is the eigenspace of D for eigenvalue -1. Let us reason geometrically, considering the vectors $e_0 \ldots e_{n-1}$ of the canonical basis of $\mathbb{R}^n$. A vector $x = \sum_{i=0}^{n-1} x_i e_i$ lies in this eigenspace iff

$$\sum_{i=0}^{n-1} x_i e_i = \sum_{i=0}^{n-1} -x_i e_{i+d} \iff \forall i = 0 \ldots n-1 \quad x_{i+d} = -x_i \tag{21}$$

(all indexes are computed modulo n).

From this we get $x_{i+k\,d \mod n} = (-1)^k x_i$. Hence $x_{i+n\,d \mod n} = x_i = (-1)^n x_i$: if n is odd then the only solution is $x = 0$, i.e. S is invertible.

Say now that n is even, $n = 2^v n'$ where n' is odd. Let k be the smallest integer such that $k\,d = 0 \mod n$, e.g. $k = n/(d, n)$.[15] If k is odd we have the same impossibility, and S is invertible. This is the case whenever d harbours the same power of 2 as n (or a larger one): we have proved that if 2^v divides d then S is invertible.

Assume now that 2^v does not divide d, i.e. $d = 2^u d'$ with d' odd and $u < v$. We can produce the eigenvectors, i.e. elements of the kernel of S, in the following way:

- Fix one coordinate – say $x_0 = 1$.

- From Equation (21), $x_d = x_{0+d} = -x_0 = -1$.

- Iterate until back to x_0: $x_{2d} = +1, x_{3d} = -1, \ldots x_0 = x_n = x_{n/(d,n) \times d} = +1$. The last value is indeed +1 because $n/(d, n)$ is an even number.

So the value of one coordinate determines the value of $n/(d, n)$ coordinates. We have thus $n/\big(n/(d, n)\big) = (d, n)$ arbitrary coordinates $x_0, x_1 \ldots x_{(d,n)-1}$, that is to say (d, n) degrees of freedom, and hence the dimension of the kernel of S is exactly (d, n). Its largest possible value (apart from $d = 0$ which is no more a dyad) is for $d = n/2$. In general, we get the rank of matrix S by way of the rank-nullity theorem: $\text{rank}(S) = n - (d, n)$, remembering though that $\text{rank}(S) = n$ when 2^v divides d.

---

[15]We denote here the gcd of d, n by (d, n) for concision.

8.2 Proof of Thm. 2

By assumption, $Sz = x$, and it is desired to find an expression for z_2 where $Sz_2 = \mathbf{1}_n - x$. This can be rewritten

$$
\begin{aligned}
\mathbf{1}_n - x &= \mathbf{1}_n - Sz \\
&= \frac{1}{m}S\mathbf{1}_n - Sz \\
&= S(\frac{1}{m}\mathbf{1}_n - z)
\end{aligned}
$$

where m is the number of nonzero coefficients in x. The second line follows because $\mathbf{1}_n$ is an eigenvector of S, that is, $S\mathbf{1}_n = m\mathbf{1}_n$.

8.3 Proof of Thm. 3

The scale vector $\mathbf{1}_n - x$ can be expressed

$$
\begin{aligned}
\mathbf{1}_n - x &= \frac{1}{m}S_x\mathbf{1}_n - S_x e_1 \\
&= S_x\left(\frac{1}{m}\mathbf{1}_n - e_1\right) \equiv S_x z
\end{aligned}
$$

The first equality uses the fact that $S_x e_1 = x$ and that $\mathbf{1}_n$ is an eigenvector of S_x, that is, $S_x \mathbf{1}_n = m\mathbf{1}_n$. Similarly, $S_{\mathbf{1}_n - x}(\frac{1}{n-m}\mathbf{1}_n - e_1) = x$.

8.4 Proof of Thm. 4

One way to compute IFunc (A, B) is to calculate AB' and extract the first column. Indeed, if A and B are circulating matrices of scales a and b, then so is $C = AB'$,[16] and the kth element of the first (or rather, 0th) column of C is

$$
\sum_{j=0}^{n-1} A'_{k,j}B_{j,0} = \sum_{j=0}^{n-1} A_{j,k}B_{j,0} = \sum_{j=0}^{n-1}
\begin{cases}
1 & \text{when } j - k \in a \text{ and } j - 0 \in b \\
0 & \text{else}
\end{cases}
$$

and the term in the sum in non zero only if k is the distance between some element in a and another element in b: we recognize IFunc(A, B). Since C is a circulating matrix, the other columns are defined by this first column.

Another way makes use of special features of the matrices, namely that $\overline{B} = B, \overline{\Omega} = \Omega'$ and for any circulating matrix S, the matrix with the Fourier coefficients

---

[16]Because the algebra of circulating matrices is stable under $\times$ and also under transposition of matrices, as its generating element J satisifies $J' = J^{-1} = J^{n-1}$.

$\Omega^{-1}S\,\Omega$ is diagonal. Hence, if we denote by $\mathcal{F}_A$ (resp. $\mathcal{F}_B$) the diagonal matrix of the Fourier coefficients of a (resp. b):

$$AB' = A\overline{B'} = \Omega\mathcal{F}_A\Omega^{-1}\overline{\Omega.\mathcal{F}_B\Omega^{-1}}' = \Omega\mathcal{F}_A\Omega^{-1}\overline{\Omega^{-1'}}\overline{\mathcal{F}_B'}\overline{\Omega'} = \Omega\mathcal{F}_A\overline{\mathcal{F}_B}\Omega^{-1}$$

and we recognize the inverse Fourier Transform of $\mathcal{F}_A\overline{\mathcal{F}_B}$, i.e. IFunc (A,B).

8.5 Proof of Thm. 5

Matrix U has rational elements and its eigenvalues are of magnitude 1. From a well-known lemma in linear algebra,[17] these eigenvalues are roots of unity. Moreover, as U is a polynomial $U = P(J), P \in \mathbb{Q}[X]$ in the matrix J, whose eigenvalues are the nth roots of unity, the eigenvalues of U are polynomials $P(e^{2ik\pi/n})$ in these roots, hence lie in the field $\mathbb{Q}[e^{2i\pi/n}] = \mathbb{Q}_n$, often called *cyclotomic field*.

Let ξ be any eigenvalue of U. We know that $\xi \in \mathbb{Q}[e^{2i\pi/n}]$ and that $\xi^m = 1$ for some integer m (chosen minimal). Hence as ξ generates $\mathbb{Q}_m$, $\mathbb{Q}_m \subset \mathbb{Q}_n$. The set M of possible values for m has a maximum element for the order of division, because if ξ (resp. ξ') with order m (resp. m') is in the field, then so is $\xi^p\xi'^q$ that can be chosen with order $\mathrm{lcm}(m,m')$.

As $n \in M$, the maximum value for m is a multiple of n. Let us call m this maximal value, any element ξ of $\mathbb{Q}_n$ lying on the unit circle must satisfy $\xi^m = 1$. Because again $n \in M$, m is a multiple of n and $\mathbb{Q}_n \subset \mathbb{Q}_m$. Finally $\mathbb{Q}_n = \mathbb{Q}_m$.

Letting $\Phi(n) = \dim[\mathbb{Q}_n \mid \mathbb{Q}]$ be Euler's totient function, $\begin{cases} n \mid m \text{ and} \\ \Phi(n) = \Phi(m) \end{cases}$.

As $\Phi(n) = n \prod_{p\mid n; p \text{ prime}} \left(1 - \dfrac{1}{p}\right)$, the only possibility is that $m = \begin{cases} n & \text{for } n \text{ even} \\ 2n & \text{for } n \text{ odd} \end{cases}$.

This proves the theorem: all eigenvalues of U are n or $2n^{th}$ roots of unity.

Hence, U could be constructed as a polynomial in the elementary circulating matrix J (as all other circulating matrices) $U = P(J)$ where P is the interpolating polynomial that sends the Fourier coefficients of J, i.e. the $e^{2ik\pi/n}$, to the Fourier coefficients chosen for U. Such a construction is easy with the Lagrange polynomials associated with the $e^{2ik\pi/n}$, as P is a linear combination of these polynomials with coefficients that are precisely the Fourier coefficients of the desired u.

The possibilities of mapping the nth roots of 1 to m^{th} roots of 1 can be somewhat reduced by noticing that U is a rational polynomial in J, and such a polynomial is stable under all field automorphisms of $\mathbb{Q}_n$. Perhaps this is a good starting point to try and answer Rosenblatt's question [17] as to the structure of the group of spectral units.

---

[17]Algebraic numbers who are of magnitude ≤ 1 together with all their conjugates are roots of unity.

8.6 Proof of Thm. 6

We start from some simple observations, denoting by S the scale matrix of $A \subset \mathbb{Z}_n$:

- The scale matrix of the inverse (i.e. $n - A$) is the transpose of the scale matrix S;

- The scale matrix (i.e., the circulant matrix) IFunc (A, A) is $S'S = SS'$;

- Recall that for any scale $s = 1_A$,

$$F_A = \Omega^{-1} S \Omega = \begin{pmatrix} \mathcal{F}_A(0) & 0 & \cdots & \cdots & 0 \\ 0 & \mathcal{F}_A(1) & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & & \cdots & 0 & \mathcal{F}_A(n-1) \end{pmatrix},$$

where $\mathcal{F}_A(k)$ are the Fourier coefficients of A.

- The Fourier matrix Ω is unitary: $\overline{\Omega'} = \Omega^{-1}$.

We prove the equivalence, denoting by F_A the diagonal matrix with coefficients the values of $\mathcal{F}_A$ above, and $\mathbf{1}_n \mathbf{1}_n'$ the square matrix with all values equal to one. A is a FLID iff the matrix of its interval content is constant (except the diagonal):

$$Mat(IC_A) = \begin{pmatrix} d & m & \cdots & m \\ m & d & & \vdots \\ \vdots & \ddots & \ddots & \\ m & \cdots & m & d \end{pmatrix} = (d - m)\, I + m \mathbf{1}_n \mathbf{1}_n'$$

$$\iff S' \times S = (d - m)I + m \mathbf{1}_n \mathbf{1}_n'$$

$$\iff \overline{S'} \times S = (d - m)I + m \mathbf{1}_n \mathbf{1}_n' \quad \text{(because } S \text{ is real valued)}$$

$$\iff \left(\overline{\Omega^{-1} F_A \Omega} \right)' \times \Omega^{-1} F_A \Omega = (d - m)I + m \mathbf{1}_n \mathbf{1}_n'$$

$$\iff \Omega^{-1} \overline{F_A} \Omega \times \Omega^{-1} F_A \Omega = \Omega^{-1} \overline{F_A} F_A \Omega = (d - m)I + m \mathbf{1}_n \mathbf{1}_n' \quad \text{(because } \Omega \text{ is unitary)}$$

$$\iff \overline{F_A} F_A = \Omega \big((d - m)I + m \mathbf{1}_n \mathbf{1}_n' \big) \Omega^{-1}$$

and as

$$\Omega \mathbf{1}_n \mathbf{1}_n' \Omega^{-1} = \overline{(\Omega \mathbf{1}_n \mathbf{1}_n' \Omega^{-1})'} = \begin{pmatrix} n & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & 0. \end{pmatrix} = F_{\mathbf{1}_n}$$

where $n\delta_0$ is the Fourier transform of $\mathbf{1}_n$. We find thus

$$\begin{pmatrix} |\mathcal{F}_A(0)|^2 & 0 & \cdots & 0 \\ 0 & |\mathcal{F}_A(1)|^2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & |\mathcal{F}_A(n-1)|^2 \end{pmatrix} = (d - m)I + m \mathcal{F}_{\mathbf{1}_n}$$

$$= \begin{pmatrix} d + (n-1)m & 0 & \cdots & 0 \\ 0 & d - m & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & d - m \end{pmatrix},$$

24

meaning that $|\mathcal{F}_A(k)|$ is constant for $k \neq 0$, i.e. that $|\mathcal{F}_A|$ is flat. Observe that the modulus of the Fourier coefficients are always square roots of integers.

References

[1] Amiot, E., *À propos des canons rythmiques*, Gazette des Mathmaticiens, 106 (2005).

[2] Amiot, E., *New Perspectives on rhythmic canons and the Spectral Conjecture*, JMM (2009), special issue on rhythmic canons, C. Agon et M. Andreatta dir.

[3] C. S. Burrus, R. A. Gopinath, and H. Guo, *Wavelets and Wavelet Transforms,* Prentice Hall, NJ (1998).

[4] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies" *IEEE Trans. Inform. Theory*, 52 5406-5425 (2004).

[5] E. J. Candès, M. Wakin and S. Boyd, "Enhancing sparsity by reweighted l1 minimization," *J. Fourier Anal. Appl.*, 14 877-905 (2007).

[6] N. Carey, *Coherence and sameness in well-formed and pairwise well-formed scales*, J. Math. Music 1, (2006), pp. 79–98.

[7] N. Carey, and D. Clampitt, *Aspects of well-formed scales*, Music Theory Spectrum 11, (1989), pp. 187–206.

[8] I. Daubechies, "Time-frequency localization operators: A geometric phase space approach," *IEEE Trans. Information Theory,* Vol. 34, No. 4, 605-612 (July 1988).

[9] R. M. Gray *Toeplitz and Circulant Matrices: A review*, Now (2002).

[10] P. J. Huber, "Projection pursuit," *Annals of Statistics,* Vol. 13, No. 2, 435-475 (1985).

[11] A. King, "Employments of the "standard pattern" in Yoruba music." *African Music*, 2 **3** (1961).

[12] M. Kolountzakis, private communication.

[13] D. Lewin, "Intervallic Relations between two collections of notes." *Journal of Music Theory*, **3:** 298-301 (1959).

[14] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing,* Vol. 41, No. 12 (Dec. 1993).

[15] A.J. Milne, W.A. Sethares, and J. Plamondon, *Isomorphic controllers and dynamic tuning: Invariant fingering over a tuning continuum,* Computer Music J. 31 (2007), 15–32.

[16] I. Quinn, Perspect. New Mus. 39, 108 (2001).

[17] Rosenblatt, J., *Phase Retrieval*, Commun. Math. Phys. 95, 317-343 (1984)

[18] W.A. Sethares, *Tuning, Timbre, Spectrum, Scale*, Springer-Verlag, London, 2nd edn (2004).

[19] W.A. Sethares, *Rhythm and Transforms*, Springer-Verlag, London (2007).

[20] G. Strang and T. Nguyen, *Wavelets and Filter Banks,* Wellesley College Press, Wellesley, MA (1996).

[21] http://homepages.cae.wisc.edu/∼sethares/AlgofScales.m

[22] Wild, J., unpublished.