

On Musical Scale Rationalization

Albert Gräf

Dept. of Music Informatics,
Johannes Gutenberg University Mainz, Germany
Dr.Graef@t-online.de

Abstract

Already the ancient Greek philosophers discovered that the consonance of musical intervals is related to the complexity of the corresponding frequency ratios. Modern psychoacoustic and neurological research provides us with explanations why simple frequency ratios might be important for our perception of intervallic relationships, even though many tunings only approximate these “just” intervals. By rationalizing a (non-rational) musical scale the harmonic relationships can be made explicit and can then be used in various applications ranging from algorithmic composition to automatic transcription. The first general scale rationalization algorithm based on a certain harmonicity measure and a complete enumeration approach was proposed by the composer Clarence Barlow. This paper introduces the notion of “sub-additive” consonance measures to give a general formulation of the problem, and proves that the problem is NP-complete. The paper also describes a new algorithm for scale rationalization, based on clique search in graphs, and points out possible applications such as the graphical visualization of scales and algorithmic composition.

1 Motivation and background

This paper discusses one of the fundamental problems in the mathematical theory of music which has been with us at least since the time of the ancient Greek philosophers. According to the written records, some 2500 years ago Pythagoras was the first to point out the importance of simple frequency ratios in the definition of the major musical intervals such as the octave (2/1), the fifth (3/2) and the fourth (4/3). This work formed the basis of many later developments of musical scales and tunings during the medieval times and the modern age until today. It has also been observed that, at least for the kinds of harmonic sounds produced by instruments commonly used in Western music, the “pleasantness” or “consonance” of an interval is somehow correlated with the complexity of its frequency ratio.

While this observation is not undisputed and is still the subject of investigations, psychoacoustic and neurological research have provided us with possible explanations for this phenomenon. Helmholtz proposed that the consonance of two complex tones depends on the beats between adjacent overtones of the two harmonic series (Helmholtz 1983). The perceived roughness of adjacent sine tones was later studied more thoroughly and explained with the concept of the critical band (Plomp and Levelt 1965). By summing up the pairwise dissonances of all involved partials one obtains a dissonance function for pairs of complex tones which, for harmonic sounds, attains its local minima at simple fundamental frequency ratios (Sethares 2005). Neurological research suggests another possible explanation, namely that the processing of sound in early stages of the auditory signal path involves nonlinear resonances leading to phase-locked activation patterns of neural oscillators, a process which tends to map frequency relationships in the auditorial input to nearby simple integer ratios (Large and Tretakis 2005).

This gives rise to an interesting mathematical problem: What exactly is a “simple” ratio, and is it possible to “rationalize” a given sequence of (possibly irrational) pitches in such a manner that only simple integer ratios are used? An abundance of different “consonance” or “harmonicity” measures have been proposed to answer the first question, such as Euler’s *gradus suavitatis* (Euler 1739), Helmholtz’s harmoniousness measure (Helmholtz 1983), James Tenney’s harmonic distance (Chalmers 1993) and Clarence Barlow’s harmonicity function (Barlow 2001b). However, to our knowledge, Barlow was the first to also point out an algorithm to rationalize arbitrary (non-rational) scales using complete enumeration of possible solutions derived from a precomputed set of possible tuning alternatives.

Our contribution in this paper is that we introduce the notion of “subadditive” harmonicity measures which allows us to treat many of the existing consonance measures in a uniform setting. We then go on to prove that a certain formulation of the scale rationalization problem is NP-complete and thus brute force algorithms like the one by Barlow are probably the best we can do if we want to find an *optimal* solu-

tion. However, if we are satisfied with “good,” though not necessarily optimal, rationalizations, then heuristic solution approaches become applicable. To these ends, we show how the scale rationalization problem relates to the clique problem on combinatorial graphs, and employ this relationship to derive a heuristic rationalization algorithm, which often uses much less computing time than the brute force method and appears to work fairly well in practice. We then sketch out two applications of these methods in some detail: graphic visualization of scales using multidimensional scaling, and employing the harmonicities computed from rationalized scales in algorithmic composition. In the conclusion we also point out a few other potential applications.

2 Harmonicity

The term “consonance” is somewhat troublesome, since it is used to denote several different phenomena in music theory and psychoacoustics (Tenney 1988). Hence we adopt Barlow’s terminology and use the artificial terms *harmonicity* and *disharmonicity* throughout this paper.

Barlow’s harmonicities, as well as Euler’s gradus suavitatis and several other consonance measures are defined in terms of sums of elementary disharmonicities (called “indigestibilities” by Barlow) of the prime factors of a rational interval. In the following we give a generic definition which covers all these approaches. So let the *prime disharmonicities* $g(p) > 0$ for all prime numbers p be given. (We assume throughout that $g(p)$ can be computed in polynomial time with respect to $\log p$.) We extend g to all positive rational numbers x as follows. If $x = p_1^{a_1} p_2^{a_2} \cdots p_n^{a_n} > 0$ is a rational number, where $p_1 < p_2 < \cdots < p_n$ are the prime factors of x and a_1, a_2, \dots, a_n their (positive or negative) multiplicities, then the *disharmonicity* $g(x)$ of x (with respect to the given prime disharmonicities) is defined as:

$$g(x) = |a_1|g(p_1) + |a_2|g(p_2) + \cdots + |a_n|g(p_n).$$

The *harmonicity* $h(x)$ of x can then be defined as the inverse of its disharmonicity, that is,

$$h(x) = 1/g(x).$$

The generic disharmonicity function defined above has some nice mathematical properties. In particular, note that g is *additive* for integer arguments, i.e., $g(xy) = g(x) + g(y)$ if x and y are positive integers. Moreover, the disharmonicity of an interval x/y given in its cancelled-down form (i.e., x and y integer, $\gcd(x, y) = 1$), is always the sum of the disharmonicities of the numerator and the denominator:

$$g(x/y) = g(x) + g(y) = g(xy).$$

Interval	Ratio	Cents	g_B	g_E
unison	1/1	0.00	0.00	0
minor semitone	16/15	111.73	13.07	10
minor whole tone	10/9	182.40	12.73	9
major whole tone	9/8	203.91	8.33	7
minor third	6/5	315.64	10.07	7
major third	5/4	386.31	8.40	6
perfect fourth	4/3	498.04	4.67	4
tritone	45/32	590.22	16.73	13
perfect fifth	3/2	701.96	3.67	3
minor sixth	8/5	813.69	9.40	7
major sixth	5/3	884.36	9.07	6
minor seventh	16/9	996.09	9.33	8
major seventh	15/8	1088.27	12.07	9
octave	2/1	1200.00	1.00	1

Figure 1: Barlow and Euler disharmonicities for some common intervals.

It is also easy to see that in the general case, for arbitrary *rational* values $x, y > 0$, the disharmonicity function is *sub-additive*:

$$g(xy) \leq g(x) + g(y).$$

Using the generic definition above, we can obtain various different harmony measures such as Barlow’s disharmonicities and Euler’s gradus function by just plugging in suitable prime disharmonicities. For instance:

- Barlow disharmonicities: $g_B(p) = 2(p-1)^2/p$
- Euler disharmonicities: $g_E(p) = p-1$

The Barlow and Euler disharmonicities for some common intervals are shown in Fig. 1.¹ Note that the factor 2 in the definition of the Barlow disharmonicities is just a normalization factor which makes the octave 2/1 have a value of 1. We also remark that Euler’s original definition of the *gradus suavitatis* actually adds an extra 1 term after summing up the prime disharmonicities, i.e., if we denote Euler’s gradus function as Γ then $\Gamma(x) = g_E(x) + 1$. But in the following we stick to our definition because it makes the measure subadditive which is not the case for the original gradus function.

3 Harmonic distance

Next we show how to derive a scale metric from a disharmonicity function. For our purposes a *scale* is simply a finite set S of positive real numbers. The members of S are called *tones* or *pitches* which are either specified as frequency ratios

¹As usual, the “cent” values in the figure are in units of a 1/100 semitone, i.e., a value of 100 Cents corresponds to the frequency ratio $2^{1/12}$. More precisely, to obtain the cent values from the frequency ratios, one just takes the base 2 logarithm and multiplies with 1200.

or Cent values, relative to a given base tone which is usually left unspecified. In this section we only consider *rational* scales, i.e., scales whose pitches are all rational. Given two scale tones $x, y \in S$, we define the *harmonic distance* $d(x, y)$ as the disharmonicity of the interval between the pitches, i.e.,

$$d(x, y) = g(x/y).$$

Note that to calculate this value we have to cancel common factors in x and y . For instance, if $x = 3/2$ (the fifth over the base tone) and $y = 5/4$ (the major third) then

$$d(x, y) = g(3/2 \times 4/5) = g(6/5),$$

which, not very surprisingly, is the disharmonicity of the minor third. In this manner we can calculate harmonic distances between all pairs of scale members.

Now it is important to note that, for *any* choice of positive prime disharmonicities, the harmonic distance function d thus defined is indeed a *metric* in the mathematical sense. That is, it obeys the following rules:

- $d(x, y) = 0$ if and only if $x = y$, for all $x, y \in S$
- $d(x, y) = d(y, x)$, for all $x, y \in S$
- $d(x, z) \leq d(x, y) + d(y, z)$, for all $x, y, z \in S$

The first two conditions are easily verified. The third condition, which is also known as the *triangle inequality*, follows from the subadditivity of the disharmonicity function.

4 The scale rationalization problem

In the following we generally assume a disharmonicity function g and the corresponding harmonic distance metric d as defined in the preceding section. We now take a closer look at the scale rationalization problem. Generally speaking, we are given an arbitrary (not necessarily rational) scale S and ask how we can transform S into a “similar” rational scale S' in which the intervals between scale pitches are as “simple” as possible. Of course, the devil lies in the details, and so we now have to specify what exactly we mean by “similar” and “simple”.

Obviously, in order to get a scale S' which sounds similar to the original scale S , we have to match the pitches of S as closely as possible. But since there are infinitely many rational numbers in even the smallest positive range around a real value we need a secondary criterion which enables us to choose the “best” among those. This can be accomplished with Barlow’s method (Barlow 2001b) which always selects a given number of alternatives x' for a given scale tone $x \in S$

with a given minimum harmonicity, $h(x') \geq h_{\min}$, and the highest *weighted* harmonicity

$$h_w(x') = w(x')h(x'),$$

where w is a weight function which weighs pitches according to their offsets from the original scale tone. Barlow defines the weights as

$$w(x') = \exp(-\Delta(x')^2 \ln(1/a)/t^2),$$

where $\Delta(x') = 1200|\log_2(x'/x)|$ is the absolute offset of x' from x in cents, t is the tolerance in cents and a is the desired *attenuation* factor which determines the weight of the harmonicity values at the edges of the tolerance range. Thus the harmonicity values are modified by superimposing the usual kind of bell-shaped Gaussian curve, which has the effect that those pitches will be preferred which either have a high harmonicity or are close to the original scale tone. We refer the reader to the cited reference (Barlow 2001b) for an example.

So in the following we may assume that a (nonempty, finite) set of alternative rational “candidate” pitches $C(x)$ is given for each $x \in S$. Next we have to decide which combination of candidate pitches is considered to form a “simple” scale. For this purpose we employ the harmonic distance metric. We will be interested in solutions in which the harmonic distances for all pairs of rational pitches are as small as possible. To accomplish this, we could start out with a global optimization method, as proposed by Barlow, which attempts to minimize the total (or average) harmonic distance of all intervals in the resulting scale. However, this approach effectively lumps all intervals together, which is often a bad idea because some intervals will be inherently more disharmonious than others. For instance, an interval of about 600 cent (a tritone) will almost surely have a higher disharmonicity than an interval of about 700 cent (a fifth), in *any* reasonable rationalization of the scale. Therefore in the following we consider a refinement of Barlow’s approach, in which different disharmonicity bounds can be specified for the individual intervals between scale members.

Problem 1 (Scale rationalization) INSTANCE: A collection of pitch candidate sets C_i , $i = 1, \dots, n$, and disharmonicity bounds b_{ij} for $i, j = 1, \dots, n$, $i \neq j$. **QUESTION:** Are there pitches $x_i \in C_i$ for $i = 1, \dots, n$ such that $d(x_i, x_j) \leq b_{ij}$ for $i, j = 1, \dots, n$, $i \neq j$?

Note that we have stated the problem in the form of a decision problem, namely the problem to decide whether *any*

²In the following we generally assume that the disharmonicity bounds are *symmetric*, i.e., $b_{ij} = b_{ji}$ for all $i \neq j$. We can do so without loss of generality since d is a metric and hence $d(x_i, x_j) = d(x_j, x_i)$ for all i, j .

rationalization exists which satisfies the given disharmonic-ity bounds. Of course, if the answer is affirmative then we will still be interested in obtaining the “best” such solution according to certain optimization criteria; we return to this question in Section 6.

But first let us show that the scale rationalization problem is indeed “difficult” in a precise mathematical sense: the problem is *NP-complete*. The class of NP-complete problems comprises many important practical problems which do not appear to have an “efficient”, i.e., polynomial-time solution. For the definition and the many ramifications of this concept we refer the reader to Garey and Johnson’s classic book on the subject (Garey and Johnson 1979). In particular, showing that the scale rationalization problem is NP-complete gives us a strong indication that only the (computationally expensive) complete enumeration of all possible solutions will ensure that we find an optimal rationalization. One way to deal with this situation is to explore alternative heuristic solution approaches which might give us a good, though not optimal solution in a reasonable amount of time, a path that we will pursue in Section 6.

In fact, we can prove that already a fairly restricted version of the scale rationalization problem is NP-complete:

Theorem 1 *The scale rationalization problem is NP-complete, for any given harmonic distance metric d , even if the disharmonic-ity bounds are all the same and each candidate set contains at most three pitches.*

For the proof of the above theorem we employ a reduction from the following 3-satisfiability problem (“3SAT”) of propositional logic; see (Garey and Johnson 1979, p. 46). 3SAT is NP-complete, and hence a polynomial-time reduction from 3SAT to Problem 1 shows that the latter problem is NP-complete as well.

Problem 2 (3SAT) INSTANCE: Set U of variables, and a set K of clauses where each clause consists of three literals of the form u or \bar{u} , $u \in U$. QUESTION: Is there a truth assignment $T : U \mapsto \{\text{true}, \text{false}\}$ which satisfies all clauses?³

Proof of Theorem 1. Problem 1 clearly belongs to NP since we can verify in polynomial time whether a given selection $x_i \in C_i$, $i = 1, \dots, n$, satisfies the disharmonic-ity bounds. We reduce 3SAT to the scale rationalization problem. Let an instance $U = \{u_1, \dots, u_m\}$, $K = \{K_1, \dots, K_n\}$ of 3SAT be given. We show how to transform this instance into a corresponding instance of the scale rationalization problem which has a solution if and only if the original 3SAT instance

³Note that in order for all clauses to be satisfied, each clause must contain a literal v such that $T(v) = \text{true}$, where $T(v) = T(u)$ if $v = u$ and $T(v) = \neg T(u)$ if $v = \bar{u}$.

has one. We assume without loss of generality that $m \geq 3$ (otherwise our instance of 3SAT can be solved in linear time, by simply enumerating all truth assignments).

We first choose mutually distinct prime numbers p_1, p_2, \dots, p_m . (This can be done in polynomial time with respect to m .) Now for $j = 1, \dots, m$ let

$$x_j = p_1 \cdots p_m / p_j^2.$$

Then for each $j, j_1, j_2 \in \{1, \dots, m\}$, $j_1 \neq j_2$, we have:

$$\begin{aligned} g(x_j^2) &= 2 \sum_{j=1}^m g(p_j) & g(x_j/x_j) &= 0 \\ g(x_{j_1}x_{j_2}) &= 2 \sum_{j \neq j_1, j_2} g(p_j) & g(x_{j_1}/x_{j_2}) &= 2g(p_{j_1}) + 2g(p_{j_2}) \end{aligned}$$

Thus, since $m \geq 3$, it is possible to choose a bound B such that $g(x_{j_1}x_{j_2}) > B$ if $j_1 = j_2$ and $g(x_{j_1}x_{j_2}), g(x_{j_1}/x_{j_2}) \leq B$ otherwise. We now construct an instance of the scale rationalization problem as follows. For each clause K_i let C_i be the set of pitches x_j for which $u_j \in K_i$ and pitches $1/x_j$ for which $\bar{u}_j \in K_i$. Furthermore, let $b_{i_1 i_2} = B$ for all $i_1, i_2 \in \{1, \dots, n\}$, $i_1 \neq i_2$. Now it is easy to verify that $S = \{y_1, \dots, y_n\}$, $y_i \in C_i$, is a solution for the scale rationalization instance if and only if T is a satisfying truth assignment for the 3SAT instance, where $T(u_j) = \text{true}$ iff $x_j \in S$, $j = 1, \dots, m$. Conversely, if T is a satisfying truth assignment then we can pick a literal $v_{j_i} \in K_i \cap \{u_{j_i}, \bar{u}_{j_i}\}$ such that $T(v_{j_i}) = \text{true}$ for $i = 1, \dots, n$. We then obtain a solution $S = \{y_1, \dots, y_n\}$ for the scale rationalization instance, where $y_i = x_{j_i}$ if $v_{j_i} = u_{j_i}$ and $y_i = 1/x_{j_i}$ otherwise. \square

5 Scale rationalization and the clique problem

Let us now see how the scale rationalization problem relates to the clique problem on graphs. First we recall some terminology: A *graph* is a pair $G = (V, E)$ where V is a finite set, the set of *nodes* or *vertices* of the graph, and E is a set of unordered pairs of nodes vw , the *edges* of G . Two nodes v and w connected by an edge vw are called *adjacent*. In this paper, all graphs are *simple* (they do not contain multiple edges between the same pair of nodes), and *loopless* (they do not contain edges connecting a node to itself). A *clique* of a graph is a subset of nodes $U \subseteq V$ such that every two distinct nodes $u, v \in U$ are connected by an edge uv . The *clique problem* can be stated as follows:

Problem 3 (Clique) INSTANCE: Graph $G = (V, E)$, $k \leq |V|$. QUESTION: Does G contain a clique of size k ?

The clique problem is also NP-complete, but there are some algorithms with which moderately sized instances of the problem can be solved fairly well in practice. To see how scale rationalization can be formulated as a clique problem we need the following definition.

Definition 1 (Harmonicity graph) *Let pitch candidate sets C_i , $i = 1, \dots, n$ and disharmonicity bounds b_{ij} for $i, j = 1, \dots, n$, $i \neq j$ be given. Then the harmonicity graph $G(C, b)$ has $|C_1| + \dots + |C_n|$ nodes v_i^x , $i = 1, \dots, n$, $x \in C_i$, and the edges $v_i^x v_j^y$ for which $d(x, y) \leq b_{ij}$, $i, j \in \{1, \dots, n\}$, $i \neq j$, $x \in C_i$, $y \in C_j$.*

We have defined the harmonicity graph in such a manner that $\{x_1, \dots, x_n\}$ is a solution to the scale rationalization problem if and only if $\{v_1^{x_1}, \dots, v_n^{x_n}\}$ is a clique of the harmonicity graph. Hence:

Theorem 2 *An instance of the scale rationalization problem is solvable if and only if the corresponding harmonicity graph has a clique of size n .*

Consider Fig. 2 for a simple example. It contains a single candidate pitch (1/1) for the unison (0 Cent), and three pitch candidates for both the minor third (300 Cent) and the major third (400 Cent). The given edges in the graph are for a (global) maximum disharmonicity bound of 25; the actual (Barlow) disharmonicity below this threshold are printed on the edges. The eligible rationalizations under these conditions are given by the different triangles (a.k.a. 3-cliques) in the graph. Note that one of these actually corresponds to the standard “just” tuning $\{1/1, 6/5, 5/4\}$.

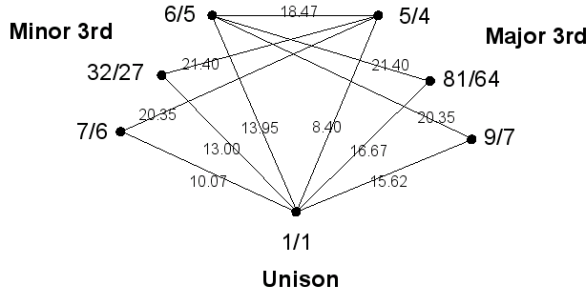


Figure 2: Harmonicity graph.

In this example it is easy to just read off the possible rationalizations from the graph, but of course this becomes much harder when the harmonicity graph is larger. Fortunately, the clique problem has been studied extensively and algorithmic methods for computing cliques are readily available; we discuss one such algorithm in the following section.

6 A scale rationalization algorithm

Using the relationship between scale rationalization and the clique problem established in the previous section, we now show how to solve the rationalization problem with Carraghan and Pardalos’ branch and bound procedure for enumerating cliques in a graph (Carraghan and Pardalos 1990). The basic algorithm is shown in Fig. 3. The algorithm starts out with an initial solution C (the empty clique) and the set V of all nodes of G . It then adds nodes from V to the current clique C , one at a time, checks whether the new configuration can still lead to a clique of the desired size, and invokes itself recursively on the new partial solution. The invariant maintained during execution of the algorithm is the fact that C always is a clique of G and the current set of candidate nodes V consists of all nodes which are adjacent to all nodes in C . The algorithm terminates when all cliques have been enumerated. (In a concrete implementation we will of course stop the algorithm as soon as the desired number of cliques has been produced.)

```

1 Inputs: Graph  $G = (V, E)$ , desired clique size  $k$ .
2 Outputs: All cliques  $C$  of  $G$  with  $|C| = k$ .
3 Method: begin clique( $\emptyset, V$ ) end
4 proc clique( $C, V$ )  $\equiv$ 
5   if  $|C| \geq k$ 
6     then output  $C$ 
7   else while  $V \neq \emptyset$  do
8     Choose a node  $v \in V$ ;
9      $C' := C \cup \{v\}$ ;  $V := V \setminus \{v\}$ ;
10     $V' := \{w \in V : uw \in E \forall u \in C'\}$ ;
11    if  $|C'| + |V'| \geq k$ 
12      then clique( $C', V'$ )
13    fi
14  od
15 fi
16 end

```

Figure 3: Carraghan/Pardalos clique algorithm.

If you compare the above algorithm to the original one in (Carraghan and Pardalos 1990), you will notice some slight modifications. First, the node selection strategy (step 8 of the algorithm) is adaptable rather than using a fixed node order determined before execution; this allows us to apply different search strategies as described below. Second, the recursion is terminated as soon as a clique of the given size is found; this accounts for the fact that we know in advance how large our cliques must be.

To apply the algorithm to the scale rationalization problem, it is invoked on the harmonicity graph for the given prob-

lem instance (cf. Definition 1) and with the desired clique size $k = n$. If we just want to take a quick look at some (not necessarily optimal) solutions, the algorithm can be run until the desired number of alternative solutions is obtained. But we can also use it to enumerate *all* solutions and list the “best” among them. Following Barlow, a suitable measure for the quality of a solution $S = \{x_1, \dots, x_n\}$ is the total harmonic distance $d(S)$ between all selected pitches $x_i \in C_i$, which is to be minimized:

$$d(S) = \sum_{i \neq j} d(x_i, x_j).$$

The crucial step in the algorithm is step 8 in which the next candidate node is selected. We can use different node selection schemes in this step to implement alternative search heuristics. The chosen search heuristic determines which solutions will be enumerated first and how long the search takes. In our computational experiments we found that the search can be sped up considerably by selecting an appropriate search heuristic for the type of search (full or partial enumeration) and given problem parameters (e.g., depending on whether the harmonicity graph is “sparse” or “dense”). Here are some search heuristics we found to be useful:

- The *first-first* strategy. Here we simply select nodes in the “natural” order $v_1^{x_{11}}, v_1^{x_{12}}, \dots, v_2^{x_{21}}, v_2^{x_{22}}, \dots$, where the different pitches x_{ij} of each candidate set C_i are sorted in some given order (e.g., by decreasing weighted harmonicities). This strategy is useful if we want the candidate pitches to be considered in the prescribed order.
- The *hardest-first* strategy. In this strategy we always pick a node of smallest degree in the graph induced by the current node set V . That is, we take a node $v \in V$ which minimizes the value $|\{w \in V : vw \in E\}|$. It has been observed by Carraghan and Pardalos that this approach tends to reduce the overall running time if the input graph is dense. We found this strategy most helpful when enumerating all solutions in order to find an optimal solution.
- The *random-first* strategy. Here we always select the next node at random (each node $v \in V$ is selected with the same probability). This strategy is useful if one wants to have a quick look at the average solution quality.
- The *best-first* strategy. In this strategy we always select a node $v \in V$ which minimizes the total weight $\sum_{w \in C} d(v, w)$ where $d(v, w)$ denotes the harmonic distance between the pitches represented by the nodes v and w . This strategy tends to enumerate solutions

first which have a lower average harmonic distance; it effectively turns the algorithm into a “greedy” heuristic. This is useful if we want to find some “good” (but not necessarily optimal) solutions quickly.

Note that all described variations of the algorithm take exponential time in the worst case, since it may have to explore the entire solution space before actually finding a clique. However, in practice the algorithm typically enumerates the first few cliques rather quickly. Another advantage of the algorithm is that it enables us to experiment with different search strategies and algorithm parameters until a good solution is obtained in a reasonable amount of time. The algorithm seems to be practical for the usual kinds of scales composers work with, which rarely have more than a few dozen pitches per octave.

7 Scale visualization

As an application of the methods developed in the previous sections, let us show how we can draw a scale in 2- or 3-dimensional space in such a manner that the visible (Euclidean) distances between the scale tones provide a good approximation for the actual harmonic distances. Note that if the input scale under scrutiny is non-rational, then our scale rationalization algorithm can be used as a preprocessing stage to first produce a reasonable rational approximation of the scale. Such visualizations of a scale make it easy to spot the harmonic relationships inside the scale and to compare different rationalizations of a scale by just taking a look at the corresponding pictures.

Our scale visualization method rests on the observation that the harmonic distance derived from a given subadditive harmonicity measure is a metric. As Barlow has first pointed out, we might thus use *multidimensional scaling* (henceforth abbreviated as “MDS”) to draw a scale. Whenever we have a metric d on a finite set $S = \{x_1, \dots, x_n\}$, MDS can be used to *embed* the members of the set into m -dimensional Euclidean space, by assigning a point u_i to each $x_i \in S$ such that the Euclidean distances $d_2(u_i, u_j) = \|u_i - u_j\|_2$ match the metric distances $d(x_i, x_j)$ as closely as possible. Of course, if we want to draw a real picture, say, on a computer screen, we better find an embedding in low-dimensional space. Hence for our purposes we concentrate on 2- and 3-dimensional embeddings.

MDS is routinely used in the social sciences to analyze statistical data involving similarity measurements. A fortunate consequence of this situation is that an abundance of different MDS methods is readily available, like, e.g., Torgerson’s “classical” algorithm (Torgerson 1952) and Kruskal’s gradient method (Kruskal 1964). However, the method which

nowadays is used most often for this purpose is a modern MDS algorithm due to De Leeuw and Heiser, called “SMA-COF” (the curious acronym stands for “scaling by majorizing a convex function”), see (Borg and Groenen 1997) for details. This algorithm, like most others, attempts to minimize the so-called *stress* of the embedding, which is a measure for the error in the representation, defined as:

$$\sigma = \sum_{i < j} (d_2(u_i, u_j) - d(x_i, x_j))^2.$$

If the stress is zero then the embedding provides a perfect visualization. This is rarely achieved since many interesting metrics are not Euclidean at all (this condition can be checked using Torgerson’s algorithm), and even if a metric is Euclidean then it might not be perfectly embeddable in low-dimensional space. This also happens with harmonic distance metrics, but we have found that the embeddings produced for many interesting scales provide fairly good representations of the harmonic distances, at least for the Barlow and Euler metrics. To determine the relative error in the representation one usually works with a relative kind of stress measure, called *stress-1*, which is defined as the ratio between σ and the total of the squared metric values, $\sum_{i < j} d(x_i, x_j)^2$. A rule of thumb says that an MDS solution is acceptable if the stress-1 value, as a percentage, does not exceed some 10%.

As an example, Fig. 4 shows the “harmonicity graph” of an Indian shruti scale which exhibits a fairly regular structure. This picture was produced with a software implementing the rationalization and MDS algorithms sketched out in this paper. The edges of the harmonicity graph are for a harmonicity threshold of 0.1 (i.e., edges are shown for all intervals with a maximum Barlow disharmonicity of 10). The picture makes it easy to spot the chains of fifths in the scale, as well as the thirds and sixths between the central chain and the two “side cords”. Note that this is just a side view of a 3-dimensional embedding; the chains of fifths are actually curved in the direction of the z axis which is invisible in this projection.

8 Algorithmic composition: Autobusk and Raptor

As a second application let us briefly sketch out how the harmonicity information might be used in algorithmic composition. In fact this is the application for which Barlow originally developed the concept of scale rationalization. In this context the harmonic distances can be used to drive the parameters of random note generation, thereby enabling the composer to vary the desired harmonicity of the produced notes in a continuous way. This has been implemented in

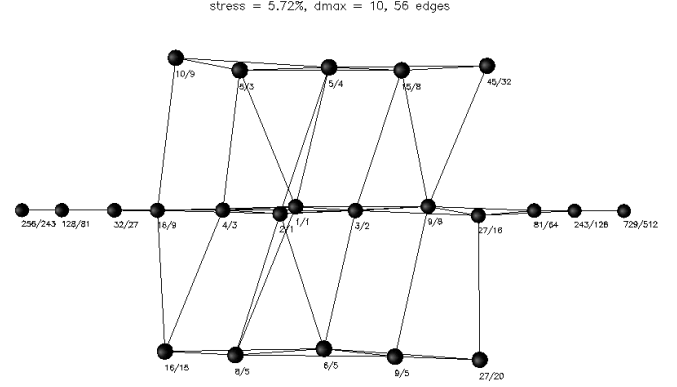


Figure 4: Indian shruti scale.

Barlow’s algorithmic composition program “Autobusk” (Barlow 2001a), and we recently developed a slightly modified, interactive version of this algorithm for our “Raptor” program. “Raptor” stands for “random arpeggiator,” so the program produces an auto-accompaniment in the form of random arpeggios in response to MIDI input.

In the Raptor program we actually use a slight variation of the definition of interval harmonicity, $h(x) = 1/(g(x) + 1)$. This maps the harmonicity of the unison to 1 (rather than infinity), which is more convenient to handle computationally. We also “factor out” octaves by just setting $g(2) = 0$, since in our application we are only concerned with tone classes rather than absolute pitches. Now, given some input notes x_1, \dots, x_n , we can assign a harmonicity measure to a “candidate note” y by computing the average harmonicity

$$\bar{h}(x_1, \dots, x_n; y) = \sqrt[n]{h(q(x_1, y)) \cdots h(q(x_n, y))},$$

where $q(x, y)$ denotes the integer ratio assigned to the interval between some notes x and y . (Note that we employ a geometric mean here, which appears to work better in this context than the arithmetic mean.) As already noted, Raptor works with MIDI notes, so in this case we can simply map the absolute difference between x and y (which is the number of semitones in the interval) to the corresponding integer ratio in a precomputed rationalization of the 12TET base scale and then compute the corresponding harmonicity value from that ratio.

The resulting average harmonicity value is then compared against minimum and maximum bounds set in the user interface of the program or via MIDI controllers, to determine whether a note is eligible to be played or not. Thus, by increasing the minimum harmonicity or decreasing the maximum harmonicity value of the program, one can increase or decrease the “harmoniousness” of the generated accompaniment in relationship to the notes played by the musician,

e.g., on a MIDI keyboard. Raptor also allows the harmonicity bounds to be varied implicitly in correspondence to the current “pulse strength” a.k.a. rhythmic accentuation, which is computed using Barlow’s “indispensability” formula (Barlow 2001b). Moreover, the average harmonicities of different eligible candidate notes can be compared to make a choice according to a “harmonic preference” criterion.

Despite the fact that Raptor uses nothing but average harmonicity values to guide the note selection mechanism, the results are fairly convincing and thus provide another experimental validation of the harmonicity concept.

9 Conclusion

We have proposed a mathematical model of scale rationalization which allowed us to show that the scale rationalization problem is NP-complete, but can be solved heuristically using graph-theoretic clique enumeration techniques. Besides the visualization of harmonic relationships in scales and the applications of harmonicity in algorithmic composition, these methods appear to have a lot of potential for other computer music applications. For instance:

- Pitch quantization: Scale rationalization allows us to map MIDI pitches to simple integer ratios depending on the current harmonic context. This enables us, e.g., to determine the correct enharmonic spelling of notes.
- Rhythmic quantization: While our methods have been developed for rationalizing scales, they also apply to other problems in which a given sequence of real numbers is to be mapped to simple integer ratios. In particular, this problem also arises when mapping the physical time values of a performance to musical time, which is known as rhythmic or time quantization and is an important part of the MIDI import facilities of sequencer and notation programs.
- Chord recognition: It should also be possible to employ scale rationalization in chord recognition algorithms, by comparing the average harmonicities between the input and a given chord table.

The last two applications in the above list have not yet been worked out and should be the subject of further investigations. Another interesting topic is the search for still better and faster scale rationalization algorithms. As we stated in Section 4, the scale rationalization problem turns out to be NP-complete already if at most three candidate pitches are given for each scale tone. This raises the question whether the problem becomes polynomial-time solvable if we only consider two alternatives per pitch. Also, the reduction from

3SAT in our NP-completeness proof (Theorem 1) employs big prime numbers and thus leaves open the possibility of a “pseudo-polynomial” algorithm (Garey and Johnson 1979, Section 4.2). Such an algorithm, if it exists, would probably be the most efficient method to solve the scale rationalization problem in practice, as the prime numbers occurring in just intervals are typically very small.

Acknowledgements

I would like to thank Clarence Barlow for introducing me to the topic of scale rationalization, for many enlightening discussions, and for commenting on earlier drafts of this paper.

References

- Barlow, C. (2001a). AUTOBUSK: A real-time pitch and rhythm generator. Technical Report 44, Musikinformatik, Johannes Gutenberg University Mainz.
- Barlow, C. (2001b). On the quantification of harmony and metre. In C. Barlow (Ed.), *The Ratio Book*, Feedback Papers 43, pp. 2–23. Cologne: Feedback Publishing Company.
- Borg, I. and P. Groenen (1997). *Modern Multidimensional Scaling*. Springer Series in Statistics. New York: Springer.
- Carraghan, R. and P. M. Pardalos (1990). An exact algorithm for the maximum clique problem. *Operations Research Letters* 9, 375–382.
- Chalmers, J. (1993). *The Divisions of the Tetrachord*. Frog Peak Music.
- Euler, L. (1739). Tentamen novae theoriae musicae ex certissimis harmoniae principiis dilucide expositae. St. Petersburg.
- Garey, M. R. and D. S. Johnson (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York: W. H. Freeman and Company.
- Helmholtz, H. v. (1983). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* (6th ed.). Hildesheim: Vieweg.
- Kruskal, J. B. (1964). Nonmetric multidimensional scaling: a numerical method. *Psychometrika* 29, 115–129.
- Large, E. W. and A. E. Tretakis (2005). Tonality and nonlinear resonance. *Annals of the New York Academy of Sciences* 1060, 1–4.
- Plomp, R. and W. J. M. Levelt (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America* 38, 548–568.
- Sethares, W. A. (2005). *Tuning, Timbre, Spectrum, Scale* (2nd ed.). London: Springer.
- Tenney, J. (1988). *A History of ‘Consonance’ and ‘Dissonance’*. New York: Excelsior Music Publishing Company.
- Torgerson, W. S. (1952). Multidimensional scaling: I. Theory and method. *Psychometrika* 17, 401–419.