Chapter 7

2D/3D Theory of Music

This chapter describes my older 2D/3D theory of music, which was formulated in response to observations about the vector representations of musical intervals and the various mappings between them.

Firstly we look at some more vector and point space mappings: a 2D to 1D vector mapping which maps both tones and semitones to "steps", and the visual 3D to 2D point space mapping which maps the 3D world to 2D (retinal) images. Then I discuss the major concept in the 2D/3D theory, which is the suggestive analogy between the musical 3D to 2D mapping and the visual 3D to 2D mapping.

7.1 More Vector Space Mappings

7.1.1 Another Mapping from 2D to 1D

We've looked at the "natural" mapping from 2 dimensions to 1 dimension, i.e. the one that maps tones and semitones to semitones. But there is another mapping from the 2-dimensional space to a 1-dimensional space that could be considered relevant to understanding music perception. This is the mapping that maps both a tone and a semitone to a step. The "step" represents a step on the diatonic scale that one takes as one goes from one note to the next note on the scale. We cannot consider the target space of this mapping to be the same as the 1-dimensional semitone space, so perhaps we can call it the **1-dimensional step space**. This mapping "forgets" the difference between a tone and a semitone, in the sense that looking at an output vector consisting of n steps, we cannot tell which of those n steps in the input vector were semitones and which were tones. It is represented by the following matrix:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

So why might this mapping be important for understanding music perception? There are many tunes where a first phrase consists of some sequence of notes played in a certain rhythm, and then a second phrase consists of the same sequence of notes *transposed along the diatonic scale*, played in the same rhythm. This transposition is different from the normal sort of transposition, which refers to an exact translation such as when a key change occurs. The exact pattern of intervals in the second phrase will be different from that in the first phrase, because some tones will change to semitones, and vice versa. To give a simple example, the first phrase might be CDEDEE, and the second phrase could be DEFEFF, which is transposed one "step" up the scale.

But if we apply the forgetful 2D to 1D mapping that we have just described, then the mapped version of the second phrase is an exact translation of the mapped version of the first phrase.

This seems a promising notion. But if it really forms an aspect of music perception, there would have to be some cortical map that performs this mapping. If we assume that the cortical maps that process music already exist to serve some other purpose, then it is unlikely that such a cortical map exists, because there is no other reason why the brain would want to process information about musical intervals in this way; in particular scales do not occur outside music, and speech melodies do not have a structure which can be factored into independent dimensions of tone and semitone. In Chapter 10, the **melodic contour cortical map** is introduced. This map ignores the difference between tones and semitones in many cases, not because there is a 2D to 1D mapping, but rather because it processes pitch information with a reduced level of precision.

7.1.2 Another Perceptual 3D to 2D Mapping

The world we live in is 3-dimensional. We make representations of parts of the world in pictures and photographs which are 2-dimensional. The images on the retinas of our eyes are 2-dimensional, and our brain reconstructs a model of the 3-dimensional world from the information in these two 2-dimensional images. The correspondence between a 3D scene and its 2D picture can be described as a mapping from a 3D point space to a 2D point space. By considering vectors defined by pairs of points in the 3D and 2D spaces, we can define a corresponding mapping from a 3D vector space to a 2D vector space. As already mentioned in Chapter 5, a mapping between point spaces that defines a corresponding well-defined linear mapping between vector spaces is called an **affine mapping**. The mapping between a 3D scene and a 2D picture is *not* an affine mapping. This has to do with the fact that things far away are smaller on the picture than things that are close. The technical name for such a mapping is a **projective mapping**.

However, if we consider a very small portion of the 3D scene ("small" in the sense of being a small volume of limited diameter), which is a large distance from the point of view that defines the picture ("large" compared to the size of the "small" portion), then the mapping is *approximately* affine, and there is a corresponding approximately linear mapping of displacement vectors.[1]

Furthermore, the human brain necessarily has an ability to process the correspondence between 3D scenes and 2D pictures of those scenes. This ability underlies our ability to perceive 3D from the 2D information provided by our retinas.

The first assumption of the 2D/3D theory of music is that there is a significant analogy between the two different 3D to 2D mappings:

- the musical 3D to 2D natural mapping which maps from the 3D representation of musical intervals to the 2D tone/semitone representation of musical intervals, and,

- the visual 3D to 2D natural mapping which maps from arbitrarily small displacement vectors in an arbitrarily small portion of a 3D scene to their images in a 2D picture (with the point of view not too close to said portion).

Translated into the language of neurons and cortical maps, this analogy suggests two possible hypotheses about the relationship between the two types of 3D/2D mapping:

1. There is a cortical map somewhere in the brain that processes the relationship between 2D and 3D in the brain, and this cortical map also processes the relationship between 2D and 3D in music, or,

2. there is a set of neurons somewhere in the brain with an intrinsic ability to process 2D/3D relationships. Most of them are recruited to process the relationship between 3D objects and 2D images, but some of them get recruited to the task of processing 2D/3D relationships in music.

The problem with the first hypothesis is that we would then expect listening to music to *feel like* visual perception of the real 3-dimensional world. We would expect this because that is the generally observed fact about cortical

[1] If a point space mapping is *not* affine, not only will the corresponding vector mapping not be linear, it won't even be well-defined (the mapped value of a vector will vary depending on which two points are used to define it). But if we assume that the point space mapping over a small enough portion of the point space is sufficiently close to affine, then the corresponding vector space mapping will be correspondingly close enough to being well-defined.

maps: two different experiences or perceptions or emotions feel the same if and only if the same neurons are active in both cases.

The second hypothesis is an attempted work-around to this problem, i.e., the same sort of neurons process visual 2D/3D and musical 2D/3D, but there is no actual overlap in which neurons are active in each case, and that is why music does not feel like visual perception of 3D space.

As stated so far, the 2D/3D theory provides an explanation for the diatonic scale, and it explains the relevance of harmonic relationships between notes in the scale, but it does not explain any other features of music.

7.2 The Looping Theory

The second assumption of the 2D/3D theory is based on two observations:

1. Music tends to go around in circles. Tunes start on a home note and a home chord (prototypically the note C and the chord C major which consists of the notes C, E and G), travel a path visiting other notes and chords, and finally return to the home note and the home chord.

2. The Harmonic Heptagon (see end of Chapter 5) defines a cyclic path around the diatonic scale.

So maybe the 3D representation of notes, as defined by the 3D representation of the intervals between different notes, travels once around the Harmonic Heptagon as it travels from the initial home note and chord to the final home note and chord. This implies that the final home note is displaced from the initial home note by the 3D vector $(-4, 4, -1)$ which represents the syntonic comma of 81/80 (or by $(4, -4, 1)$ representing 80/81, depending on which way we go around the loop). In 3 dimensions the tune travels along something like a spiral, and the 2-dimensional picture is seen from a point of view such that the spiral looks like a closed circle. To close the gap corresponding to the syntonic comma, the point of view has to be one such that points separated by a multiple of the $(-4, 4, -1)$ vector are in the same line of sight, and thus occupy the same position in the 2D image.

The looping theory adds some extra constraint into the 2D/3D theory. Furthermore, we can relate common chord sequences to a trip around the Harmonic Heptagon. For example, a common chord sequence is C major, F major, G7, C major. To make the theory work there has to be some method of determining where each chord would be placed on the Harmonic Heptagon relative to previous chords that have already occurred in the tune. The tune starts with C major (CEG). Next is F major (FAC). It seems reasonable to regard the F major as being connected to the C major via the shared note C. Moving on to G7 (GBDF), it seems reasonable again to connect it to F major by the shared note F. And then the G7 will be connected to the final C major by the shared note G, which completes a full circle going clockwise around the heptagon. In 3D space, the final C major chord is located in a position displaced from the initial C major chord by the vector $(4, -4, 1)$.

7.3 Outlook for the 2D/3D Theory

Unfortunately my development of the 2D/3D theory has not made any further progress. And I have now developed the newer **super-stimulus theory**, which has a much better foundation in biological theory, and is able to explain many aspects of music in plausible and convincing detail. But given uncertainty about some parts of the super-stimulus theory, and the incompleteness of that theory, I can't rule out the possibility that the older 2D/3D theory has some relevance to a final and complete explanation of music.

The concept of the Harmonic Heptagon does turn out to be important for developing certain aspects of the super-stimulus theory, in particular the theory of home notes and home chords. And the 1D/2D/3D vector theory of intervals gives a complete picture of all the relationships between intervals described as tones plus semitones and intervals described in terms of simple fractional ratios (if those ratios are considered not to have any prime factors in the numerators and denominators other than 2, 3 and 5). So the analysis of intervals as vectors was a useful analysis to do, even if the full 2D/3D theory turns out to be incorrect.

I will finish this section with a list of unresolved issues around the 2D/3D theory:

- The analysis of chordal movement around the heptagon doesn't say anything about melody. We have to find a way to relate the notes of the melody to the notes of the harmony within the framework of the theory.

- One can attempt to place or locate notes of the melody in 3D space in the same sort of way that I described chords being located. This requires us to define rules as to which harmonic intervals between which notes are to be used to locate notes relative to each other. The desired result is that the final home note is located at a position in 3D space displaced from the position of the initial home note by the syntonic comma. Presumably the displacement calculated by calculating the locations of notes in the melody should be consistent with the rules for calculating the locations of chords, particularly if the chords are implied by the melody.

- Consecutive chords do not always share notes, so shared notes cannot always be used as a basis for determining where to locate chords relative to each other in 3D space. They can also share more than one note, in some cases giving rise to two different choices of relative location.

- The theory doesn't say much about time and rhythm. The best it can do is suppose that the times that notes occur play a role in the rules that determine which relationships between which pairs of notes determine relative locations in 3D space. A bigger difficulty is that there is some degree of musicality in music that consists only of rhythmical percussion—something that a theory based on frequency ratios cannot possibly explain. (The super-stimulus theory does better here, as it can explain the musicality of music that has no melody or harmony at all.)

- The 2D/3D theory depends too much on specific features of the well-tempered diatonic scale, in particular that the steps are all one of two sizes.

- The theory assumes that ratios involving 7 (or higher prime numbers) are musically unimportant. For example, adding 7 would increase the number of dimensions from 3 to 4. This is less of an issue with the super-stimulus theory. The construction of the Harmonic Heptagon is based on powers of 2, 3 and 5; and the super-stimulus theory does make use of the Harmonic Heptagon to analyse some aspects of Western diatonic music. But the super-stimulus theory does not *depend* on the existence of this heptagon to explain *all* music—it only makes use of the heptagon to explain relevant properties of music based on the scale that the Harmonic Heptagon is constructed from.

- As already mentioned in the introduction, the 2D/3D theory is analogous to the paradoxical drawings of M.C. Escher, which exploit the ambiguity in 3D space of the location of points represented on a 2D image. But looking at an Escher drawing does not "feel like" listening to music, whereas one might expect it to do so if the same paradox applied to perceptions processed by the same cortical maps in each case.