Chapter 5

Vector Analysis of Musical Intervals

The intervals between musical notes can be regarded as **vectors** in a **vector space**. Intervals in the diatonic scale have natural 1-, 2- and 3-dimensional vector representations, and there are also natural mappings from 2 to 1, 3 to 1 and 3 to 2 dimensions. The **kernel** of the natural 3D to 2D mapping is generated by the **syntonic comma** which equals 81/80. The **Harmonic Heptagon** provides a compact visualisation of all the consonant relationships between notes in the diatonic scale, and a trip once around the heptagon corresponds to one syntonic comma.

5.1 Three Different Vector Representations

When I first started trying to understand all the relationships between notes on the diatonic scale, there seemed to be almost too many different ways to describe the intervals between pairs of notes.

(The analysis in this chapter applies to the diatonic scale; however, for the sake of concreteness, all examples are based on the white notes scale, i.e. C, D, E, F, G, A, B.)

Firstly, an interval between two notes can be described as the logarithm of the frequency ratio between the notes. On the well-tempered scale all intervals are integral powers of $\sqrt[12]{2}$, so this is equivalent to a simple count of semitones. For example, the interval from a C to the next higher G is 7 semitones (i.e. the ratio $2^{7/12}$).

The representation as a count of semitones can describe all possible intervals, but it does not take into account the structure of the diatonic scale. Some steps from one note to the next are tones, and some are semitones. So a second formulation is to count the number of tones and semitones in an interval separately. For example, the interval from a C to the next higher G would be 3 tones and 1 semitone.

So far, we have two ways to describe intervals between notes, yet neither of them says anything about the most important feature of musical intervals, which is that chords, harmony and the repeating structure of the scale are all based on intervals that correspond to simple fractional ratios of frequencies. For example, the interval from C to the next higher G corresponds to a ratio of approximately 3/2.

These simple fractional ratios form the basis of a third representation. The third representation is different from the other two, because it only applies to some intervals—only the consonant intervals have obvious representations as fractional ratios. Any ratio assigned to other intervals is somewhat arbitrary, and there is no best way of making such an assignment to all intervals.

To analyse the relationships between these three representations of musical intervals—semitones, semitones plus tones, and fractional ratios—we need a common framework for specifying them. Luckily there is a ready-made mathematical structure that we can use: each of the three representations defines a **vector space**.

5.1.1 What is a Vector Space?

Vectors are mathematical objects with **magnitude** and **direction**. Vectors can also be formulated in terms of **components**. The component formulation will turn out to be more useful for the current analysis. Also the vector spaces that we will define are all **finite dimensional**, which makes everything a lot easier.

We can define a **finite dimensional vector space** V as follows:

- The vector space has some number n of dimensions. We say that V is n-dimensional, or nD for short. (The only values for n that we use in this chapter are 1, 2 and 3.)

- A vector belonging to an n-dimensional vector space V has n components. Each component is a number.[1] We can write the components as a comma-separated list in brackets; for example, $(2, 3)$ is an example of a vector belonging to a 2-dimensional vector space, and $(-1, 0, 5)$ is an example of a vector belonging to a 3-dimensional vector space.

- Two vectors from the same space V are equal if and only if all their corresponding components are equal. (We can say that equality is defined

[1] Generally a **real** number, although most of the components of the vectors we are dealing with will be integers.

componentwise.) For example, $(3,2,1) = (3,2,1)$ but $(3,2,-1) \neq (3,1,-1)$ because the second components 2 and 1 are not equal.

- Vectors from a vector space V can be added together by adding their corresponding components. For example, $(1,0,-2) + (3,3,4) = (1+3, 0+3, -2+4) = (4,3,2)$, and $(3,4)+(1,-3) = (4,1)$, as in Figure 5.1. (Thus addition is componentwise.)

- A vector from a vector space can be multiplied by a number, often called a **scalar** to distinguish it from a vector, by multiplying each of its components by the number. This is called **scalar multiplication**. The scalar is normally written on the left of the multiplication. For example, $4\times(1,0,-2) = (4\times1, 4\times0, 4\times-2) = (4,0,-8)$, and $3\times(2,1) = (6,3)$ (see Figure 5.2).[2] (Scalar multiplication is also componentwise.) Note that the definition of scalar multiplication is consistent with the definition of addition, in that, for example, $2\mathbf{x} = \mathbf{x} + \mathbf{x}$ for any vector $\mathbf{x}$ belonging to a vector space V.

We also want to define an n-dimensional **point space**. Just like a vector in a vector space, a **point** in an n-dimensional point space can be written as a list of n components, where each component is a number. The important difference between a point space and a vector space is that they have different operations defined on them. There is no way to add points to each other or to multiply points by a scalar. We can, however, add a point to a vector to get another point, which we do by adding corresponding coordinates (exactly as for vector addition—see Figure 5.3).

In any vector space there is a well-defined zero vector $\mathbf{0}$ which is the vector whose components are all zero. For every vector $\mathbf{x}$, $\mathbf{x} + \mathbf{0} = \mathbf{x}$, and for every point p, $p + \mathbf{0} = p$. In a point space there will be a point called the **origin**, which has all components 0, but if we are choosing coordinates for a space, it is somewhat arbitrary which point in the space we choose to be the origin.[3]

In our musical point spaces, the points in each point space will represent musical notes, and the vectors will represent intervals between pairs of musical notes. We will generally choose the note middle C to be the origin in the coordinate systems of our point spaces. Note that some of our point spaces

[2] In this book I use three different notations for multiplication. For example, to multiply 2 by x we can write $2x$ or $2\cdot x$ or $2\times x$. Sometimes in mathematics different multiplicative notations are used for different types of multiplication, but here we are not defining more than one kind of multiplication for any pair of mathematical objects that can be multiplied together. The first notation is the most compact, but it cannot be used to multiply numbers ($3\times 2 \neq 32$); the "dot" notation is the next most compact, and is OK as long as there is no danger of confusing the dot with a decimal point; and the traditional "$\times$" is the least compact but most explicit notation. ("$\cdot$" and "$\times$" are also standard notations for different ways of multiplying vectors together, but there is no multiplication of vectors by other vectors in this book, so we can get away with using them to represent numerical and scalar multiplication.)

[3] To put it another way, there is no operation that we are allowed to define on the point space that can actually tell us whether or not a point is the origin.

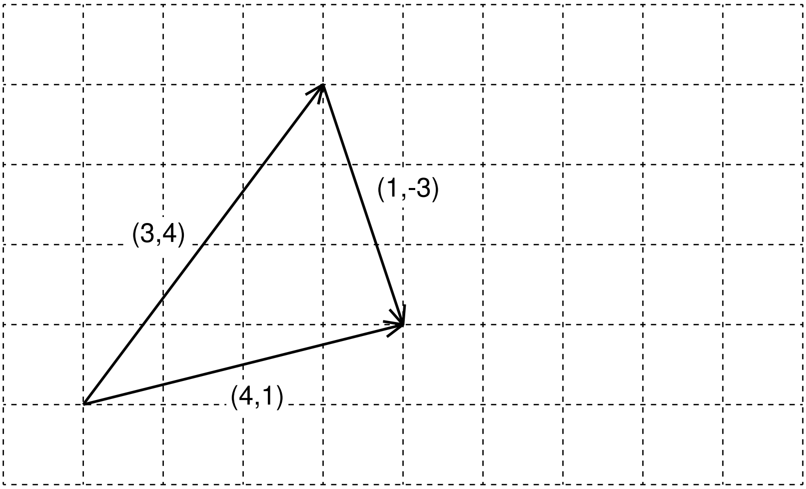

Figure 5.1. Vector addition: $(3,4)+(1,-3)=(4,1)$.

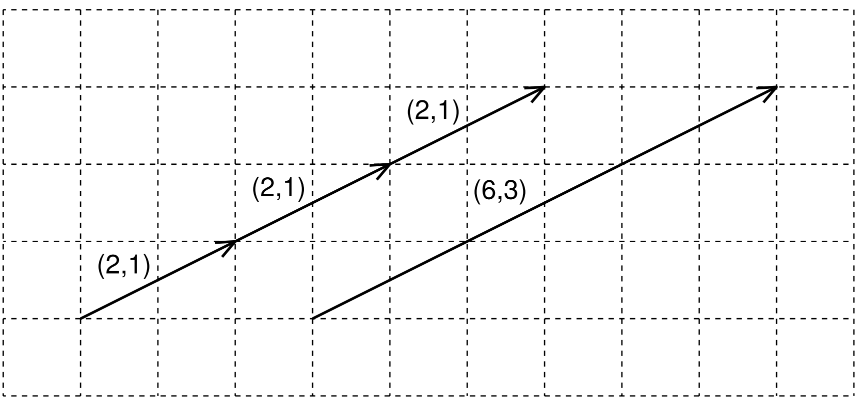

Figure 5.2. Vector scalar multiplication: $3 \times (2,1) = (6,3)$.

will have multiple points representing each note, in which case it will be more precise to say that we will choose an origin such that the origin represents middle C (and other points in the point space may also represent middle C).

Having done the theory, we can see what vector spaces and corresponding point spaces we get from our three representations of musical intervals.

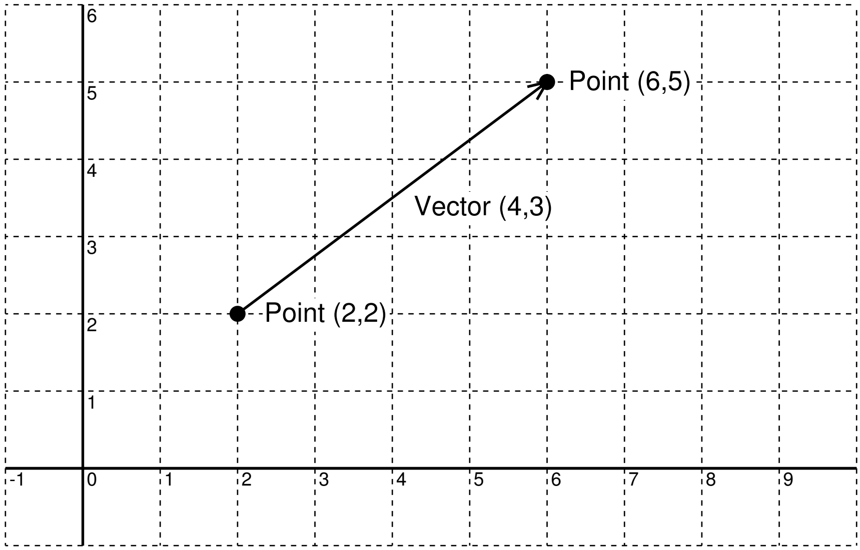

Figure 5.3. Addition of a vector to a point: point (2,2) + vector (4,3) = point (6,5).

5.1.2 1D Semitones Representation

The semitones representation is the simplest. Imagine all the notes of the chromatic scale laid out evenly along a straight line. An interval from one note to another is represented by a vector containing one component, which is the number of semitones it takes to get from the start note to the end note (positive if we are going up, negative if we are going down). For example, middle C is the point (0), the G above it is (7), and the interval from C to G is represented by the vector (7), which means 7 semitones.

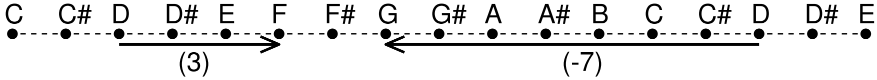

Figure 5.4. 1D semitones representation. Two vectors are shown: a 3 semitones interval (a minor third) going from D up to F, and a −7 semitones interval (a "negative" perfect fifth) going down from D to G.

5.1.3 2D Tones/Semitones Representation

For this representation we take the number of tones in an interval to be the first component, and the number of semitones to be the second component. So for our interval C $\to$ G, the vector is $(3, 1)$, which means 3 tones + 1 semitone. We can imagine the point space as a slightly irregular stairway: for each tone step we travel one unit to the right, for each semitone step we travel one unit upwards.

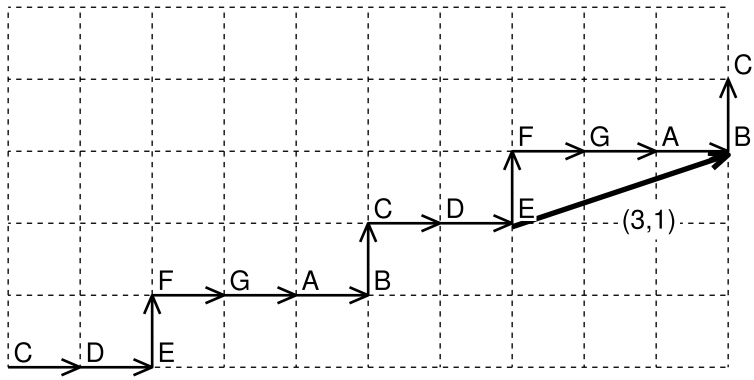

Figure 5.5. 2D tones/semitones representation. The perfect fifth interval from E to B is represented by the vector $(3, 1)$ = 3 tones + 1 semitone.

5.1.4 3D Consonant Interval Representation

When we combine consonant intervals, we have to *multiply* the corresponding fractions. For example, the equation

minor 3rd + major 3rd = perfect fifth

is represented by $6/5 \times 5/4 = 3/2$.

The way to convert multiplication of fractions to addition of vectors is to create vectors where each component is the power of a prime number that occurs in the fraction (as a factor of the numerator or the denominator). As it happens, we assume that the only primes occurring in the relevant fractions are 2, 3 and 5. The three different primes are where our three dimensions come from. For example, $6/5 = 2^1 \times 3^1 \times 5^{-1}$, so we can represent it by the vector $(1, 1, -1)$. Similarly $5/4 = 2^{-2} \times 3^0 \times 5^1$ is represented by $(-2, 0, 1)$, and $3/2 = 2^{-1} \times 3^1 \times 5^0$ is represented by $(-1, 1, 0)$. We can check that the vector addition gives the right answer: $(1, 1, -1) + (-2, 0, 1) = (1 - 2, 1 + 0, -1 + 1) = (-1, 1, 0)$. This corresponds to the multiplication $6/5 \times 5/4 = (2^1 \times 3^1 \times 5^{-1}) \times (2^{-2} \times 3^0 \times 5^1) = (2^1 \times 2^{-2}) \times (3^1 \times 3^0) \times (5^{-1} \times 5^1) = 2^{1+-2} \times 3^{1+0} \times 5^{-1+1} = 2^{-1} \times 3^1 \times 5^0 = 3/2$.

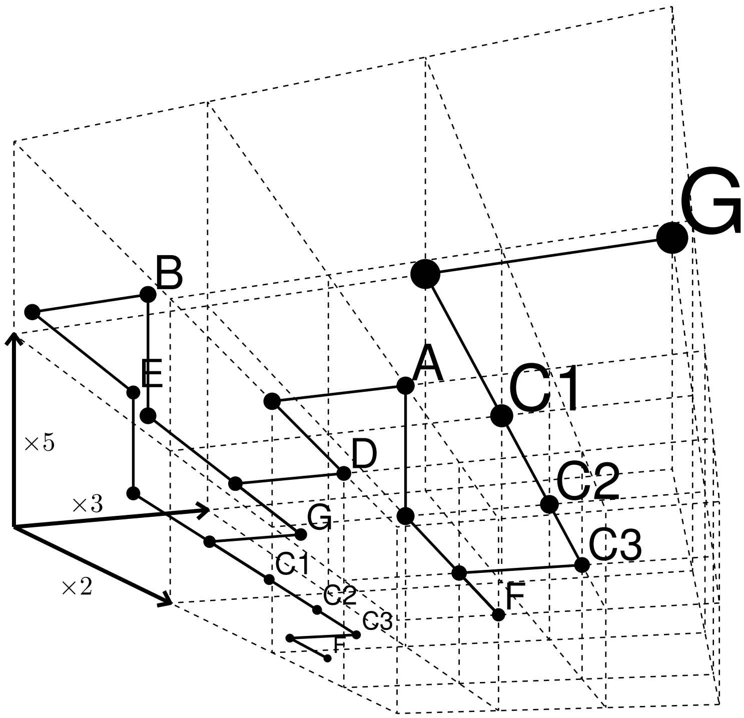

Figure 5.6. 3D musical point space. The C's are labelled according to which octave they are in. The repetition of the pattern is shown for F, C1, C2, C3 and G.

Figure 5.6 attempts to show notes in the 3D point space. Unfortunately 3D visualisation is difficult to do on 2-dimensional paper. One way we can simplify the 3D representation is by removing the least important dimension. The best dimension to remove is the ×2 dimension, i.e. the octaves, because the 3D representation represents harmonic relationships between notes, and as we will see, the brain represents harmonic relationships modulo octaves anyway.

Figure 5.7 shows the notes on the scale in this "flattened" 3D representation. Octaves have been flattened to zero, so the ×3 unit vector is equivalent to a perfect fifth (×3/2) and the ×5 unit vector is equivalent to a major third (×5/4).

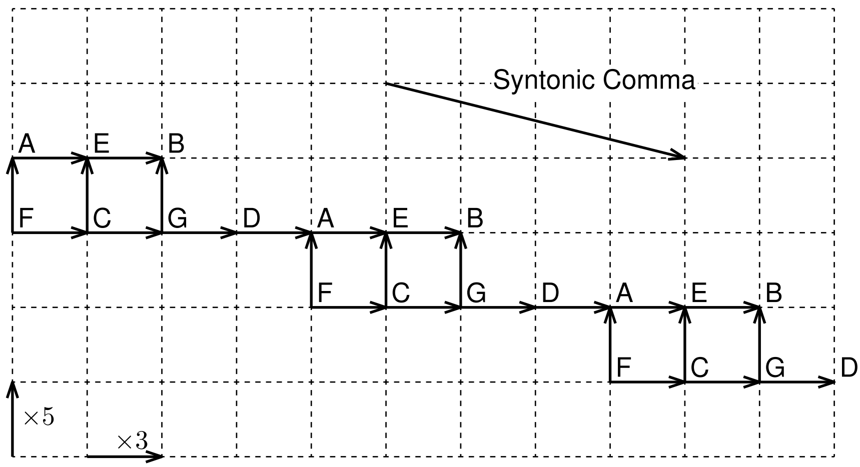

Figure 5.7. "Flattened" 3D musical point space. It is much easier to see the repeated representations of notes in this representation. The period of repetition is the **syntonic comma**, as shown in the diagram, which is discussed in more detail later in this chapter.

5.2 Bases and Linear Mappings

An important fact about an n-dimensional vector space is that we can **generate** all the elements of the vector space from a suitably chosen **basis** (plural **bases**) of n vectors. The required generation is by means of addition and scalar multiplication. For example, in a 2-dimensional vector space, we can choose $(1,0)$ and $(0,1)$ as a basis. Then for any vector (x,y), we can generate it as $x(1,0) + y(0,1)$. (Calculating: $x(1,0) + y(0,1) = (x\times 1, x\times 0) + (y\times 0, y\times 1) = (x\times 1 + y\times 0, x\times 0 + y\times 1) = (x,y)$). Similarly, for a 3-dimensional space we can choose $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$ as a basis.

These examples of vector bases are formed by selecting n vectors e_i, $i = 1..n$ where the jth component of e_i is 1 if $i = j$ and 0 otherwise. But there are many other bases that we can choose. In fact, almost any randomly chosen set of n vectors will form a basis for an n-dimensional vector space. The only requirement is that the n vectors must be **independent**, which means that none of the basis vectors can be generated from any or all of the other basis vectors.

Vectors spaces are a type of **mathematical structure**. The structure consists of the vectors in the vector space and the operations of addition and scalar multiplication. Give mathematicians a structure and they will always ask: what **mappings** are there from one space to another that preserve (or "respect") the structure? A **mapping** (or **function**) is some rule that specifies for each input value from one set called the **domain** a single output value from another set called the **codomain**.

What exactly do mathematicians mean by "preserving" a structure? In the case of a mapping from a vector space to another vector space, what this means for a mapping f is that if we add two input values and apply f to the result, it's the same as if we applied f to the input values first, and then added those two output values together. Writing this is as an equation we get:

$$f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$$

Similarly the mapping must respect scalar multiplication: if we multiply the input value by a scalar and apply f, we get the same result as if we applied f first and then multiplied by the scalar. As an equation this is:

$$f(a \cdot \mathbf{x}) = a \cdot f(\mathbf{x})$$

A mapping with these properties is a **linear mapping**. If we have a mapping from a point space to another point space which generates a corresponding linear mapping between the associated vector spaces, then the point space mapping is called an **affine mapping**. If we already have a linear mapping f between vector spaces associated with a pair of point spaces, then for any pair of points p_i in the input space and p_o in the output space, we can define an affine function g such that $g(p_i) = p_o$, and for any vector $\mathbf{v}_i$ in the input vector space, $g(p_i + \mathbf{v}_i) = g(p_i) + f(\mathbf{v}_i)$. In other words, we can extend the mapping of vector spaces to a corresponding affine mapping of point spaces that is consistent with the vector mapping, and which maps the chosen point in the input space to the chosen point in the output space.

In the case of our musical point and vector spaces, we are only going to be interested in affine mappings that consistently map points representing notes to points representing exactly the same notes. Similarly for vector mappings and intervals, we are interested in vector mappings that map vectors representing intervals to vectors representing the *same* intervals. These mappings are **natural mappings**, where "natural" can be interpreted to mean the most obvious or meaningful.

5.2.1 2D to 1D Natural Mapping

How do we naturally map from the 2D tone/semitones representation of intervals to the 1D semitones representation? This is an easy one: we want to reduce an interval described as x tones and y semitones to the form of z semitones. We know that 2 semitones = 1 tone, so the answer is $z = 2x + y$.

As mathematicians are never happy with anything that is *too* easy, given a question that was easy to answer, they try to think of a related question that might be a bit harder. A good question for this mapping is: can we *reverse* the mapping? In other words, given the output value, can we determine an input value, and will this input value be unique?

In the first instance it seems obvious that there are many possible input values for each output value. For example, $(2, 0)$, $(1, 2)$ and $(0, 4)$ all map to (4). In other words, an interval of 4 semitones could be 2 tones + 0 semitones, or 1 tone + 2 semitones, or 0 tones + 4 semitones. But, only one of these input vectors represents an actual interval between two notes on the diatonic scale, i.e. $(2, 0)$. So the question that follows is: is there a unique input value for each possible output value if we restrict input values to those vectors that correspond to intervals between notes on the diatonic scale? To give an exhaustive answer to this question, the following tables list all non-negative intervals on the diatonic scale which are not greater than an octave, showing them as tone/semitone vectors and as semitone vectors, grouped by equality of the semitone vector:

Interval	Tone/Semitone	Semitone
C → C	(0, 0)	(0)
D → D	(0, 0)	(0)
E → E	(0, 0)	(0)
F → F	(0, 0)	(0)
G → G	(0, 0)	(0)
A → A	(0, 0)	(0)
B → B	(0, 0)	(0)

Interval	Tone/Semitone	Semitone
E → F	(0, 1)	(1)
B → C	(0, 1)	(1)

Interval	Tone/Semitone	Semitone
C → D	(1, 0)	(2)
D → E	(1, 0)	(2)
F → G	(1, 0)	(2)
G → A	(1, 0)	(2)
A → B	(1, 0)	(2)

Interval	Tone/Semitone	Semitone
D → F	(1, 1)	(3)
E → G	(1, 1)	(3)
A → C	(1, 1)	(3)
B → D	(1, 1)	(3)

Interval	Tone/Semitone	Semitone
C → E	(2, 0)	(4)
F → A	(2, 0)	(4)
G → B	(2, 0)	(4)

Interval	Tone/Semitone	Semitone
C → F	(2, 1)	(5)
D → G	(2, 1)	(5)
E → A	(2, 1)	(5)
G → C	(2, 1)	(5)
A → D	(2, 1)	(5)
B → E	(2, 1)	(5)

Interval	Tone/Semitone	Semitone
F → B	(3, 0)	(6)
B → F	(2, 2)	(6)

Interval	Tone/Semitone	Semitone
C → G	(3, 1)	(7)
D → A	(3, 1)	(7)
E → B	(3, 1)	(7)
F → C	(3, 1)	(7)
G → D	(3, 1)	(7)
A → E	(3, 1)	(7)

Interval	Tone/Semitone	Semitone
E → C	(3, 2)	(8)
A → F	(3, 2)	(8)
B → G	(3, 2)	(8)

Interval	Tone/Semitone	Semitone
C → A	(4, 1)	(9)
D → B	(4, 1)	(9)
F → D	(4, 1)	(9)
G → E	(4, 1)	(9)

Interval	Tone/Semitone	Semitone
D → C	(4, 2)	(10)
E → D	(4, 2)	(10)
G → F	(4, 2)	(10)
A → G	(4, 2)	(10)
B → A	(4, 2)	(10)

Interval	Tone/Semitone	Semitone
C → B	(5, 1)	(11)
F → E	(5, 1)	(11)

Interval	Tone/Semitone	Semitone
C → C	(5, 2)	(12)
D → D	(5, 2)	(12)
E → E	(5, 2)	(12)
F → F	(5, 2)	(12)
G → G	(5, 2)	(12)
A → A	(5, 2)	(12)
B → B	(5, 2)	(12)

And the answer to our question, as to whether we can reverse the mapping from the tone/semitone representation to the semitone representation, is that the mapping can be inverted uniquely for all intervals except 6 semitones. There are two possible input values—$(3, 0)$ and $(2, 2)$—which map to an interval of 6 semitones. The tables above only list interval sizes from 0 semitones to 12 semitones. However, an octave is always represented by $(5, 2)$, and all other possible intervals can be created (from one of the intervals listed) by adding a whole number of octaves to either the start note or the end note. So a full answer is: the mapping can be inverted uniquely for all intervals, except $12n + 6$ semitones (for any integer n). It perhaps should be noted that this "almost" inverse mapping is not at all linear; for example, $f^{-1}((4)) = (2, 0)$ but $f^{-1}((8)) = (3, 2) \neq 2 \times (2, 0)$.[4]

This answer will turn out to help us when we come to define a natural 3D to 2D mapping. We will want to know that the 2D to 1D mapping can be uniquely reversed for all consonant intervals, which is the case, because the only interval for which it cannot be uniquely reversed is 6 semitones, and 6 semitones is not a consonant interval.

Finally, the natural 2D to 1D mapping can be represented by a **matrix**:

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Each row of the matrix represents a component position in the input value, and each column of the matrix represents a component position in the output value. Each number represents the contribution that the corresponding input component makes to the corresponding output component. This matrix has 2 rows and 1 column because our mapping is from a 2D space to a 1D space.

[4]Or rather the reverse mapping is only linear if we restrict it to head-to-tail addition of vectors corresponding to pairs of intervals on the diatonic scale.

5.2.2 3D to 1D Natural Mapping

How do we naturally map from the 3D fractional ratio representation of intervals to the 1D semitones representation?

We determined the 2D to 1D mapping by calculating the value we expected to get for basis vectors $(1, 0)$ and $(0, 1)$. We can do the same thing to construct a 3D to 1D map. $(1, 0, 0)$ represents a ratio of 2, which equals 12 semitones. $(0, 1, 0)$ represents a ratio of 3—more precisely an *approximate* ratio of 3—which equals 19 semitones. And $(0, 0, 1)$ represents a ratio of approximately 5, which equals 28 semitones. So our corresponding matrix is:

$$\begin{pmatrix} 12 \\ 19 \\ 28 \end{pmatrix}$$

5.2.3 3D to 2D Natural Mapping

We can continue to use the same pattern to construct a natural 3D to 2D mapping: specify suitable values for the basis vectors. In keeping with our need to interpret intervals as intervals between actual notes on the scale, we want to choose 2D vectors that represent such intervals. Our chosen basis vectors in the 3D space represent consonant intervals, and we have already seen that all consonant intervals have unique representations as tone/semitone vectors. 12 semitones (an octave) equals 5 tones plus 2 semitones, 19 semitones equals 8 tones plus 3 semitones and 28 semitones equals 12 tones plus 4 semitones. We can write these numbers out as another matrix:

$$\begin{pmatrix} 5 & 2 \\ 8 & 3 \\ 12 & 4 \end{pmatrix}$$

The question remains as to whether this mapping gives correct answers for all consonant intervals. Why is this important? We have proven that the mapping works for our basis vectors. But does this proof automatically extend to *all* consonant intervals? One way to be sure is to check all the possibilities (for consonant intervals less than an octave):

$$\begin{aligned} &3 \text{ semitones} \approx 6/5 = 2^1 \times 3^1 \times 5^{-1} : \\ &\quad f(1, 1, -1) = (1 \cdot \mathbf{5} + 1 \cdot \mathbf{8} - 1 \cdot \mathbf{12}, 1 \cdot \mathbf{2} + 1 \cdot \mathbf{3} - 1 \cdot \mathbf{4}) \\ &\qquad = (5 + 8 - 12, 2 + 3 - 4) \\ &\qquad = (1, 1) \end{aligned}$$

$$\begin{aligned}
&\text{4 semitones} \approx 5/4 = 2^{-2} \times 3^0 \times 5^1 : \\
&\quad f(-2,0,1) = (-2\cdot\mathbf{5} + 0\cdot\mathbf{8} + 1\cdot\mathbf{12}, -2\cdot\mathbf{2} + 0\cdot\mathbf{3} + 1\cdot\mathbf{4}) \\
&\qquad = (-10+0+12, -4+0+4) \\
&\qquad = (2,0) \\
&\text{5 semitones} \approx 4/3 = 2^{2} \times 3^{-1} \times 5^0 : \\
&\quad f(2,-1,0) = (2\cdot\mathbf{5} - 1\cdot\mathbf{8} + 0\cdot\mathbf{12}, 2\cdot\mathbf{2} - 1\cdot\mathbf{3} + 0\cdot\mathbf{4}) \\
&\qquad = (10-8+0, 4-3+0) \\
&\qquad = (2,1) \\
&\text{7 semitones} \approx 3/2 = 2^{-1} \times 3^{1} \times 5^0 : \\
&\quad f(-1,1,0) = (-1\cdot\mathbf{5} + 1\cdot\mathbf{8} + 0\cdot\mathbf{12}, -1\cdot\mathbf{2} + 1\cdot\mathbf{3} + 0\cdot\mathbf{4}) \\
&\qquad = (-5+8+0, -2+3+0) \\
&\qquad = (3,1) \\
&\text{8 semitones} \approx 8/5 = 2^{3} \times 3^{0} \times 5^{-1} : \\
&\quad f(3,0,-1) = (3\cdot\mathbf{5} + 0\cdot\mathbf{8} - 1\cdot\mathbf{12}, 3\cdot\mathbf{2} + 0\cdot\mathbf{3} - 1\cdot\mathbf{4}) \\
&\qquad = (15+0-12, 6+0+-4) \\
&\qquad = (3,2) \\
&\text{9 semitones} \approx 5/3 = 2^{0} \times 3^{-1} \times 5^{1} : \\
&\quad f(0,-1,1) = (0\cdot\mathbf{5} - 1\cdot\mathbf{8} + 1\cdot\mathbf{12}, 0\cdot\mathbf{2} - 1\cdot\mathbf{3} + 1\cdot\mathbf{4}) \\
&\qquad = (0-8+12, 0-3+4) \\
&\qquad = (4,1)
\end{aligned}$$

These all give the right answer. Alternatively, we could have realised that all other consonant intervals can be constructed from our basis vectors by doing **head-to-tail**[5] addition of vectors between points on the scale, so the answers would have had to come out right anyway (because we can simultaneously perform the head-to-tail additions in the 3D space and 2D space, so the answers always have to match).

5.2.4 Images and Kernels

If we tell a mathematician that we have a linear mapping from one vector space to another, there are two questions that he or she is likely to ask us about our mapping:

- What is the **image** of the mapping?

[5] "Head-to-tail" refers to adding vectors by directly adding them represented as displacements from one point to another. For example, if $p_1 = p_0 + \mathbf{x}_{01}$, $p_2 = p_1 + \mathbf{x}_{12}$ and $p_2 = p_0 + \mathbf{x}_{02}$, it follows that $\mathbf{x}_{02} = \mathbf{x}_{01} + \mathbf{x}_{12}$. p_1 is both the "head" of $\mathbf{x}_{01}$ and the "tail" of $\mathbf{x}_{12}$. See also Figure 5.1.

- What is the **kernel** of the mapping?

The image of a mapping is the set of all output values for the mapping (or function). If a point y is in the image of a function f, then there must be some value x in the domain such that $f(x) = y$.

The set of output values for our natural 3D to 2D mapping is indeed the whole of the 2D tone/semitone vector space. To show this is true it is enough to choose a basis for the output space, and show that all the basis vectors are in the image.

For example, we can choose $(-3, 2, 0)$ as an input value that maps to $(1, 0)$, i.e. 1 tone, and $(4, -1, -1)$ as an input value that maps to $(0, 1)$, i.e. 1 semitone, as checked by the following calculations:

$$\begin{aligned} f(-3,2,0) &= (-3\cdot\mathbf{5}+2\cdot\mathbf{8}+0\cdot\mathbf{12}, -3\cdot\mathbf{2}+2\cdot\mathbf{3}+0\cdot\mathbf{4}) \\ &= (-15+16+0, -6+6+0) \\ &= (1,0) \end{aligned}$$

$$\begin{aligned} f(4,-1,-1) &= (4\cdot\mathbf{5}-1\cdot\mathbf{8}-1\cdot\mathbf{12}, 4\cdot\mathbf{2}-1\cdot\mathbf{3}-1\cdot\mathbf{4}) \\ &= (20-8-12, 8-3-4) \\ &= (0,1) \end{aligned}$$

(These choices correspond to asserting that a tone is equal to a ratio of 9/8 and a semitone is equal to 16/15, but these ratios are not the only possible choices. Two alternative choices are tone = 10/9 and semitone = 27/25. We will analyse the mathematics behind this non-uniqueness shortly.)

Having found input vectors that map to all the basis vectors, we can use them to construct input vectors that map to any output vector according to the construction of that output vector from the basis vectors. For any vector (n, m) in the 2D space, the vector $n(-3, 2, 0) + m(4, -1, -1) = (-3n + 4m, 2n - m, -m)$ is mapped by f onto (n, m).

We can conclude that our 3D to 2D mapping is an **onto** mapping, which means that its image is the whole of the codomain.

We can find an input value for any chosen output value, but this choice of input value is not unique. Uniqueness of input values for output values is exactly what the concept of **kernel** is about. The domain of our natural mapping has 3 dimensions, and the image has 2 dimensions. The theory of vector spaces tells us that we can do some simple arithmetic on these dimensions:

$$3 - 2 = 1$$

to conclude that our mapping has a kernel of dimension 1. So what is the kernel? It is the **subspace**[6] of the domain which is reduced to zero by the mapping. It's the 1 dimension that we "lose" as we go from 3 dimensions to 2 dimensions. The kernel is a measure of the non-uniqueness of input values for a given output value. If two input values differ by an element of the kernel, then they will map to the same output value. Conversely, if two input values map to the same output value, their difference must belong to the kernel.

The kernel of our natural 3D to 2D mapping has 1 dimension, so we must be able to find a basis for this subspace consisting of a single unit vector such that all elements of the kernel are multiples of that unit vector. If the unit vector is (x, y, z), then x, y and z must satisfy the following equations:

$$\begin{array}{ccccccc} 5x & + & 8y & + & 12z & = & 0 \\ 2x & + & 3y & + & 4z & = & 0 \end{array}$$

One solution is: $x = -4, y = 4, z = -1$, i.e. the vector $(-4, 4, -1)$ maps to zero. And to check:

$$\begin{aligned} f(-4,4,-1) &= (-4 \cdot \mathbf{5} + 4 \cdot \mathbf{8} - 1 \cdot \mathbf{12}, -4 \cdot \mathbf{2} + 4 \cdot \mathbf{3} - 1 \cdot \mathbf{4}) \\ &= (-20 + 32 - 12, -8 + 12 - 4) \\ &= (0, 0) \end{aligned}$$

The most general solution is $x = -4t$, $y = 4t$ and $z = -t$ for arbitrary t. So if $\mathbf{y}$ is an output value, and $\mathbf{x}$ is an input value such that $f(\mathbf{x}) = \mathbf{y}$, then $f(\mathbf{x} + t(-4, 4, 1)) = \mathbf{y}$ for any number t. However, if we restrict ourselves to 3D vectors that can be constructed by adding together 3D representations of intervals between notes on the diatonic scale, the components of $\mathbf{x}+t(-4, 4, 1)$ will always be integers, and if the components of $\mathbf{x}$ are all integers, the only way this can happen is if t is an integer. (Because if t is not an integer, then the last component of $t(-4, 4, 1) = (-4t, 4t, t)$ will not be an integer.)

$(-4, 4, -1)$ corresponds to the fractional ratio $2^{-4} \times 3^4 \times 5^{-1} = 81/80$. The zero vector $(0, 0, 0)$, which also maps to zero, corresponds to $2^0 \times 3^0 \times 4^0 = 1$. So the statement that $(-4, 4, -1)$ generates the kernel of the natural 3D to 2D mapping in effect tells us that 81/80 represents an interval of 0 semitones, and that in some sense 81/80 is the same as 1. Now 81/80 could perhaps be considered close to 1, but it's definitely not equal to 1. In fact if we listen to two notes whose frequencies have a ratio of 80 to 81, we will be able to tell the difference, as they will separated by about 22% of a semitone.

[6] A **subspace** of a vector space is a vector space consisting of a subset of the vectors in the original vector space with the same operations of addition and scalar multiplication defined. Any sum of two vectors in the subspace must also be in the subspace, and any scalar multiple of a vector in the subspace must be in the subspace.

The value of 81/80 is an example of what is called a **comma**; in particular it is called the **syntonic comma**, **Ptolemaic comma** or **comma of Didymus**. In general a comma is something that happens when we define musical scales according to rules that require one note (or more than one note) to be in two places at once. The comma is the interval between the two different places where the note wants to be.

We can summarise all this analysis of the relationships between consonant intervals and intervals on the diatonic scale as follows:

> The kernel of the natural linear mapping from the 3D representation of music intervals, as natural ratios based on powers of 2, 3 and 5, to the 2D representation of musical intervals, as sums of tones and semitones, is the vector space generated by the vector representing the ratio of the syntonic comma, which is equal to 81/80.

5.2.5 Visualising the Syntonic Comma

There are many sequences of intervals (between notes on the diatonic scale) that we can follow to realise the syntonic comma, and "prove" that 81 = 80. One example is (with all steps going up in frequency):

$$\begin{array}{ccccccccc} C & \rightarrow & G & \rightarrow & D & \rightarrow & A & \rightarrow & C \\ & 3/2 & \times & 3/2 & \times & 3/2 & \times & 6/5 & \end{array}$$

versus

$$\begin{array}{ccccc} C & \rightarrow & C & \rightarrow & C \\ & 2 & \times & 2 & \end{array}$$

By equating these two paths, which both start at a C and arrive at a C two octaves higher, we get $162/40 \approx 4$, which reduces to $81/80 \approx 1$.

But this is not the only path. For example, we could replace the first path with:

$$\begin{array}{ccccccccccc} C & \rightarrow & E & \rightarrow & B & \rightarrow & D & \rightarrow & F & \rightarrow & C \\ & 5/4 & \times & 3/2 & \times & 6/5 & \times & 6/5 & \times & 3/2 & \end{array}$$

Reconciling with the second path above this tells us $1620/400 \approx 4$, again reducing to $81/80 \approx 1$.

To find the full set of such paths that can tell us $81/80 \approx 1$, we need to start by categorising all pairs of notes whose frequencies are related to each other by simple fractional ratios.

We have already noted that all consonant ratios between notes on the diatonic scale come from intervals of 0, 3, 4, 5, 7, 8, 9 semitones, or from one of those values with a multiple of 12 semitones (i.e. 1 octave) added on. We also noted that the consonances of x semitones and $12 - x$ semitones are directly related. For example, A → C is 3 semitones which is the approximate ratio 6/5, and C → A is $12 - 3 = 9$ semitones which is the approximate ratio $2 \div 6/5 = 5/3$. Independently of the operations of adding octaves to intervals and subtracting intervals from octaves, there are only 3 distinct consonant intervals. We can choose one representative from each of the pairs 3 and 9, 4 and 8, and 5 and 7.

It will turn out to be simpler to choose 3 (minor third), 4 (major third) and 7 (perfect fifth). The reasons for this are as follows:

- Every interval of 3 or 4 semitones is 2 steps on the diatonic scale, and conversely, every interval between notes separated by 2 steps on the diatonic scale is either 3 or 4 semitones.
- Every interval of 7 semitones is 4 steps on the diatonic scale, which in all cases divides up into parts of 2 steps plus 2 steps, where one part is 3 semitones and the other is 4 semitones.
- Every interval between a pair of notes on the diatonic scale separated by 4 steps is either an interval of 7 semitones that divides up into intervals of 3 and 4 semitones as just stated, or it is the dissonant interval of 6 semitones that divides up into two portions of 3 semitones each.

It follows that we can visualise all consonant intervals and the relationships between them by stepping along the white notes scale two notes at a time, i.e. C, E, G, B, D, F, A, C. Every consonant interval is represented by either one step on this "double-stepped scale" (i.e. 2 steps on the white note scale), or by two steps (i.e. 4 steps on the white note scale). And for every case where three notes are related pairwise to each other by consonant intervals, the three notes will be found arranged in consecutive order on the double-stepped scale.

The double-stepped scale and the relationships between notes are shown in the following diagram:

$$
\begin{array}{ccccccccccccccc}
C & \Rightarrow & E & \rightarrow & G & \Rightarrow & B & \rightarrow & D & \rightarrow & F & \Rightarrow & A & \rightarrow & C \\
C & & \longrightarrow & & G & & \longrightarrow & & D & & \longrightarrow & & A & \rightarrow & C \\
C & \Rightarrow & E & & \longrightarrow & & B & & & & F & & \longrightarrow & & C
\end{array}
$$

$$
\begin{array}{lccccc}
\text{(Key:} & \rightarrow & = & \text{minor third} & = & 6/5, \\
 & \Rightarrow & = & \text{major third} & = & 5/4, \\
 & \longrightarrow & = & \text{perfect fifth} & = & 3/2 \quad)
\end{array}
$$

All possible paths from C to C that "travel along" the syntonic comma can be stepped along these lines. Starting on the left, step a minor third, a major third or a perfect fifth to the right each time, switching lines as necessary. Notice that there is no arrow from B to F on the third line—this is the dissonant 6 semitones interval.

5.3 The Harmonic Heptagon

If we take the double-stepped scale, and close it up into a loop, we get a diagram that I call the **Harmonic Heptagon**,[7] as shown in Figure 5.8.

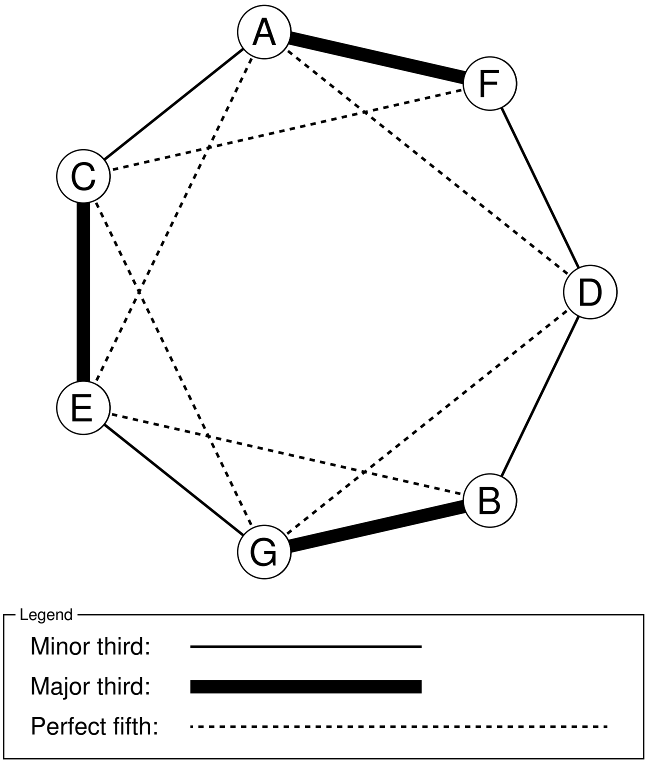

Figure 5.8. The Harmonic Heptagon

It is very easy to explain the syntonic comma in terms of the Harmonic Heptagon:

> Every time we travel once around the Harmonic Heptagon, we travel a distance in 3D space corresponding to the syntonic comma. And every time we travel a distance in 3D space corresponding to the syntonic comma, we must have travelled once around the Harmonic Heptagon.

[7] Those familiar with music theory might know about the **Circle of Fifths**. The Harmonic Heptagon looks a bit similar, but it is a different diagram. The Circle of Fifths contains all the notes (both black and white), and it only shows connections representing a perfect fifth (7 semitones $\approx 3/2$). The Harmonic Heptagon contains only the white notes, but it shows connections for all consonant intervals between those notes.

There are a few other things we might note in passing about this heptagon:

- Every note belongs to one minor third interval and one major third interval, except for D, which belongs to two minor third intervals.
- Every note belongs to two perfect fifth intervals, except for B and F which both belong to one perfect fifth interval and one dissonant 6 semitones interval.
- Every major and minor chord that can be played on the white notes consists of 3 consecutive notes going around the heptagon. These are: C major (CEG), E minor (EGB), G major (GBD), D minor (DFA), F major (FAC) and A minor (ACE). The only such triad that is not a chord is (BDF) which is a dissonant combination of notes, although this group of notes does appear as part of the G7 chord (GBDF).
- Other common chords form sequences of 4 consecutive notes on the heptagon, including G7 (GBDF), Dmin7 (DFAC), Fmaj7 (FACE), Amin7 (ACEG), Cmaj7 (CEGB) and Emin7 (EGBD).
- There is a reflective symmetry between D,F,A,C and D,B,G,E.
- Alternative home chords C major (CEG) and A minor (ACE) are opposite the D, and are symmetrical with respect to the symmetry mentioned in the previous point.

The significance of some of these points will be explained when I develop aspects of the theory in later chapters.