Three

Voice Leading, Partial Permutations and Geodesics

Counterpoint represents the melodic point of view of composition, reflecting a horizontal way of thinking. In the particular case of simultaneous motion of voices, the attention is centered on the composition of multiple, independent melodies that end up forming a sequence of chords. This choice allows to compose melodies affecting the listener both as a whole (chords) and as different autonomous fluxes of notes (parts). In the following sections, we shall focus on the formalization of the *voice leading* process, also called *part writing*, that is the evolution and the interaction of parts or voices in a sequence of chords.[1] Intuitively, we can think of it as of the assignment of a melody to a certain instrument, when more than one melody is played by more than one instrument at the same time.[2]

3.1 Defining the Voice Leading

In general, it is possible to describe a melody as a finite sequence of ordered pairs $(p_i, p_{i+1})_{i\in I}$, where I is a finite set of indices. In order to model the voice leading in a mathematical way it is necessary to introduce first the concept of *multiset*, a generalisation of the idea of set. (This approach was already considered by D. Tymoczko in Tymoczko (2006).) Roughly speaking, we can think of it as of a list where an object can appear more than once, whilst the elements of a set are necessarily unique. More formally, a *multiset* M is a couple (X, μ) composed of an *underlying set* X and a map $\mu : X \to \mathbb{N}$, called the *multiplicity* of M, such that for every $x \in X$ the value $\mu(x)$ is the number of times that x appears in M. We define the *cardinality* $|M|$ of M to be the sum of the multiplicities of each element of its underlying set X. Observe, however, that a multiset is in fact completely defined by its multiplicity function: it suffices to set $M \coloneqq (\mathrm{dom}(\mu), \mu)$.

If we interpret a set of n singing voices (or parts played by n instruments, or both) as a multiset of pitches of cardinality n, then a voice leading can be mathematically described as follows.

Definition 3.1.1. Let $M \coloneqq (X_M, \mu_M)$ and $L \coloneqq (X_L, \mu_L)$ be two multisets of pitches with same cardinality n and arrange their elements into n-tuples $(x_1, \dots, x_n)$

[1] Here the term "chord" is used in the musical sense, not necessarily as a point of the space $\mathbb{A}_n$.

[2] It is possible to think in terms of voice leading even in non-compositional contexts: for instance, a guitarist reading a partition makes a part-writing choice, deciding to play a note on a certain string. Thus we can imagine the six strings as a choir composed by six singers playing together.

and $(y_1, \dots, y_n)$ respectively.[3] A *voice leading* of n voices between M and L, denoted by $(x_1, \dots, x_n) \to (y_1, \dots, y_n)$, is the multiset

$$Z \coloneqq \{(x_1, y_1), \dots, (x_n, y_n)\},$$

whose underlying set is $X_Z \coloneqq X_M \times X_L$ and whose multiplicity function μ_Z is defined accordingly, by counting the occurrences of each ordered pair.

Remark 1. Observe that the definition just given is not linked to the particular type of object (pitches): it is possible to describe voice leadings also between pitch classes, for instance.

Note that it is also possible to describe a voice leading as a bijective map from the multiset M to the multiset L, i. e. as a partial permutation of the *union multiset*

$$M \cup L \coloneqq (X_M \cup X_L, \mu_{M \cup L}),$$

where

$$\mu_{M \cup L} \coloneqq \max\{\mu_M \chi_M, \mu_L \chi_L\}$$

and χ_M and χ_L are the characteristic functions of X_M and X_L, respectively.[4]

3.2 Partial Permutations

We recall that a *partial permutation* of a finite multiset S is a bijection between two subsets of S. In general, if S has cardinality n then this map can be represented as an n-tuple of symbols, some of which are elements of S and some others are indicated by a special symbol — we use $\diamond$ — to be interpreted as a "hole" or an "empty character". However, since we are not dealing with subsets of a fixed multiset, we shall use the cycle notation to avoid ambiguity and confusion.

Remark 2. In order to be able to do computations with partial permutations, it is fundamental to fix an *ordering* among the elements of the union multiset $M \cup L$. We henceforth give $M \cup L$ the natural ordering $\leqslant$ of real numbers, being its elements pitches. Indeed, in classical music with equal temperament, one defines the pitch p of a note as a function of the fundamental frequency ν (measured in Hertz) associated with the sound; more precisely, as the map $p : (0, +\infty) \to \mathbb{R}$ given by

$$p(\nu) \coloneqq 69 + 12 \log_2 \left(\frac{\nu}{440} \right).$$

This can be done also in the case where the elements of the union multiset are pitch classes: the ordering is induced by the ordering of their representatives belonging to a same octave. However, in this paper we shall not follow this practice and shall instead restrict to pitches only.

[3] These are in fact the images of two bijective maps $\psi_M : \{1, \dots, n\} \to M$ and $\psi_L : \{1, \dots, n\} \to L$.

[4] For a multiset S we assume that $\mu_S(x) = 0$ if $x \notin X_S$. With this understanding, the function $\mu_{M \cup L}$ is defined on the whole of $X_M \cup X_L$.

Example 3.2.1. The voice leading

$$(G_2, G_3, B_3, D_4, F_4) \to (C_3, G_3, C_4, C_4, E_4) \tag{3.2.1}$$

is described by the partial permutation of the ordered union multiset

$$(G_2, C_3, G_3, B_3, C_4, C_4, D_4, E_4, F_4)$$

defined by

$$\begin{pmatrix} G_2 & C_3 & G_3 & B_3 & C_4 & C_4 & D_4 & E_4 & F_4 \\ C_3 & \diamond & G_3 & C_4 & \diamond & \diamond & C_4 & \diamond & E_4 \end{pmatrix}. \tag{3.2.2}$$

Thus, a voice leading between two multisets of n voices can be seen as a partial permutation of a multiset whose cardinality is less than or equal to $2n$.

The next step is to associate a representation matrix with the partial permutation. Let V be an n-dimensional vector space over a field $\mathbb{F}$ and let $\mathcal{E} := \{e_1, \ldots, e_n\}$ be a basis for V. The symmetric group S_n acts on $\mathcal{E}$ by permuting its elements: the corresponding map $S_n \times \mathcal{E} \to \mathcal{E}$ assigns $(\sigma, e_i) \mapsto e_{\sigma(i)}$ for every $i \in \{1, \ldots, n\}$. We consider the well-known *linear representation* $\rho : S_n \to \mathrm{GL}(n, \mathbb{F})$ of the group S_n given by

$$\rho(1\ i) := \begin{pmatrix} 0 & & & & 1 & & & \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & & & & \\ 1 & & & & 0 & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix},$$

where the 1's in the first row and in the first column occupy the positions $1, i$ and $i, 1$ respectively. The map ρ sends each 2-cycle of the form $(1\ i)$ to the corresponding permutation matrix that swaps the first element of the basis $\mathcal{E}$ for the i-th one. Note that each row and each column of a permutation matrix contains exactly one 1 and all its other entries are 0. Following this idea and (Horn and Johnson, 1991, Definition 3.2.5, p. 165), we say that a matrix $P \in \mathrm{Mat}(m, \mathbb{R})$ is a *partial permutation matrix* if for any row and any column there is at most one non-zero element (equal to 1). When dealing with a voice leading $M \to L$, the dimension m of the matrix P is equal to the cardinality of the multiset $M \cup L$.

Remark 3. In general, the partial permutation matrix associated with a given voice leading is not unique. This is due to the fact that we are dealing with multisets: if $M \to L$ is a voice leading it is possible that some components of L have the same value, i. e. that different voices are playing or singing the same note.

For this reason we introduce the following convention.

Convention 1. Let $M := (x_1, \ldots, x_n) \to L := (y_1, \ldots, y_n)$ be a voice leading and suppose that more than one voice is associated with a same note of L. To this end, let $(x_{i_1}, \ldots, x_{i_k})$ be the pitches of M (with $i_1 < \cdots < i_k$) that are mapped to the pitches $(y_{j_1}, \ldots, y_{j_k})$ of L, with $y_{j_1} = \cdots = y_{j_k}$ and $j_1 < \cdots < j_k$. In order to uniquely associate a partial permutation matrix $P := (a_{ij})$ with the above voice leading, we assign the value 1 to the corresponding entries of P by following the order of the indices, that is by setting $a_{i_1 j_1} = 1, \ldots, a_{i_k j_k} = 1$.

Thus, we shall henceforth speak of *the* partial permutation matrix associated with a given voice leading.

Example 3.2.2. The partial permutation matrix associated with the cycle representation (3.2.2) of voice leading (3.2.1) is

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Therefore, if $M \to L$ is a voice leading, if both M and L are thought of as ordered tuples and if P is its partial permutation matrix, we have that $PM = L$; in addition, the "reversed" voice leading $L \to M$ is obviously described by the transpose P^{T} of P: $P^{\mathsf{T}} L = M$.

This representation has the advantage of providing objects that are much handier than a multiset of couples, speaking in computational terms. Algorithm algorithm 3.1 presents the pseudocode for the computation of the partial permutation matrix of a voice leading.

Algorithm 3.1 Computing the partial permutation matrix.

Input:
$M \to L$ ▷ Source (M) and target (L) multisets describing the voice leading
Output:
P ▷ Partial permutation matrix associated with the voice leading

Evaluate multiplicities of all $x \in M$ and all $y \in L$;
Generate the *ordered* multiset $U := M \cup L$;
Initialise $P \in \mathrm{Mat}(|U|, \mathbb{R})$ by setting $P(i,j) = 0$ for all i, j;
for $i, j \in \{1, \ldots, |U|\}$ **do**
  if $U(i) \to U(j)$ **then**
    $P(i,j) = 1$
  end if
end for

3.3 Voice leading and piecewise geodesic paths

We can imagine a voice leading of n voices as a sequence of n-dimensional vectors (points in $\mathbb{R}^n$), whose components are the pitches associated with each note *sung* by each voice. An important feature of this visualisation is that the melody of a certain voice is always represented by the *same* coordinate (say the ith one) in every vector of the sequence: we can thus read it very simply by looking at the projections

$$\begin{aligned}\pi_i : \mathbb{R}^n &\to \mathbb{R}\\ (v_1, \dots, v_n) &\mapsto v_i \text{ for } i = 1, \dots, n.\end{aligned}$$

A useful way to represent this idea is to take an oriented segment joining two consecutive points u and v of $\mathbb{R}^n$, that is a path $\gamma : [0,1] \to \mathbb{R}^n$ given by

$$\gamma(s) \coloneqq u + s(v-u). \tag{3.3.1}$$

Note that this is just a convenient graphical tool and does not mean at all that every point constituting the path is effectively "played": the only ones that are involved in the melody are the endpoints $\gamma(0) = u$ and $\gamma(1) = v$.

The main characteristic of the path presented just above is that it is a *geodesic* between the points u and v, being the n-dimensional Euclidean space flat. There are infinitely many ways to connect two points in $\mathbb{R}^n$, and we are not interested in the particular way they are joined. Anyway, it makes sense to set the convention that they be linked in the simplest way possible; this choice will bring advantages also in the following, as the reader will see.

If we iterate this process for each note and for each voice we obtain a polygonal chain in $\mathbb{R}^n$, which is *not* a geodesic but rather a *piecewise geodesic*. This is not surprising and in fact quite desirable, because if we considered a melody of more than 2 notes (per voice) and if we joined the endpoints with a segment, then we would lose all the information between the two, that is we would erase the melody itself! For this reason it is meaningful to consider a *concatenation of geodesics*, which allows to reproduce every step of the music.

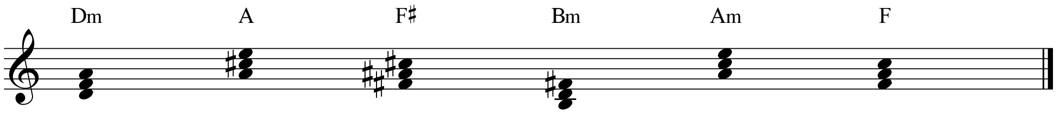
This is actually the geometric representation of what has been presented above in the algebraic form through partial permutation matrices. Indeed, if we regard a melody as a finite sequence of points (say k) in $\mathbb{R}^n$, with n the number of voices, then we can describe it geometrically through a piecewise linear path and algebraically as the product $P_k \cdots P_1$, where P_i is the partial permutation matrix of the i-th voice leading. As an example, consider the progression of triads in Figure figure 3.1a on the next page: each of them is represented as a triple (p_1, p_2, p_3) in $\mathbb{R}^3$, with $p_1 < p_2 < p_3$. In general it is possible to build a voice leading by associating each note of a given chord with a note of the following one, respecting the order induced by $<$. This rule has been used to draw the path in figure 3.1b.

Let us now consider the four voice leadings

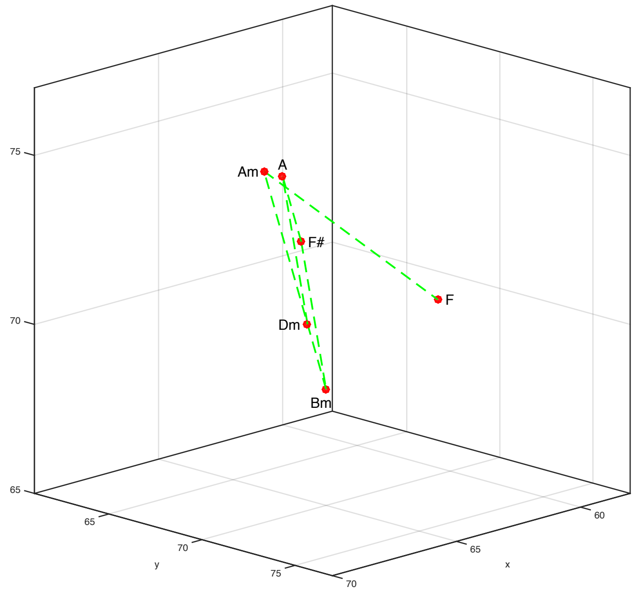
$$\begin{aligned}(B_3, F_5) &\to (C_4, E_4) \quad \text{and} \quad (F_5, B_3) \to (E_4, C_4),\\ (B_3, F_5) &\to (E_4, C_4) \quad \text{and} \quad (F_5, B_3) \to (C_4, E_4),\end{aligned}$$

depicted in Figure figure 3.2 on page 45 (black, red, green and blue arrow respectively): from the musical and perceptive viewpoint they are completely equivalent in pairs (each row describes the same voice leading). Generalising this fact to n voices, it is natural to identify the paths in $\mathbb{R}^n$ that are symmetric with respect to the diagonal of this space.

An immediate generalisation of this situation leads to the conclusion that if we apply the *same* permutation to both the endpoints of the paths representing a voice leading,



(a) Voice leading among triads in root position.



(b) Graphic representation in $\mathbb{R}^3$ of the above voice leading.

Figure 3.1: Voice leading and corresponding piecewise geodesic path.

then we obtain in actual fact the *same* voice leading. This discussion translates into the following proposition.

Proposition 3.3.1. *The moduli space of a voice leading of n voices is isomorphic to the quotient $(\mathbb{R}^n \times \mathbb{R}^n)/\Delta_{S_n}$ of the $2n$-dimensional linear space $\mathbb{R}^n \times \mathbb{R}^n$ with respect to the diagonal action Δ_{S_n} of the symmetric group S_n.*

Proof. Let $u, v \in \mathbb{R}^n$ and let $u \mapsto v$ be the voice leading between them, parametrised[5] by path (3.3.1), that needs $2n$ real parameters to be defined. If $\sigma \in S_n$ is a permutation acting on the canonical base of $\mathbb{R}^n$, a necessary and sufficient condition for the voice leading to be invariant under σ is that the action of σ be diagonal on the product. □

Note that in the case $n = 1$ the moduli space is isomorphic to $\mathbb{R}^2$, for we have that $S_1 = \{\mathrm{id}\}$.

The above discussion about symmetry, points out that it makes sense to represent a voice leading of n voices as a geodesic on the Riemannian manifold with boundary $\mathbb{R}^n/S_n$. In the special case $n = 2$ the space $\mathbb{R}^2/S_2$ is isomorphic to the half-plane

$$\mathbb{H} \coloneqq \{(x, y) \in \mathbb{R}^2 \mid x \leqslant y\}.$$

[5]The parametrisation holds modulo the action of a uniform permutation, i. e. a permutation acting on both u and v in the same way.

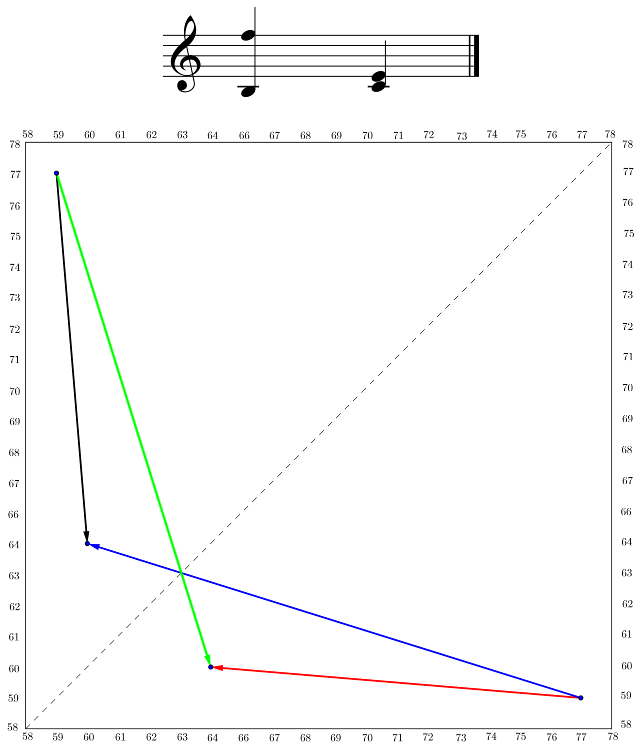


Figure 3.2: All possible voice leadings between the notes of the score depicted above. Observe the symmetric nature of the paths with respect to dashed line ($y = x$).

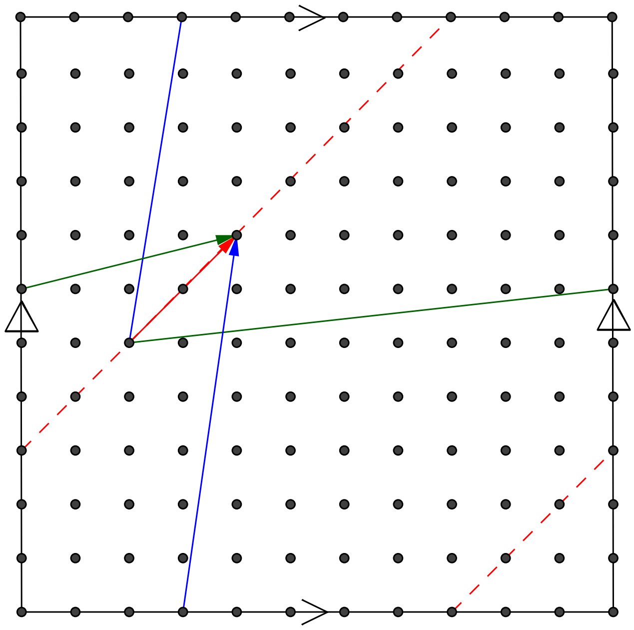
Figure 3.2 shows the voice leading between two dyads in $\mathbb{R}^2$, the universal covering space of $\mathbb{H}$.

It possible to represent and to analyse voice leadings as paths on more harmony-oriented spaces than $\mathbb{R}^n$, such as the pitch class space $\mathbb{T}^n$ or the chord space $\mathbb{A}_n$. From the harmonic point of view it is indeed admissible to ignore the octave which a certain note of a chord belongs to, and in the same fashion it is possible to identify each chord with the whole set of its possible voicings. We are therefore interested in geodesics on these spaces, as they will be the representation of voice leading also in this fairly general setting. The paths that we are seeking will be easily obtained once we note that $\mathbb{T}^n$ and $\mathbb{A}_n$ are obtained as identification spaces from $\mathbb{R}^n$. Therefore it suffices to draw the segments connecting the endpoints of the voice leading in $\mathbb{R}^n$, just like before, and then project them via the covering map that gives rise to the desired space. Here are some illustrated examples.

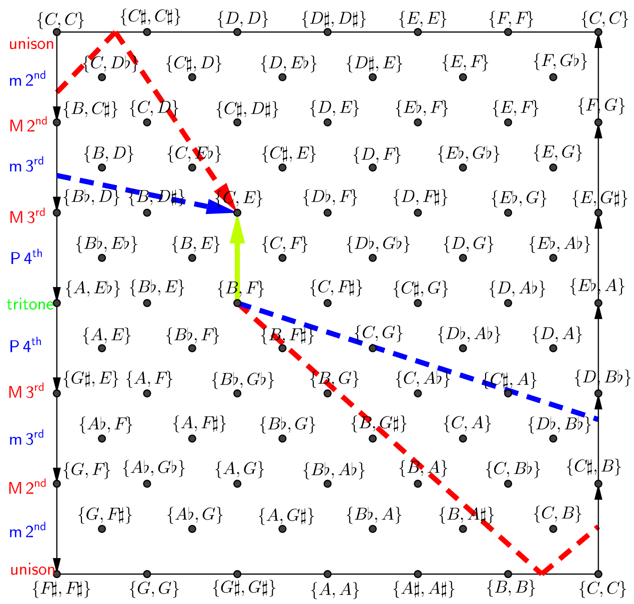
Example 3.3.1 (Voice Leading on $\mathbb{T}^2$)**.** We show how the generators of the torus can be used to retrieve information about possible *octave leaps* of one or more voices (which *a priori* is lost, since we have identified all the octaves with each other). Let us consider for this purpose the following four voice leadings in $\mathbb{R}^2$:

$$\text{i) } (D_0, F_0) \to (E_0, G_0), \qquad \text{ii) } (D_0, F_0) \to (E_1, G_0),$$
$$\text{iii) } (D_0, F_0) \to (E_0, G_1), \qquad \text{iv) } (D_0, F_0) \to (E_1, G_1).$$

They all represent the same voice leading $(D, F) \to (E, G)$ in the pitch class space, but their path realisations are different: Figure figure 3.3a on the next page displays



(a) Geodesics and octave leaps in $\mathbb{T}^2$.



(b) Geodesics in $\mathbb{A}_2$.

Figure 3.3: Voice leading paths in the pitch class space $\mathbb{T}^2$ and in its relative chord space $\mathbb{A}_2$.

these four paths on $\mathbb{T}^2$ (represented as a square with the usual identification rule on its sides expressed by the symbols $>$ and $\triangle$):

Path i) is drawn as the shortest red arrow, since the jump between the dyads does not actually exceeds the 0-th octave;

Path ii) is represented by the green arrow, exiting the square from its right side and coming back in from its left side: this reflects the fact that the first voice makes a leap of one octave;

Path iii) is associated with the blue arrow, pointing to the top and re-entering the figure from the bottom: in this case it is the first voice to jump to the next octave;

Path iv) is rendered by the dashed red arrow: it jumps from the top to the bottom and then from the right to the left of the square, because both voices exceed the 0-th octave.

Example 3.3.2 (Voice Leading on $\mathbb{A}_2$)**.** When we identify each chord with its musical inversions we are operating in the chord space, i. e. in the orbifold $\mathbb{A}_n := \mathbb{T}^n/S_n$. In the 2-dimensional case of dyads $\mathbb{T}^2/S_2$ is the *Möbius strip* (see Bergomi et al. (2014) for details about the construction and the positioning of the chords on the lattice). Figure 3.3b shows three different geodesic paths corresponding to the voice leading $\{B, F\} \to \{C, E\}$, where the curly braces mean that we are identifying all possible assignments of parts to each voice. Observe the identification of the left and right side of the square with inverted orientation and note the *singular boundary* of the unisons, constituted by the upper and lower side of the square. The paths bounce back when touching it because of the quotient operation by the symmetric group. Here it is clear that harmony is favoured over melody, because neglecting both octaves and ordering leads to focus on the *ensemble* of voices.

These examples share and show one important feature: the shortest paths joining the two pairs of pitch classes ($\mathbb{T}^2$) or dyads ($\mathbb{A}_2$), i. e. the *minimal* geodesics between those two points, represent voice leading *with neither crossing nor octave leaps*, whilst the paths that touch the singular boundary correspond to part writings where at least one of these phenomena occurs.

Although the voice crossing is not advised as a standard practice in Harmony manuals, it is a useful technique to avoid repeated notes, parallel fifths and hidden octaves and to assure a high degree of independence to each voice. For further details see Prout (2012); Boland and Link (2012); Russo (1997); Sussman and Abene (2012); Notley (2007).

3.4 Simultaneous motions of the voices and complexity of a voice leading

We have seen in the previous section how the partial permutation matrix associated with a voice leading contains the information about the passage from one note to the next one for each voice. Here we are going to illustrate that, in fact, the tool that we have built also encodes the direction of motion of the different voices, including the crossings.

On the one hand, in Music one distinguishes between three main behaviours (cf. Figure 3.4; we omit parallel motion because that is not involved in our analysis):

- *Similar motion*, when the voices move in the same direction;

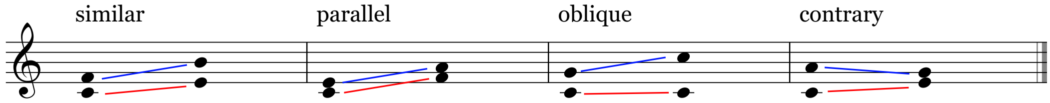


Figure 3.4: Motion classes for two voices. *Similar:* same direction but different intervals; *parallel:* same direction and same intervals; *oblique:* only one voice is moving; *contrary:* opposite directions.

- *Contrary motion*, when the voices move in opposite directions;
- *Oblique motion*, when only one voice is moving.

On the other hand, with reference to a partial permutation matrix (a_{ij}), it is possible to describe the motion of a voice by noting three conditions, which are immediate consequences of the ordering of the union multiset:

1) If there exists an element $a_{ij} = 1$ for $i < j$ then the i-th voice is moving "upwards";
2) If there exists an element $a_{ij} = 1$ for $i > j$ then the i-th voice is moving "downwards";
3) If there exists an element $a_{ii} = 1$ then the i-th voice is constant.

The connection between the two worlds is the following:

- If either Condition 1) or Condition 2) is verified by two distinct elements then we have *similar motion*;
- If both Condition 1) and Condition 2) hold for two distinct elements then we are facing *contrary motion*;
- The case of *oblique motion* involves Conditions 1) and 3) or Conditions 2) and 3), for at least two distinct elements.

As we mentioned in section 1.1, *voice crossing* is a particular case of these motions where the voices swap their relative positions. This phenomenon can be described in terms of multisets as follows.

Definition 3.4.1. Let $(x_1, \dots, x_n) \to (y_1, \dots, y_n)$ be a voice leading ($n \in \mathbb{N}$). If there exist two pairs (x_i, y_i) and (x_j, y_j) such that $x_i < x_j$ and $y_i > y_j$ or such that $x_i > x_j$ and $y_i < y_j$ then we say that a *(voice) crossing* occurs between voice i and voice j.

The partial permutation matrix retrieves even this information, as the following proposition shows.

Proposition 3.4.1. *Consider a voice leading of n voices and let $P := (a_{ij})$ be its associated partial permutation matrix. Choose indices $i, j, k, l \in \{1, \dots, n\}$ such that $a_{ij} = 1$ and $a_{kl} = 1$. Then there is a crossing between these two voices if and only if one of the following conditions hold:*

i) $i < k$ and $j > l$;

ii) $i > k$ and $j < l$.

Furthermore, the total number of voices that cross the one represented by a_{ij} is equal to the number of 1's in the submatrices (a_{rs}) and (a_{tu}) of P determined by the following restrictions on the indices: $r > i$, $s < j$ and $t < i$, $u > j$.

Proof. In a partial permutation matrix the row index of a non-zero entry denotes the initial position of a certain voice in the ordered union multiset, whereas the column index of the same entry represents its final position after the transition. It is then straightforward from Definition 3.4.1 that for a voice crossing to exist either condition i) or condition ii) must be verified. Every entry a_{kl} satisfying one of those conditions refers to a voice that crosses the one represented by a_{ij}, hence the number of crossings for a_{ij} equals the amount of 1's in positions (r, s) such that $r > i$ and $s < j$, summed to the number of 1's in positions (t, u) such that $t < i$ and $u > j$. $\square$

Remark 4. The fact that the number of crossings with a given voice equals the number of 1's in the submatrices determined by the entry corresponding to that voice (as explained in the previous proposition) holds true only because we assumed Convention 1. Indeed, if we did not make such an assumption, the submatrices could contain positive entries referring to voices ending in the same note but that do not produce crossings.

From what we have shown thus far it emerges that it is possible to characterise a voice leading by counting the voices that are moving upwards, those that are moving downwards, those that remain constant and the number of crossings. We summarise these features in a 4-dimensional *complexity vector* c defined by

$$c := (\#\text{upward voices}, \#\text{downward voices}, \#\text{constant voices}, \#\text{crossings}), \quad (3.4.1)$$

so that we are now able to classify and distinguish voice leadings by simply looking at these four aspects.

Example 3.4.1. *Similar motion.* The voice leading $(C_1, E_1, G_1) \to (D_1, F_1, A_1)$ is represented by

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and its complexity vector is $(3, 0, 0, 0)$.

Oblique motion. The voice leading $(G_2, G_2, C_3) \to (C_3, C_3, C_3)$ is associated with

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and its complexity vector is $(2, 0, 1, 0)$.

Voice crossing. The voice leading $(C_1, E_1, G_1) \to (G_1, C_1, E_1)$ is represented by

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

and its complexity vector is $(1, 2, 0, 2)$.

By virtue of these tools it is straightforward to analyse an entire first specie counterpoint: it is enough to divide it into pairs of notes for each voice and apply the procedure described above for each passage. The concatenation of all the consecutive passages results then in a sequence of partial permutation matrices, whence one can extract a sequence of complexity vectors. This last piece of information can be visualised as a set of points in a 4-dimensional space — or rather as one or more of its 3-dimensional projections (see Subsection 3.5). In fact, if one wants to represent the complexity of the whole composition as a point cloud, one should take into account that different matrices can produce the same complexity vector; therefore we have a *multiset* of points in $\mathbb{R}^4$ (with non-negative integer components).

3.5 Complexity analysis of two *Chartres Fragments*

We are going to analyse two pieces that are parts of the *Chartres Fragments*, an ensemble of compositions dating back to the Middle Ages: *Alleluia, Angelus Domini* and *Dicant nunc Judei*; both of them are counterpoints of the first species and involve only two voices. The musical interest in these compositions consists in the introduction of a certain degree of independence between the voices and the use of a *parsimonious voice leading*, i. e. an attempt to make the passage from a melodic state to the next as smooth as possible. Note how the independence of the voices is reflected by the presence of contrary motions and crossings, which can then be interpreted as a rough measure of this feature. For a complete treatise on polyphony and a historical overview we refer the reader to Taruskin (2009).

In what follows, we represent the multiplicity of each complexity vector c as a circle of centre $c \in \mathbb{R}^4$ and radius equal to the *normalised multiplicity* $\mu(c)/n$ of c, where $\mu(c)$ is the number of occurrences of c in the analysed piece and n is the total number of notes played or sung by each voice in the whole piece.

Alleluia, Angelus Domini. The fragment under examination is depicted in Figure 3.5; here is the list of its first four voice leadings, as they are generated by the pseudocode described in Algorithm 3.1:

```
Voice Leading: ['F4', 'C4'] ['G4', 'D4']
[2, 0, 0, 0] - similar motion up
Voice Leading: ['G4', 'D4'] ['A4', 'E4']
[2, 0, 0, 0] - similar motion up
Voice Leading: ['A4', 'E4'] ['G4', 'F4']
[1, 1, 0, 0] - contrary motion
Voice Leading: ['G4', 'F4'] ['F4', 'G4']
[1, 1, 0, 1] - contrary motion - 1 crossing
```

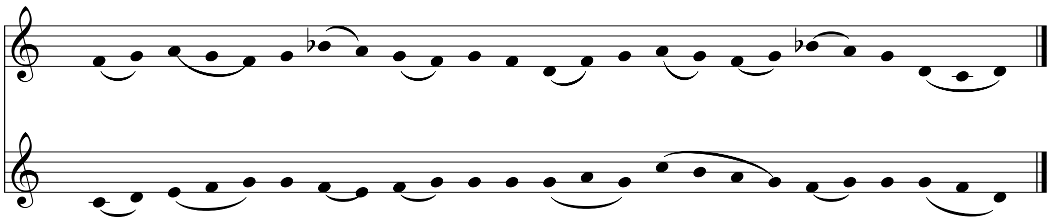
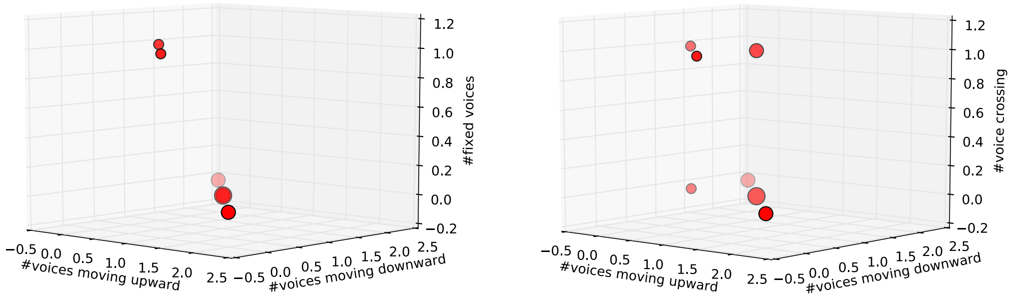


Figure 3.5: *Alleluia, Angelus Domini*, Chartres fragment n. 109, fol. 75.



(a) Projection neglecting the *crossing* component of the complexity vectors.

(b) Projection neglecting the *constant voices* component of the complexity vectors.

Figure 3.6: Three-dimensional projections of the complexity cloud of the paradigmatic voice leading *Alleluia, Angelus Domini*. The radius of each circle represents the normalised multiplicity of the corresponding complexity vector.

Table table 3.1 on the following page contains the the complexity vectors and their occurrences in the piece; the point cloud associated with this multiset is represented in Figure 3.6. Observe how the projection that neglects the component of c corresponding to the number of constant voices (Figure 3.6b) gives an immediate insight on the relevance of voice crossing in the piece.

Dicant nunc Judei. The first part of the output of Algorithm 3.1 produces the following analysis:

```
Voice Leading: ['F4', 'C4'] ['G4', 'E4']
[2, 0, 0, 0] - similar motion up
Voice Leading: ['G4', 'E4'] ['F4', 'D4']
[0, 2, 0, 0] - similar motion down
Voice Leading: ['F4', 'D4'] ['E4', 'C4']
[0, 2, 0, 0] - similar motion down
Voice Leading: ['E4', 'C4'] ['D4', 'D4']
[1, 1, 0, 1] - contrary motion - 1 crossing
```

The complexity vectors arising in the whole piece and their multiplicities are again collected in Table 3.1; see Figure 3.8 instead for a visualisation of the point cloud describing the piece. Note how the voice crossing is more massive than in the point cloud describing *Alleluia, Angelus Domini*. In addition, the point $(0, 0, 0)$ in

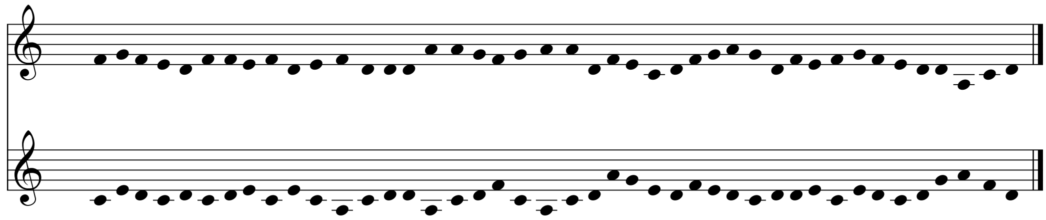
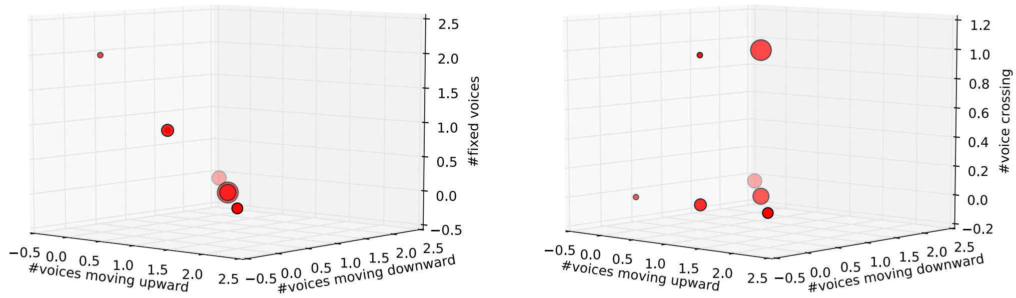


Figure 3.7: *Dicant nunc Judei*, Chartres fragment.



(a) Projection neglecting the *crossing* component of the complexity vectors.

(b) Projection neglecting the *constant voices* component of the complexity vectors.

Figure 3.8: Three-dimensional projections of the complexity cloud of the paradigmatic voice leading *Dicant nunc Judei*. The radius of each circle represents the normalised multiplicity of the corresponding complexity vector.

Table 3.1: Complexity vectors of the analysed fragments and their occurrences.

Alleluia, Angelus Domini	
c	$\mu(c)$
$(0,1,1,0)$	2
$(0,1,1,1)$	2
$(0,2,0,0)$	4
$(1,0,1,1)$	2
$(1,1,0,0)$	6
$(1,1,0,1)$	4
$(2,0,0,0)$	4

Dicant nunc Judei	
c	$\mu(c)$
$(0,0,2,0)$	1
$(0,2,0,0)$	7
$(1,0,1,0)$	5
$(1,0,1,1)$	1
$(1,1,0,0)$	9
$(1,1,0,1)$	15
$(2,0,0,0)$	4

Figure 3.8b corresponds to the point $(0,0,2,0) \in \mathbb{R}^4$, that represents trivial voice leadings where both parts do not vary.

Our analysis showed that our definition of complexity in terms of the relative movements of the voices and especially of crossing is suitable for characterising a musical piece. Point-cloud representation yields a “photograph” of complexity, a sort of fingerprint that lets clearly emerge what are the main features of the examined composition, noticeable even at first glance. DTW provides then further support to this evidence by directly measuring the distance between the complexities of two pieces, and it proved itself a good tool that is able to recognise and distinguish quite well the aforementioned pictures.

Four

A Braid-oriented Representation of Voice Leadings

Partial permutations and their representation as sparse matrices are a comfortable computational tool to describe voice leadings. There is a natural link among braids described as a collection of strands connecting two copies of a set of points and a permutations acting on this set. Considering partial permutations and allowing the strands to overlap either at their starting or ending point, it is possible to visualize the information encoded by partial permutations as a particular class of singular partial braids. We shall also describe how these entities are capable of encoding additional information in their geometry, such us leaps among voices and their restriction to the analysis of *voice leadings* in a pitch-class setting.

4.1 The Braid Group and the partial singular Braid Monoid

In this section the basic background concerning braids and partial singular braids are recalled. Our basic references are (Hansen, 1989) and (East, 2007, 2010).

The Braid Group

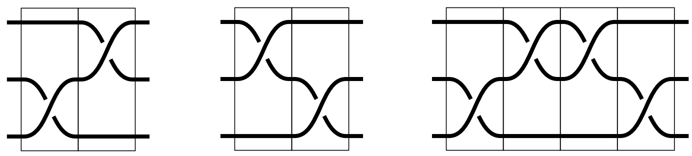
Definition 4.1.1. A *braid* β on n strands is a collection of embeddings

$$\mathscr{B} := \{\beta^\alpha ; \beta^\alpha : [0,1] \to \mathbb{R}^3,\ \alpha = 1, \dots, n\}$$

with disjoint images such that:

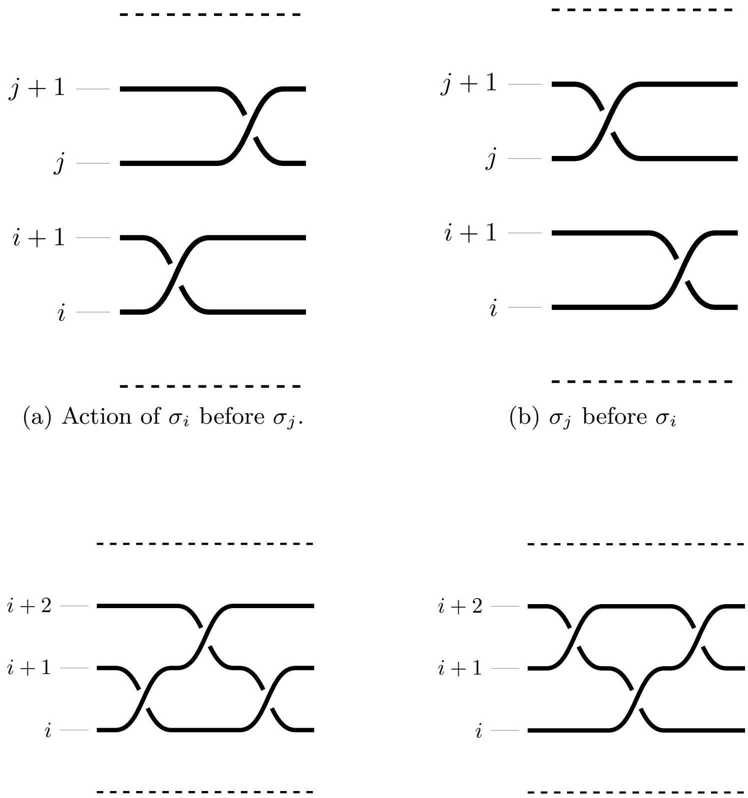
- $\beta^\alpha(0) = (0, \alpha, 0)$;
- $\beta^\alpha(1) = (1, \tau(\alpha), 0)$ for some permutation τ;
- the images of each β^α is transverse to all planes $\{x = \mathit{const}\}$, where $\mathbb{R}^3$ is equipped with the (x, y, z) coordinates.

Definition 4.1.2. Two such braids are said to lie in the same *topological braid class* if they are homotopic in the sense of braids: one can deform one braid to the other without any intersection among the strands.



(a) First braid (b) Second braid (c) Their concatenation

Figure 4.1: Concatenation of braids



(a) Action of σ_i before σ_j. (b) σ_j before σ_i

(c) $\sigma_i\sigma_{i+1}\sigma_i$ (d) $\sigma_{i+1}\sigma_i\sigma_{i+1}$

Figure 4.2: Graphical representation of the Braids properties in equations (4.1.1) and (4.1.2).

There is a natural group structure on the space of topological braids with n strands, B_n, given by concatenation. Using generators σ_i which interchanges the i-th and $(i+1)$-th strands with a positive crossing yields the presentation for B_n.

Let

$$p_1: \quad \sigma_i\,\sigma_j = \sigma_j\,\sigma_i; \quad |i-j| > 1 \tag{4.1.1}$$

$$p_2: \quad \sigma_i\,\sigma_{i+1}\,\sigma_i = \sigma_{i+1}\,\sigma_i\,\sigma_{i+1}; \quad i < n-1. \tag{4.1.2}$$

The two properties are represented in figure 4.2. We are ready to give a presentation of the group B_n as

$$B_n := \langle \sigma_1, \ldots, \sigma_{n-1} : \ p_1 \text{ and } p_2 \text{ hold} \rangle.$$

Let s_i be the permutation (i i+1), and the symmetric group $\mathcal{S}_n$ can be presented as

$$<s_1, \ldots, s_{n-1} \mid s_i s_j = s_j s_i \text{ for } |i-j| > 1,$$
$$s_i^2 = 1,\ s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1} \text{ for } 1 \leqslant i < n\ >,$$

thus the projection defined on the group of braids on n strands, on the symmetric group is given by

$$\pi : \mathcal{B}_n \to \mathcal{S}_n$$
$$\sigma_i \mapsto (i\ i+1).$$

Partial Braids and Partial Permutations

The braid group is too structured to represent voice leadings where voices can *collide* in unisons and can be rested. We introduce the inverse monoid of partial permutation and the partial singular braid monoid.

Definition 4.1.3 (Monoid)**.** A *monoid* is a couple $(S, +)$, where S is a set, $+$ is an associative binary operation and it exists $e \in S$ such that it is the identity element.

We associate a monoid to a given set in the following way.

Definition 4.1.4 (Inverse monoid)**.** Given a set S the inverse monoid $\mathcal{I}_S$ is the set of all the partial bijection of S.

It is possible to build a braid inverse monoid $\mathcal{IB}_n$ which is the analogous of the symmetric inverse monoid $\mathcal{I}_n$ of partial permutations on at most n strands. There exists an epimorphism $\mathcal{IB}_n \to \mathcal{I}_n$, extending the projection $\pi : \mathcal{B}_n \to \Sigma_n$ defined above. The epimorphism

$$\pi^* : \mathcal{IB}_n \to \mathcal{I}_n$$
$$\beta \mapsto \overline{\beta}$$

is defined by construction. The inverse braid monoid $\mathcal{IB}_n$ is the collection of partial braids with at most n strands, thus $\beta \in \mathcal{IB}_n$ induces naturally a partial permutation on n symbols. In figure 4.3 a partial braid is depicted, it is naturally associated to the partial permutation in $\mathcal{I}_5$, such that

$$\begin{array}{ccc} 2 & \mapsto & 1 \\ 3 & \mapsto & 4 \\ 5 & \mapsto & 2 \end{array}$$

The operation defined on $\mathcal{IB}_n$ is the concatenation of partial braids, given two elements of $\{\ \beta, \beta_1\ \} \subset \mathcal{IB}_n$ (see figure 4.4a), the multiplication of the two partial braids is depicted in figure 4.4b and in figure 4.4c any string fragments which do not connect the upper plane to the lower plane are removed. See (East, 2007) for further details.

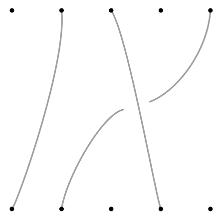
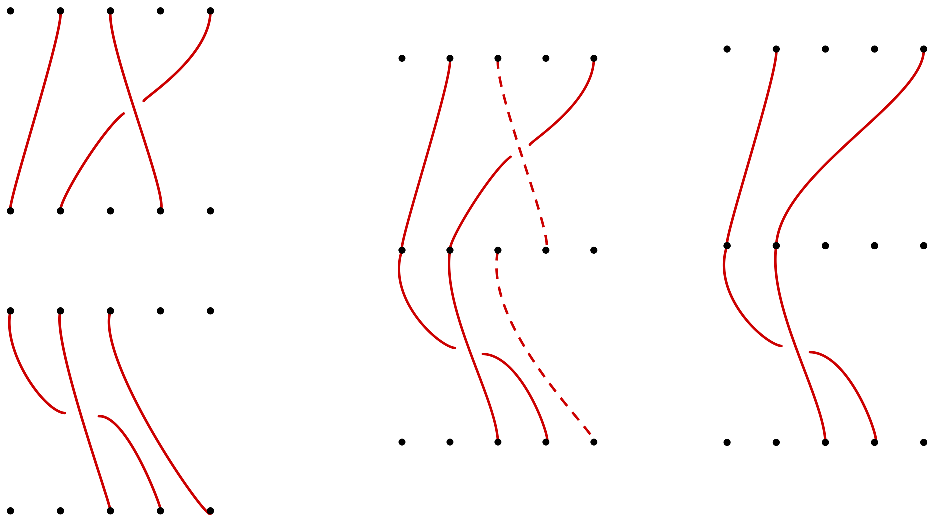


Figure 4.3: A partial braid $\beta \in \mathcal{IB}_5$



(a) Two elements β and β_1 of $\mathcal{IB}_5$.

(b) Concatenation $\beta * \beta_1$.

(c) Strands deleting rule.

Figure 4.4: Concatenation of partial braids in $\mathcal{IB}_5$.

The Singular Braids Monoid

It is possible to generalize braids to singular braids containing a finite number of singularities. It is not possible to undone a singularity, then the set $\mathcal{S}B_n$ of singular braids on n strands is not endowed with a group structure, it is a monoid and it is possible to define it as follows.

Definition 4.1.5. $\mathcal{S}B_n$ is generated by $s_1, \ldots, s_{n-1}, s_1^{-1}, \ldots, s_{n-1}^{-1}, t_1, \ldots, t_{n-1}$, due to the relations

1. $\forall i < n,\ s_i s_i^{-1} = e$.
2. For $|i-j| > 1$ the compositions $s_i s_j$, $t_i t_j$ and $s_i t_j$ commute;
3. $\forall i < n-1$

$$\begin{aligned} s_i s_{i+1} s_i &= s_{i+1} s_i s_{i+1} \\ t_i s_{i+1} s_i &= s_{i+1} s_i t_{i+1} \\ t_{i+1} s_i s_{i+1} &= s_i s_{i+1} t_{i+1}. \end{aligned}$$

Geometrically the singular generator has to be included, see figure 4.5.

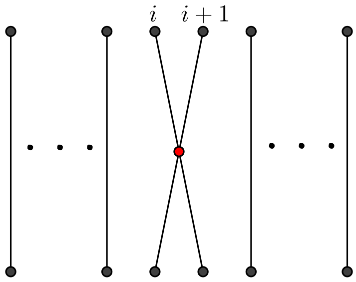


Figure 4.5: Singular generator of $\mathcal{S}B_n$.

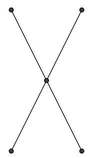


Figure 4.6: From two distinct voices to unison and back two distinct voices, singular braid representation.

Musically, a singular braid represents unisons: two voices collapsing in the same pitch will be represented as in figure 4.6.

The Partial Singular Braids Monoid

To model all of the possible leadings of voices including voices identification and rests, we need to consider the monoid of partial singular braids $\mathcal{PS}B_n$ containing either the partial and the singular braid monoid defined above. The theory about $\mathcal{PS}B_n$ will be sketched in this section, however we refer to (East, 2010) for a detailed analysis.

Intuitively an element of the partial singular braid monoid is a singular braid whose strands can be removed and the multiplication is inherited by $\mathcal{IB}_n$. An element $\beta \in \mathcal{PS}B_n$ induces a partial permutation $\bar{\beta} \in \mathcal{I}_n$. The domain of $\bar{\beta}$ corresponds to the collection of the starting points of β and symmetrically the codomain of $\bar{\beta}$ is represented by the set of ending points of each strand. Let denote these sets as $\mathrm{dom}(\bar{\beta})$ and $\mathrm{im}(\bar{\beta})$. Thus for $a \in \mathrm{dom}(\bar{\beta})$ and $b \in \mathrm{im}(\bar{\beta})$, the relation $\bar{\beta}(a) = b$ holds, if and only if it exists a strand of β connecting a to b. The operation defined on $\mathcal{PS}B_n$ is defined as an extension of the multiplication in $\mathcal{IB}_n$.

More formally, consider the set $\{ 1, \ldots, n \}$ with $n \in \mathbb{N}$. Consider a singular braid $b \in \mathcal{S}B_n$, a partial singular braid β is obtainable by removing some strands of b, observe that following the definition of partial braid given above $b \in \mathcal{PS}B_n$, and the remotion of the whole set of strands is allowed in $\mathcal{PS}B_n$. In this particular situation β is said to be a sub-braid of b.

A partial singular braid β induces a partial permutation $\bar{\beta} \in \mathcal{I}_n$ exactly as a partial braid does. Let $\{ \beta_1, \beta_2 \} \in \mathcal{PS}B_n$, the two partial singular braids are said to be equivalent if the partial permutations they induced are the same partial permutation and if $\beta_1 \subset \gamma_1$ and $\beta_2 \subset \gamma_2$ with γ_1 and γ_2 singular braids equivalent in the sense of rigid-vertex-isotopy (Birman, 1993).

The multiplication of two partial singular braids follows the same rule we introduced for partial braids (see figure 4.4). In particular denoting as $|\beta|$ the number of strings of β and as $N(\beta)$ the number of its singular points, the two submonoids of partial braids and singular braids of $\mathcal{PS}B_n$ can be described as $\mathcal{I}B_n = \{\ \beta \in \mathcal{PS}B_n \ \mid\ N(\beta) = 0\ \}$ and $\mathcal{S}B_n = \{\ \beta \in \mathcal{PS}B_n \ \mid\ |\beta| = n\ \}$.

4.2 Modelling voice leading in $\mathcal{PS}B_n$

The first approach we describe is a mere translation in the braid formalism of the model described in section 3.2. We define a voice leading of at most n voices as a partial singular braid $\beta \in \mathcal{PS}B_m$, where m is the cardinality of the underlying set of the multiset $M \cup L$ as it has been introduced in section 3.1.

The introduction of singularities allows to simplify the model described in section 3.2, since given a voice leading $v : M \to L$, where M and L are two multisets allows to identify the elements of $M \cup L$ labeled with the same symbol and encode their multiplicity in the braid's visualization as it is depicted in figure 4.7a, where the voice leading

$$(C_4, E_4, G_4) \to (D_4, F_4, F_4)$$

is represented as a singular braid.

A crossing of the voices in $\mathcal{PS}B_n$ is represented as a crossing of the strands of the partial singular braid. The voice leading

$$(C_4, E_4, G_4) \to (F_4, D_4, F_4)$$

represented in figure 4.7b shows a singular partial braid whose musical interpretation is surprisingly clear: the voice leading contains a voice crossing corresponding to the crossing of the strands and it is *singular*, since two voices collapse on their target chord in a unison.

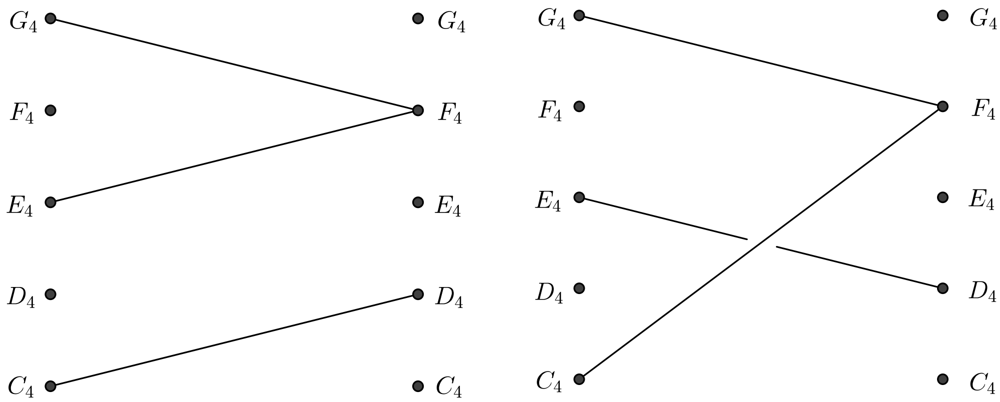
Observe that in both figures the strands $\{\ \beta^\alpha : [0,1] \to \mathbb{R}^3\ \}_\alpha$ of the braid have been chosen to be the line segment connecting the points $\beta^\alpha(0)$ and $\beta^\alpha(1)$. Such a braid is said to be piecewise linear. This particular representation is suitable for representing simultaneous motions of voices, since the slope of each strand will tell us if a voice is moving upward, downward or if it is fixed. Hence, we model voice leadings utilizing braids that are

1. *piecewise linear*: strands are geodesics in $\mathbb{R}^3$;

2. *positive*: the same sign is assigned to all crossings.

Let $s(\beta^\alpha)$ be the slope of the line segment β^α, and b^+, b^-, b^0 the number of strands of the piecewise braid with positive, negative and zero slope respectively; cr the number of crossings; $r = n - N(\beta)$, where n is the number of voices involved in the voice leading including rested ones; and finally $s = |\beta|$ the number of singularities appearing the braid associated to the voice leading. We can now rewrite the complexity vector associated to voice leadings in braid notation as

$$c = \left(b^+, b^-, b^0, cr, r, s\right)$$



(a) Unison. (b) Unison and crossing.

Figure 4.7: Partial singular braid representation of voice leadings.

Including the information coming from the topological structure of the braid concerning the crossings (geometrically represented and thus visible in this representation rather than in the one based on partial permutation matrices) and singularities. Observe that in this context the only admitted singularities are unisons.

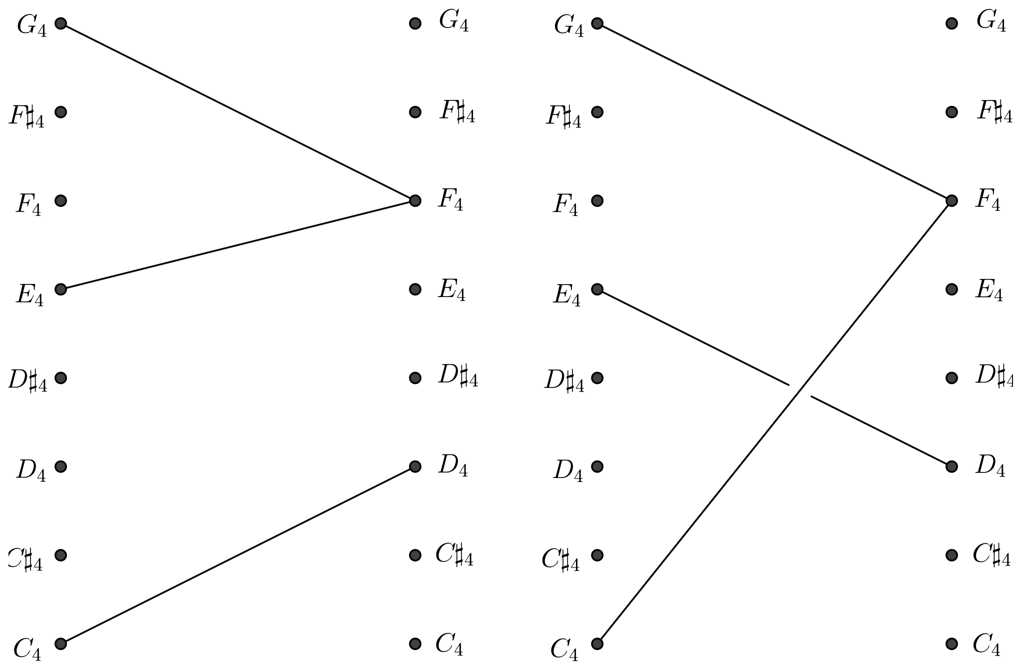
Leaps

In section 1.1, leaps are pointed out as an important feature to determine the *quality* of a voice leading. The partial permutation matrix model described in section 3.2 does not take them into account: to do that, we should define a partial permutation whose domain is made of $n \times m$ elements, where n is the number of voices involved in the voice leading and m is the number of half-steps from the lower to the higher pitch involved in the voice leading. So far, we did not introduced this representation for its high dimensionality.

Partial singular braids allow to encode this information defining a dominion of cardinality m, since singularities can be used to represent repeated voices in a chord. In this case the slope of each strands corresponds univocally to a musical interval, in figure 4.8 two piecewise linear singular partial braids encoding the leaps information are depicted. It is possible to store this information in the complexity vector either writing explicitly the slope of each strand, or splitting them into two classes of *consonant* and *dissonant* intervals, always referring to the definitions given in section 1.1, or defining them following one's particular needs.

Voice leading, partial singular braids and partial permutations

As we stated in the mathematical introduction to this section, there exists a natural projection $\pi^* : \mathcal{PS}B_n \to \mathcal{I}_n$, meaning that a class of partial singular braids β, in the sense of braid's homotopy (definition 4.1.2) and modulo the sign of the braid's crossings, describes a particular partial permutation $\bar{\beta}$ on the elements of the dominion of β in an obvious sense. As it is shown in figure 4.9 the braids β_1, β_2



(a) Voices' leaps. (b) Voices' leaps and crossing.

Figure 4.8: Partial singular braid representation of voices leaps.

and β_3 induce the same partial permutation represented by the cycle

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & \diamond & \diamond & 5 & \diamond & 4 \end{pmatrix}.$$

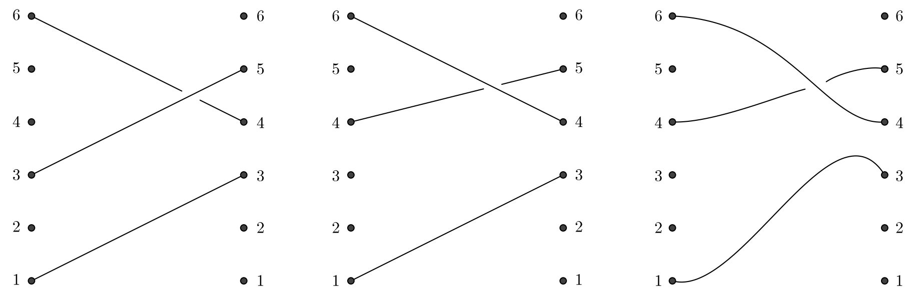
The particular choice of dealing with piecewise linear, positive, partial singular braids makes the projection a bijection, since the choice among the homotopic strands in the sense of braids is reduced to the one given by geodesic strands (which is unique in $\mathbb{R}^3$) and the crossings are always positive. Hence in this context it will always be possible to transform a partial permutation in a braid and vice versa. In particular the only admitted singularity, i. e. unisons is treated using the convention stated in section 3.2 on the dominion given by the whole range of chromatic pitches involved in the voice leading with multiplicity equal to the number of considered voices.

Partial singular braids on pitch classes

It is possible to model voice-leadings modulo octave considering pitch classes in $\mathbb{R}/12\mathbb{Z}$ instead of pitches. In this case the partial singular braids dominion is given by the chromatic set of pitch classes

$$\{\, [C]\,, [C\sharp]\,, \ldots, [B]\,\} \cong \{\, [0]\,, [1]\,, \ldots, [11]\,\}$$

and the strands are defined on a cylinder $C = \mathbb{R}/12\mathbb{Z} \times [0, 2\pi]$. Following the piecewise linear approach we used in the previous paragraphs, we assume the braid's strands



(a) β_1 (b) β_2 (c) β_3.

Figure 4.9: Partial braids inducing the same partial permutation.

to be lines wrapped around a cylinder i. e. geodesics on a cylinder, corresponding to helix segments parametrized as

$$\gamma : [0, 2\pi] \to \mathbb{R}^3$$
$$\gamma(t) = (\cos(at), \sin(at), t).$$

Consider the voice leading among pitch classes

$$\mathcal{C}_1 = (C, E, E) \to \mathcal{C}_2 = (F\sharp, C\sharp, E),$$

where, although the information concerning the octave is neglected working with pitch classes, we can deduce relative leaps among the voices. In figure 4.10 it is possible to observe how the curve connecting C and $F\sharp$ makes a complete round along the cylinder before connecting to $F\sharp$, meaning that the two notes are more than one octave far. The singularity at E in the top face of the cylinder represents the unison of the second and third voice of the chord $\mathcal{C}_1$, and the two voices lead to $C\sharp$ and E respectively without octave leaps.

Simultaneous voice motions are still well represented in this model: taking into account the orientation of the wrapping (clockwise or counterclockwise/right-handed or left-handed) of the helix segments on the cylinder. In our example we can deduce that C and one of the E move downward to $F\sharp$ and $C\sharp$ respectively, while the last E is fixed since the trajectory is a straight line. Topologically, this information is encoded in the fundamental group (groupoid) of the cylinder C: $\pi_1(C) = \mathbb{Z}$ meaning that m positive or negative turns around the cylinder encode the octave leaps information.

Remark 5. Considering the class of partial singular braids wit geodesics strands and positive crossings on pitch-classes, we cannot associate a unique braid to a partial permutation, in fact in this context geodesic are not unique, however it is always possible to consider minimal geodesics to represent strands among pitch classes, respecting the direction of the voice leading if it is known.

True and false crossings

The pitch class representation does not allow to distinguish among *true* and *false* voice crossings, unless the distance among the voices of the first chord is known *a*

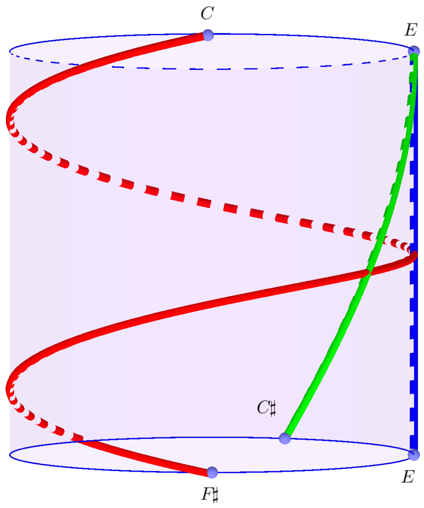


Figure 4.10: The partial singular braid representation of a voice leading defined in $\mathbb{R}/12\mathbb{Z}$.

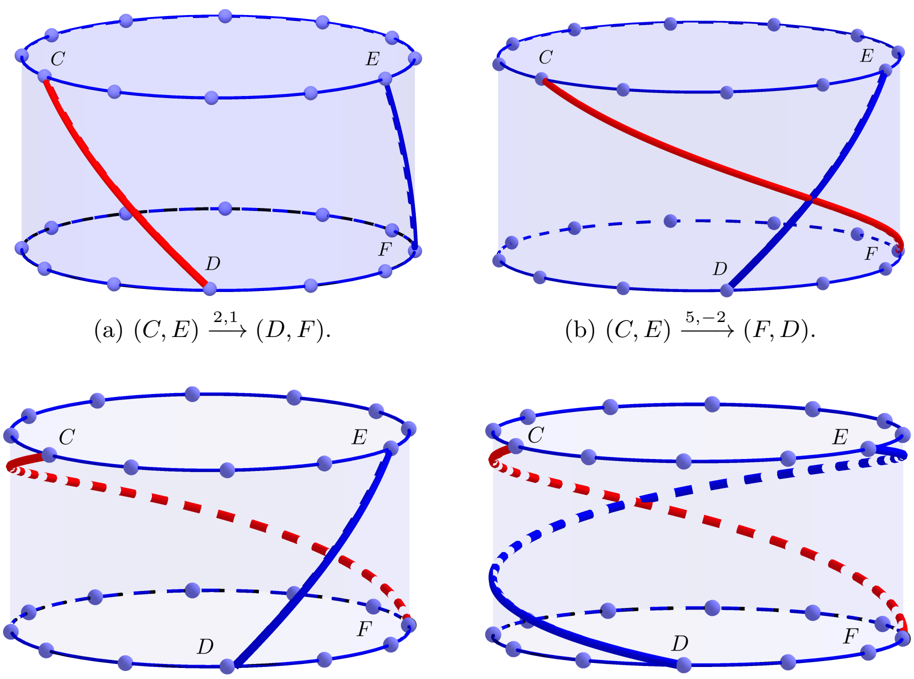
priori: a trajectory representing a leap of more than an octave and less than two, makes one complete turn around the cylinder, and hence crosses all the other strands involved in the voice leading, even if voices do not truly cross in musical sense.

In figure 4.11 four different voice leadings among the 2-pitch class chords $\mathcal{C}_1 = (C, E)$ and $\mathcal{C}_2 = (D, F)$ are depicted, see (Tymoczko, 2011, p. 76) for a representation of the same voice leadings in the chord space $\mathbb{A}_2$. As we stated few lines above, the lack on information given by the identification of the octaves does not allow to distinguish among true or false voice crossings, as it could be shown analyzing the four voice leading represented in figure 4.11:

a) In figure 4.11a the two strands does not make a complete tour of the cylinder and does not cross, meaning that the target pitch-classes lie in the same octave of the pitch-classes of the first chord and that there is no *topological* and musical crossing among the voices.

b) Figure 4.11b shows the *crossed* alternative of the previous voice leading $\mathcal{C}_1 \to \sigma_{12}\mathcal{C}_2$. Since the helix segments are left-handed and right-handed respectively and the strands of the braid do not complete the tour of the cylinder, what we can deduce from this configuration is that F lies in the same octave as C and symmetrically D lies in the same as E. To establish if the voices cross in a musical sense mirroring the strands' crossing, it is necessary to know the distance among the voices of the first chord: assuming C and E to belong to the same octave the voices actually cross in musical terms; however, if the two notes belong to two different octaves, for instance C_4 and E_5, no crossing occurs among them.

c) In figure 4.11c, C and E moves downward to reach F and D respectively and both voices moves of less than one octave. Considering the natural ordering of $\mathbb{Z}/12\mathbb{Z}$ being the class $[C] = [0] < [E] = [4]$, and the voices moving downward and upward respectively no crossing can occur among those voices as it is mirrored by the trajectories of the braid.



(a) $(C, E) \xrightarrow{2,1} (D, F)$.

(b) $(C, E) \xrightarrow{5,-2} (F, D)$.

(c) $(C, E) \xrightarrow{-7,-2} (F, D)$.

(d) $(C, E) \xrightarrow{-7,10} (F, D)$.

Figure 4.11: Simultaneous motions of two voices. The movement of each voice is written above the arrow in half-steps, the sign distinguish among upward and downward movements.

d) In the last figure, voices move in contrary motion, downward and upward respectively always targeting pitch-classes less than one octave distant from them. If the pitch classes of the first chord lie in the same octave, then no music crossing occurs, despite the topological configuration of the braid's strands, however it suffices to choose the representative of C and E to be C_4 and E_3 to have an actual crossing corresponding to the one depicted on the figure.

In conclusion, the identification modulo octave gives a representation of the no-crossing voice leading as the collection of shortest paths among multisets of pitch-classes[1] and maximize the number of crossings in the other cases, hence the pitch class braid model for voice leadings encodes all the information concerning both octave leaps and voice crossings as they are described in Hughes (2015) where a model of voice leading through the fundamental groupoid of the chord space $\mathbb{A}_n$, is described. See section 2.1 for a description of this subject.

Concatenation of voice leadings in $\mathcal{PSB}_n$

As we point out in section 3.3, representing several ordered voice leadings, it is not desirable to compose the braids representing each of them, but to concatenate

[1]Hence, neglecting the order in which voices are associated, we can always retrieve a no-crossing voice leading connecting the voices of the first chord to the ones of the second through minimal geodesics on the cylinder, as in figure 4.11a.

them one after the other: the composition $\mathcal{PS}B_n$ inherits from $\mathcal{P}B_n$ impose to delete strand fragments not connecting the first braid to the second, see figure 4.4. Thus, using the multiplication defined on $\mathcal{PS}B_n$ to compose braids, one would delete the strands representing any melody containing a rest, losing the information concerning the whole piece of music.

The idea is to represent a succession of voice leadings as a time series $\{\ \beta_i\ \}_{i\in\{\ 1,\dots,n\ \}}$, such that $\beta_i \in \mathcal{PS}B_n$ for each i, corresponding to a concatenation of braids as it is shown in figure 4.12 where both the pitches and pitch classes braids for the first seven voice leadings (corresponding to eight melodic states) of *Alleulia: Angelus Domini* are depicted. The fragment we analyse is given by the superposition of the two voices

$$v_1 = (F_4, G_4, A_4, G_4, F_4, G_4, B\flat_4, A_4)$$
$$v_2 = (C_4, D_4, E_4, F_4, G_4, G_4, F_4, E_4)\,,$$

represented by the blue and red trajectory respectively.

In this case, being the pitches involved in the part of the piece we represented contained in a octave, the pitch and the pitch class diagram are equivalent. It is possible to observe how this kind of representation gives a friendly access to the information describing the simultaneous motion of voices retrieving the special case of parallel motion and representing voice crossings and unisons as strands crossings and singularities. In this particular case, the concatenation of partial singular braids corresponds to the multiplication defined in $\mathcal{PS}B_n$.

This last representation allows to compare at first sight, voice leading *motives*, meaning patterns $\{\beta_p, \dots, \beta_q\} \subset \{\beta_1, \beta_n\}$ where $p < q \leqslant n$ evaluating their features in terms either of topological and motion complexity. The advantage of this braid representation, is the possibility to encode the whole information concerning the voice leading in a three dimensional braids, despite the number of voices composing it.

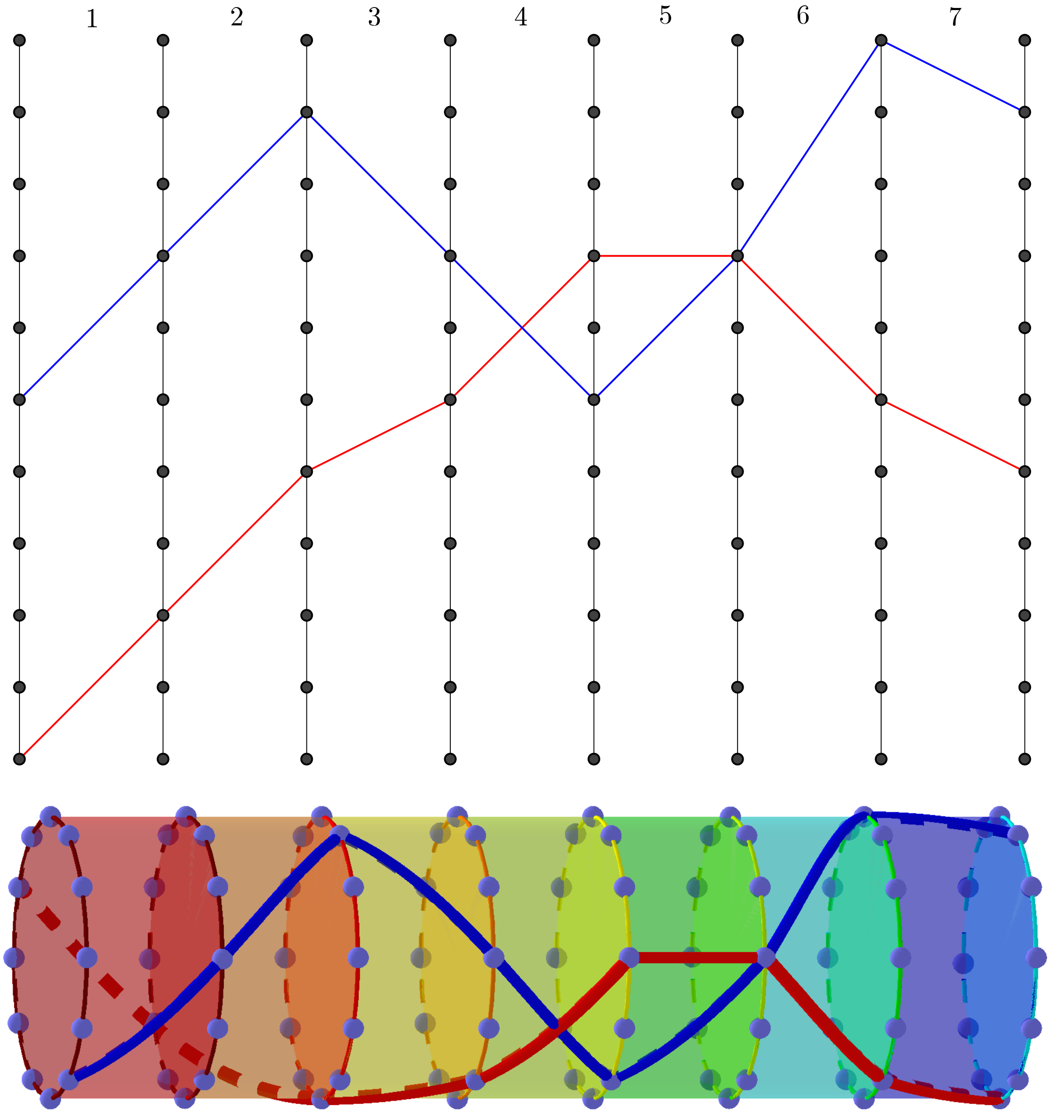


Figure 4.12: Concatenation of pitch and pitch-class partial singular braids.

Five

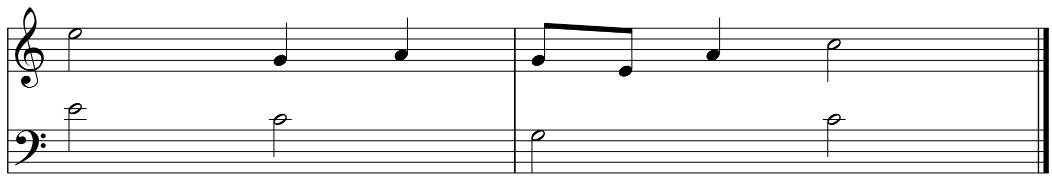
A naïve extension to other counterpoint species

5.1 Rhythmic independence and rests

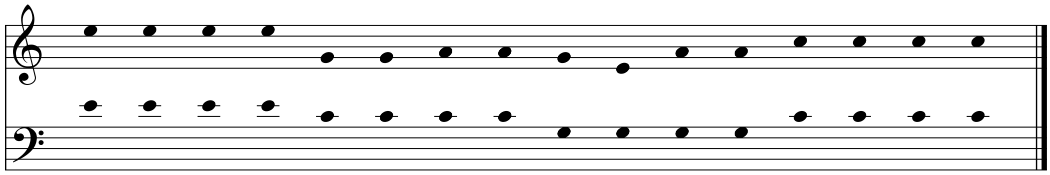
The examples analysed in Subsection 3.5 are counterpoints of the first species — which is the simplest case, in that the voices follow a *note-against-note* flow. It is however possible to study more complex scenarios by introducing *rhythmic independence* between voices and *rests* in the melody, in any case reducing non-simultaneous voices to the simplest case.

If the voices play at different rhythms or follow rhythmically irregular themes, we consider the minimal rhythmic unit u appearing in the phrase and *homogenise* the composition based on that unit: if a note has duration ku, with $k \in \mathbb{N}$, we represent it as k repeated notes of duration u (see Figure 7.5 for an example). This transformation of the original counterpoint introduces only oblique motions and does not alter the number of the other three kinds of motion.

In musical terms, if a voice is silent it is neither moving nor being constant and it cannot cross other voices. Therefore, in order to include rests in our model it is necessary to slightly modify Algorithm 3.1 by introducing a new symbol (p) in the



(a) Counterpoint of the fifth species.



(b) Reduction to the first species.

Figure 5.1: Reduction of rhythmically independent voices to a counterpoint of the first species.

dictionary of pitches; we also choose to indicate a rest in the matrices associated with a voice leading by the entry -1. We adopt the following convention concerning the ordered union multiset.

Convention 2. We choose rests to be the last elements in the ordered union multiset associated with a voice leading. In other words, we declare p to be strictly greater than any other pitch symbol.

Example 5.1.1. The voice leading $(p, D_4, D_5) \to (D_4, C_3, C_3)$ corresponds to the matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{pmatrix}.$$

Remark 6. Note that when introducing the -1's in the matrix associated with a voice leading we are no longer dealing with partial permutation matrices. However, to study voice leadings with rhythmic independence of the voices as before (thus ignoring rests) it is enough to consider the minor of the matrix obtained by deleting all rows and columns containing -1 (which is obviously again a partial permutation matrix).

We extend the complexity vector defined previously in Formula (3.4.1) by adding a fifth component that counts the number of voices that are silent at least once in the voice leading, i. e. it counts the number of negative (-1) entries of the associated matrix. Furthermore, we slightly modify also the notion of *normalised multiplicity* of a complexity vector c, needed for the representation of the complexity of a piece in the form of a point cloud, now dividing the number $\mu(c)$ of occurrences of c in the piece by the total number of notes per voice *after the homogenisation*.

Example: the *Retrograde Canon* by J. S. Bach

We consider the *Retrograde Canon* (also known as *Crab Canon*), a palindromic canon with two voices belonging to the *Musikalisches Opfer* by J. S. Bach, the beginning of which is reproduced in Figure 5.2.

We homogenise the rhythm by expressing each note in eighths and we apply Algorithm 3.1. Here is the output of the first four meaningful voice leadings:

```
Voice Leading: ['D4', 'D4'] ['D4', 'F4']
c = [1, 0, 1, 0, 0] - oblique motion
Voice Leading: ['D4', 'F4'] ['F4', 'A4']
c = [2, 0, 0, 0, 0] - similar motion up
Voice Leading: ['F4', 'A4'] ['F4', 'D5']
c = [1, 0, 1, 0, 0] - oblique motion
Voice Leading: ['F4', 'D5'] ['A4', 'C#5']
c = [1, 1, 0, 0, 0] - contrary motion
```

Table 5.1 collects the complexity vectors and their multiplicities; they are displayed in the form of point clouds in Figure 5.3.

Figure 5.2: The *Retrograde Canon* (bars 1–4), a palindromic canon belonging to the *Musikalisches Opfer* by J. S. Bach, and its reduction to first species counterpoint (unisons have been omitted).

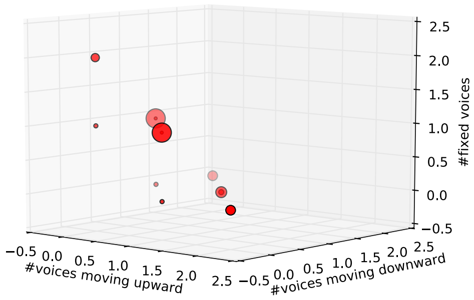
Table 5.1: Complexity vectors of the *Retrograde Canon* and their occurrences.

Retrograde Canon			
c	$\mu(c)$	c	$\mu(c)$
$(0,0,1,0,1)$	2	$(1,0,0,0,1)$	2
$(0,0,2,0,0)$	8	$(1,0,1,0,0)$	43
$(0,1,0,0,1)$	2	$(1,0,1,1,0)$	1
$(0,1,1,0,0)$	43	$(1,1,0,0,0)$	14
$(0,1,1,1,0)$	1	$(1,1,0,1,0)$	3
$(0,2,0,0,0)$	11	$(2,0,0,0,0)$	11

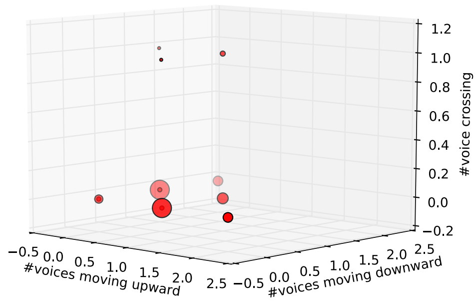
5.2 Concatenation of voice leadings and time series

The paradigmatic point cloud associated with a voice leading gives a useful 3-dimensional representation of the piece; however, this analysis is just structural, as it does not take into account the way in which voice leadings have been concatenated by the composer. It is possible to introduce this temporal dimension by looking at the sequence of complexity vectors from a different viewpoint.

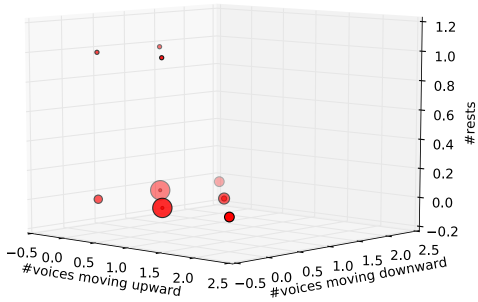
The concatenation of observations in time can be seen as a *time series*, that is a sequence of data concerning observations ordered according to time. In our case each piece of music can be described as a 5-dimensional time series, whose observations are the complexity vectors associated with each voice leading. More specifically, we use the so-called *dynamic time warping (DTW)*, a method for comparing time-dependent sequences of different lengths: it returns a measure of similarity between two given sequences by "warping" them non-linearly (see Figure **??** for an intuitive representation). We invite the reader to consult Senin (2008) for a detailed review of DTW algorithms.



(a) Projection on the first three components of the complexity vector.



(b) Projection on the *upward, downward* and *crossing* components of c.



(c) Projection on the *upward, downward* and *rest* components of c.

Figure 5.3: Three-dimensional projection of the 5-dimensional point cloud representing the complexity of the *Retrograde Canon*. The radius of each circle represents the normalised multiplicity of each complexity vector.

Dynamic time warping analysis

Let $\mathcal{F}$ be a set, called the *feature space*, and take two finite sequences $X \coloneqq (x_1, \ldots, x_n)$ and $Y \coloneqq (y_1, \ldots, y_m)$ of elements of $\mathcal{F}$, called *features* (here n and m are natural numbers). In order to compare them, we need to introduce a notion of distance between features, that is a map $\mathcal{C} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$, also called a *cost function*, that meets at least the following requirements:

i. $\mathcal{C}(x, y) \geqslant 0$ for all $x, y \in \mathcal{F}$;

ii. $\mathcal{C}(x, y) = 0$ if and only if $x = y$;

iii. $\mathcal{C}(x, y) = \mathcal{C}(y, x)$ for all $x, y \in \mathcal{F}$.

Now, if we apply $\mathcal{C}$ to the features X and Y, we can arrange the values in an $n \times m$ real matrix $C \coloneqq \big(\mathcal{C}(x_i, y_j)\big)$, where i ranges in $\{1, \ldots, n\}$ and j in $\{1, \ldots, m\}$.

A (n, m)-*warping path* in C is a finite sequence $\gamma \coloneqq (\gamma_1, \ldots, \gamma_l) \in \mathbb{R}^l$, with $l \in \mathbb{N}$, such that:

1. $\gamma_k \coloneqq (\gamma_k^x, \gamma_k^y) \in \{1, \ldots, n\} \times \{1, \ldots, m\}$ for all $k \in \{1, \ldots, l\}$;

2. $\gamma_1 \coloneqq (1, 1)$ and $\gamma_l \coloneqq (n, m)$;

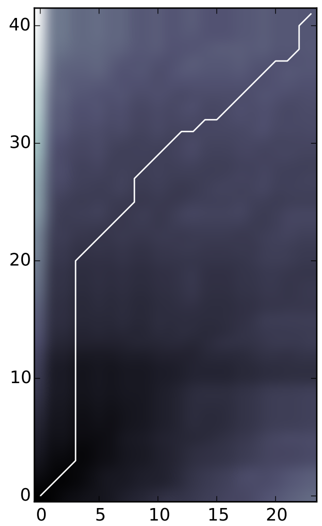


Figure 5.4: Optimal warping path on *Alleluia, Angelus Domini* and *Dicant nunc Judei*.

Table 5.2: DTW distance matrix for the three time series of complexity vectors.

	Alleluia	*Dicant*	*Canon*
Alleluia	0.00	0.62	1.34
Dicant	0.62	0.00	1.16
Canon	1.34	1.16	0.00

3. $\gamma_k^x \leqslant \gamma_{k+1}^x$ and $\gamma_k^y \leqslant \gamma_{k+1}^y$ for all $k \in \{1, \ldots, l-1\}$;
4. $\gamma_{k+1} - \gamma_k \in \{(1,0),(0,1),(1,1)\}$ for all $k \in \{1, \ldots, l-1\}$.

The *total cost* of a (n, m)-warping path γ over the features X and Y is defined as

$$\mathcal{C}_\gamma(X, Y) \coloneqq \sum_{k=1}^{l} \mathcal{C}(x_{\gamma_k^x}, y_{\gamma_k^y}).$$

An *optimal warping path* on X and Y is a warping path realising the minimum total cost (see Figure 5.4). We are now ready to define the *DTW distance* between X and Y:

$$DTW(X, Y) \coloneqq \min\left\{ \mathcal{C}_\gamma(X, Y) \mid \gamma \text{ is a } (n, m)\text{-warping path} \right\}.$$

Remark 7. Note that the minimum always exists because the set is finite.

We computed the DTW distance between each pair of the three examples that we analysed in Subections 3.5 and 5.1, choosing as cost function the Euclidean distance in $\mathbb{R}^5$. We embedded the 4-dimensional complexity vectors in $\mathbb{R}^5$ by adding a fifth component and setting it to 0. The results of the comparison are shown in Table 5.2. Although we analysed only three compositions, it is possible to observe how the DTW distance segregates the two pieces belonging to the Chartres fragments.

Conclusions and future works

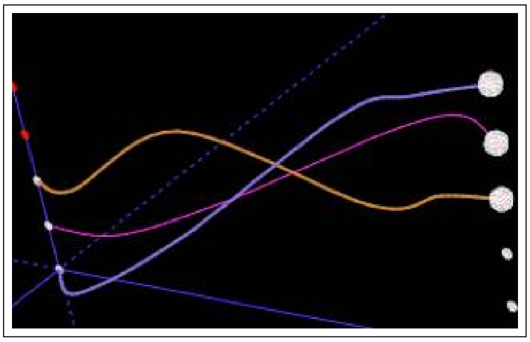


Figure 5.5: Particles on a partial braid.

Thanks to a mathematical formalization of the concept of voice leading, we deduced a model to describe voices' motions as low dimensional partial permutation matrices. Giving a geodesics-oriented interpretation of voice leadings and their concatenations highlighted the link among the analysis of these objects in both the algebraic and geometric context.

The information carried by the voice-leading partial permutation matrices has been decoded under the form of complexity vector used to characterize first specie counterpoints as multisets of 4-dimensional points.

In order to visualize voice leadings in a 3-dimensional space, a braid-like representation has been introduced allowing to extend the first model, measuring the intervallic leap of each voice in the passage from a chord to the next one. In addition, a pitch-class version of the braid representation provides an environment to visualize voice leadings among n-chords, mirroring the properties of trajectories in the space of chords.

In conclusion, the model has been extended to other counterpoint species. The suite of complexity vectors is a multi-dimensional time series describing how different kind of motions has been concatenated in time. Dynamic time warping provides a measure of the distance between the complexities of two pieces, and giving a quantitative description of the dissimilarity of the time series describing the pieces.

A straight-forward development offered by the model we describe, is the possibility to classify the collection of possible voice leadings among two chords in terms of the length of the geodesic strands of the braid representing it. Connecting the notes of the two chords with minimal paths corresponds to avoiding voice crossings, variations of this configuration could be classified considering the length of each strand of the braid associated to the voice leading. The method we described in order to bend a fifth specie counterpoint to our model is a simplistic one. It could be possible to describe the independency of voices losing the isomorphism among partial permutations and partial braids, considering a wider class of strands, rather than the geodesics ones. Discretized braids allow to describe leaps and voice crossings, which in this first model seemed to be an essential feature. Thinking to notes as particle moving on the strand at a constant speed, would allow to model longer notes in rhythmical terms, simply deforming their strand.

In addition, the model we introduced as a visualization tool, has topological properties that could be investigated, for instance in terms of knot theory. To do that one should weaken our assumption on the crossings among the strands. A possible definition could involve, for discretized braids, the slope of the line segment describing the voices, forcing a strand describing the bigger leap to pass above the others.