# The Circle of Fifths as Latent Geometry in Bach's Well-Tempered Clavier

**Najla Sadek**          NSS32@MAIL.AUB.EDU  and  **Joseph Bakarji**          JB50@AUB.EDU.LB
*American University of Beirut, Lebanon*

## Abstract

Can unsupervised deep learning methods encode fundamental music-theoretic features? We answer this question by training an autoencoder on J.S. Bach's *Well-Tempered Clavier* and analyzing its latent space via principal component analysis. Sequences in the first two principal components are clustered hierarchically into pieces and keys that spontaneously arrange in a circle-of-fifths geometry. Quantitatively, relative major-minor key pairs (sharing pitch collections) lie more than three times closer than non-relative pairs, and circle-of-fifths distance correlates strongly with learned distances. This structure emerges entirely from reconstruction loss, with no harmonic labels or supervision. Our results suggest that the circle of fifths is an intrinsic property of tonal relationships, demonstrating that un-supervised representation learning can recover harmonic principles that open the door for interpretable data-driven exploration of latent spaces across diverse musical traditions.

**Keywords:** Deep Learning, Music Information Retrieval, Autoencoders, Latent Space Analysis, Circle of Fifths, Bach

## 1. Introduction

Deep learning has revolutionized music modeling, analysis and generation, enabling systems that compose in a wide range of styles (Hadjeres et al., 2017), harmonize melodies Lim et al. (2017), and generate novel polyphonic music Huang et al. (2018). But despite their capabilities, most models remain opaque, which makes it hard to separate what is being emulated from what is being generated. Furthermore, without some level of interpretability, the iterative process of music production, and the incremental understanding required for musicological analyses can cut the (human) musical creator out of the creative process and defeat the purpose of music creation altogether. In this paper, we shed light on a way to address this challenge.

Many deep learning approaches involve a form of encoding to a latent space where musical information is compactly represented to make downstream tasks possible. Those latent spaces are rarely interpretable, if they're ever analyzed or characterized explicitly, leaving open questions about whether they capture true music-theoretic structure that enables true generalization. If we understand the type of structure they create, we can understand what they generalize over, and can potentially push them to break those regularities to explore musical spaces inaccessible to them or us.

Recent work, such as MeasureVAE by Pati and Lerch (2021) and MiniBach by Briot (2021) demonstrate that compact embeddings can retain rhythmic and harmonic features to enable transformation and generation. Recent work by Wang et al. (2020) showed that musicologically consistent representations can be extracted from transcription data using variational autoencoders. But as noted by Bryan-Kinns et al. (2023), explainable AI for

the arts still faces the challenge of linking abstract representations to human-interpretable concepts.

In this study, we show that even a simple feedforward autoencoder, trained solely on a reconstruction loss, can spontaneously reveal familiar harmonic structure observed in a 12-TET tuning system. By training on tempo-normalized piano-roll encodings of J.S. Bach's *Well-Tempered Clavier*, Book I, and projecting its 128-dimensional latent space using Principal Component Analysis (PCA), we reveal that the first two principal components (PC1, PC2) encode interpretable harmonic relationships. Specifically, the emergent geometry organizes pieces hierarchically from sequences to keys, and clusters keys around a nearly exact circle of fifths in the two-dimensional latent space.

These results suggest that fundamental music-theoretic organization emerges naturally from a combination of linear and nonlinear data compression, opening the door for interpretable melodic, modal and harmonic analysis on time series data in encoded latent spaces.

## 2. Related Work

Deep learning approaches for music have primarily focused on generation quality and stylistic control. MeasureVAE uses variational autoencoders to compress MIDI into structured latent spaces, preserving rhythmic and harmonic features. To improve interpretability, the authors introduced latent space regularization (LSR) that explicitly ties latent dimensions to musical attributes like rhythmic complexity and note density. Bryan-Kinns et al. (2023) extended this by integrating XAI techniques, demonstrating the value of real-time user interaction in music generation systems.

MiniBach, a supervised autoencoder for Bach chorales, predicts four-voice counterpoint from soprano melodies using piano-roll representations. While it's effective for music generation, the authors do not attempt to interpret the latent structure. Similarly, MusicVAE by Koh et al. (2018) combines convolutional neural networks with variational recurrent neural networks for sequence generation, using Information Rate as an evaluation metric.

In contrast, our work demonstrates that fundamental music-theoretic structure (specifically the circle of fifths) emerges spontaneously from a purely unsupervised reconstruction training. Rather than imposing interpretability through regularization, we analyze the geometric organization that arises naturally in the latent space of a simple feedforward autoencoder couple with a linear dimensionality reduction through PCA.

## 3. Method

We train a full-connected feedforward autoencoder on tempo-normalized MIDI sequences from J.S. Bach's *Well-Tempered Clavier*, Book I. Our pipeline consists of three stages: (1) encoding MIDI into a piano-roll representation, (2) learning compressed latent representations via reconstruction loss, and (3) analyzing the emergent geometric structure through principal component analysis. Figure 1 illustrates the complete architecture.
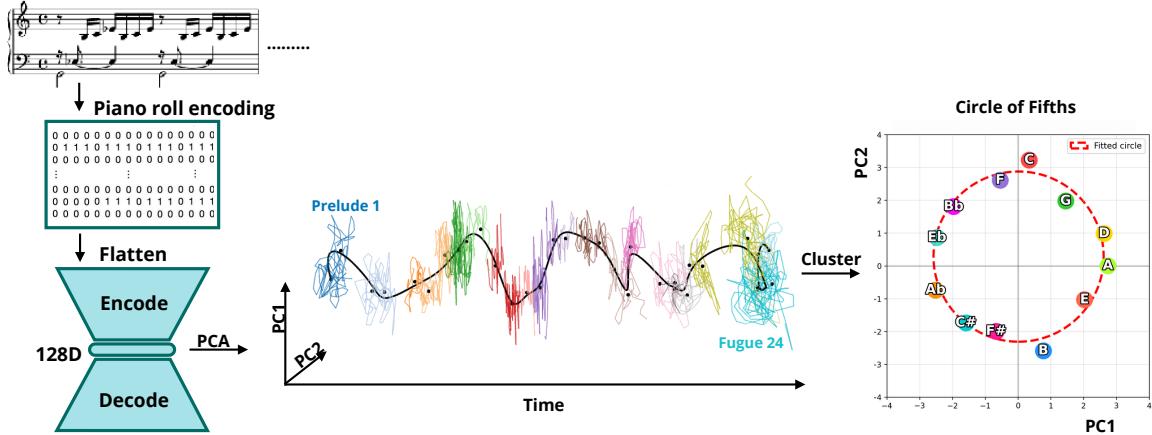
Figure 1: Latent space trajectory through Bach's *Well-Tempered Clavier*, Book I, visualized in PC1-PC2-time space. Each of the (preludes and fugues) is pairs shown as a colored trajectory, with the black curve representing a smooth fitted spiral through piece centroids. The spiral structure reveals progressive harmonic evolution across the corpus, with time flowing left to right along the horizontal axis. PC1 and PC2 encode the two dominant harmonic dimensions captured through unsupervised learning, collectively explaining 8.68% of the latent space variance.

### 3.1. Piano Roll Encoding

We represent MIDI data as a binary piano roll matrix, similarly to what was done by Briot (2021). In this encoding, each entry indicates whether a note is active at a given moment via a binary value. All pieces are normalized to 60 BPM and sampled at 4 frames per second (one frame per sixteenth note). Each piano roll $\mathbf{P} \in \{0,1\}^{T \times 88}$ spans $T$ timesteps and 88 pitches (MIDI notes 21–108, A0 to C8).

We extract overlapping sequences of length $L = 16$ frames with stride 1, producing $S = T - L + 1$ sequences per piece. Each sequence is flattened into a vector $\mathbf{x}_i \in \{0,1\}^{1408}$, where $1408 = 16 \times 88$. Sequences from all pieces are concatenated, yielding a dataset of 32171 sequences total. Piece boundaries are tracked via an index vector to enable per-piece analysis.

One of the shortcomings of the piano roll encoding is that it doesn't explicitly specify the difference between onsets and notes that are have been played before. But we don't expect this limitation to affect our current analysis significantly.

### 3.2. Architecture

The autoencoder consists of symmetric encoder and decoder networks with dense layers. The encoder $\mathbf{f_\theta} : \{0,1\}^{1408} \to \mathbb{R}^{128}$ progressively compresses each input sequence through the encoder layers of sizes: 1408 (input), 1024, 512, 128 (latent space), using Swish activations, batch normalization, and 10% dropout for regularization. The decoder $\mathbf{g}_\phi$ mirrors the structure of the encoder, going from 128-dimensional latent vector $\mathbf{z}_i$ back to the original dimensionality with a sigmoid output layer that predicts a value between 0 and 1 for each

3

piano roll value. That values are then rounded for projecting back on midi. The size of the latent variable was chosen through hyperparameter tuning to be the minimum viable dimension that generalizes on a 20% hold-out set.

The model is trained to minimize binary cross-entropy reconstruction loss, where $\hat{\mathbf{x}}_i = \mathbf{g}_\phi(\mathbf{f}_\theta(\mathbf{x}_i))$ is the reconstruction. We use the AdamW optimizer with learning rate $5 \times 10^{-4}$, weight decay, early stopping, and adaptive learning rate scheduling.

### 3.3. Principal Component Analysis and Latent Geometry

After training, we get the 128-dimensional latent vectors $\mathbf{z}_i$ for non overlapping sequences and apply PCA to reduce dimensionality for visualization and analysis. The first two principal components (PC1 and PC2) collectively explain only 8.68% of the latent space variance (4.63% and 4.05% respectively), yet they capture the dominant harmonic structure.

This low explained variance is expected. The autoencoder's 128 dimensions most probably encode many musical attributes (rhythm, texture, register, and local melodic patterns) beyond harmony alone, and PCA identifies the directions of highest variance, which in this corpus correspond to harmonic progressions and key relationships. The remaining 91% of variance might contain other musical dimensions orthogonal to the tonal organization analyze.

When visualized in PC1-PC2-time space (Fig. 1), the latent trajectories form a clear cyclical structure. This emergent geometry, arising purely from reconstruction training without harmonic supervision, is the focus of the analysis that follows.

## 4. Experiments

### 4.1. Hierarchical Emergence of Circle-of-Fifths Geometry

The circle of fifths, a geometric arrangement where adjacent keys differ by a perfect fifth interval is a fundamental organizing principle and harmonic framework in equal-temperament 12-TET tonal music. Keys close on this circle share common notes and have harmonic affinity, while distant keys sound unrelated, and their associated chords are typically harder (within classical tonal music) to modulate to within the same key.

We test whether this structure emerges spontaneously in our learned latent space through hierarchical aggregation. First, we compute piece-level centroids by averaging all sequence representations within each piece. Second, we aggregate piece centroids into key-level centroids by averaging major and minor pieces sharing the same tonic (e.g., C Major and C Minor). This two-step process mirrors the hierarchical organization predicted by the theory.

Figure 2 visualizes this convergence across three stages. Panel A shows raw sequences clustered by piece. Panel B reveals piece centroids with relative key pairs connected, demonstrating the first level of aggregation. Panel C shows key-level centroids with the a fitted circle, confirming a near-perfect circle-of-fifths geometry.

### 4.2. Circle-of-Fifths Distance Correlations

To quantify circle-of-fifths organization, we analyzed pairwise Euclidean distances between the 12 key centroids in PC1-PC2 space. Music theory predicts that fifth-related keys (neigh-
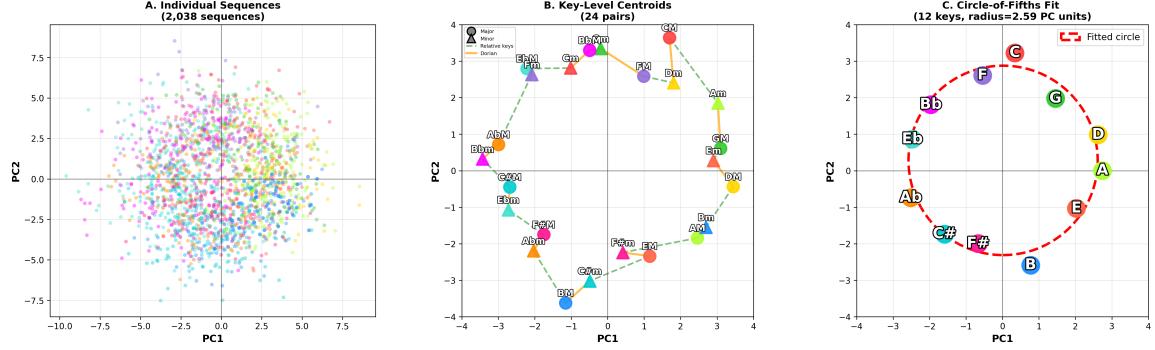
Figure 2: Hierarchical emergence of circle-of-fifths geometry through progressive aggregation through averaging. **(A)** Individual sequences (2,038 points) colored by key, showing the raw latent space distribution. **(B)** Piece-Pair level centroids (24 prelude-fugue pairs centroids) with connections showing two types of key relationships: relative keys (green dashed lines, 12 pairs) share similar diatonic collections (e.g., C Major and A Minor), while close-tonic pairs (yellow solid lines, 12 pairs) have tonics a whole step apart (e.g., C Major and D Minor). Close-tonic pairs are notably closer (mean distance $0.79 \pm 0.38$) than relative pairs ($1.35 \pm 0.62$), despite having *more* accidentals, suggesting the latent space encodes harmonic function rather than just pitch overlap. **(C)** Key-level centroids (12 keys) obtained by averaging Major and Minor pieces with the same tonic, with fitted circle overlay (radius = 2.59 PC units, CV $\approx$ 8.5%) demonstrating near-perfect circle-of-fifths geometry. Pearson correlation between circle-of-fifths distance and Euclidean distance: $r = 0.95$ ($p = 1.24 \times 10^{-33}$).

bors on the circle, e.g., C $\rightarrow$ G $\rightarrow$ D) should be close, while tritone-related keys (opposite on the circle, e.g., C and F$\sharp$) should be distant.

Table 1 confirms these predictions: tritone-related keys are 2.25$\times$ further apart than fifth-related keys. A Pearson correlation between circle-of-fifths distance (shortest path, 0-6 steps) and Euclidean distance in PC space yielded $r = 0.95$ ($p = 1.24 \times 10^{-33}$), demonstrating strong alignment between the learned geometry and music-theoretic structure. An algebraic least-squares circle fit reveals near-perfect circular arrangement (radius $\approx$ 2.6 PC units, coefficient of variation $\approx$ 8.5%; see Appendix for details).

| Relationship | Example | Mean Distance | SD |
|---|---|---|---|
| Fifth-related | C Major $\rightarrow$ G Major | 2.54 | 1.34 |
| Tritone-related | C Major $\rightarrow$ F$\sharp$ Major | 5.71 | 0.71 |
| **Ratio** | — | **2.25$\times$** | — |

Table 1: Distances between pieces in PC space by circle-of-fifths relationship. Mean and SD computed over all piece pairs for each relationship type (N=48 pairs each). Examples show individual instances from each category. Tritone-related pieces are 2.25$\times$ further apart than fifth-related pieces.

### 4.3. Relative Key Proximity and Harmonic Context

In music theory, relative keys are major and minor keys sharing the same diatonic pitch collection (e.g., C Major and A Minor both use C-D-E-F-G-A-B). We identified all 12 relative pairs in the *Well-Tempered Clavier* and computed pairwise distances for all possible 276 piece pairs.

Relative key pairs ($n = 12$) have mean distance $1.35 \pm 0.64$, while non-relative pairs ($n = 264$) have mean distance $4.29 \pm 1.58$; a 3.17 times difference. This effect is significant across multiple tests: $t = 7.44$, $p = 1.43 \times 10^{-8}$ (two-sample $t$-test); $U = 3{,}168$, $p = 3.27 \times 10^{-7}$ (Mann-Whitney); $KS = 0.833$, $p = 1.52 \times 10^{-7}$ (Kolmogorov-Smirnov).

All 12 relative pairs exhibited the predicted proximity pattern, with 75% having distances below 2.0 PC units; well below the mean non-relative distance. Table 2 lists all pairs sorted by distance.

| Rank | Relative Pair | Distance | Interpretation |
|:---:|:---|:---:|:---:|
| 1 | G Minor ↔ B♭ Major | 0.31 | Extremely close |
| 2 | E Minor ↔ G Major | 0.39 | Extremely close |
| 3 | D Minor ↔ F Major | 0.84 | Very close |
| 4 | C♯ Major ↔ B♭ Minor | 1.07 | Close |
| 5 | E♭ Minor ↔ F♯ Major | 1.19 | Close |
| 6 | C Minor ↔ E♭ Major | 1.20 | Close |
| 7 | D Major ↔ B Minor | 1.34 | Moderate |
| 8 | A♭ Minor ↔ B Major | 1.68 | Moderate |
| 9 | C♯ Minor ↔ E Major | 1.78 | Moderate |
| 10 | F♯ Minor ↔ A Major | 2.08 | Further |
| 11 | F Minor ↔ A♭ Major | 2.13 | Further |
| 12 | C Major ↔ A Minor | 2.23 | Further |

Table 2: All 12 relative key pairs sorted by Euclidean distance in PC1-PC2 space. Even the furthest pair (2.23) is significantly closer than the mean non-relative distance (4.29).

A natural question follows from our analysis: does the latent space encodes keys (tonal centers with harmonic function) or merely pitch collections (chords)? We tested this by comparing two types of key relationships visible in Figure 2B.

Relative key pairs (green dashed lines) share similar diatonic pitch collections: C Major and A Minor both primarily use C-D-E-F-G-A-B (with A Minor adding occasional chromatic inflections like G♯). These pairs have mean distance $1.35 \pm 0.62$ PC units. Close-tonic pairs (yellow solid lines) have tonics separated by a whole step: C Major and D Minor. Surprisingly, these pairs are *even closer*: mean distance $0.79 \pm 0.38$ PC units; despite D Minor using more accidentals (20.8%) than A Minor (12.2%) relative to C Major's diatonic collection.

Analysis of the actual piano roll data reveals that Bach uses harmonic and melodic minor forms with raised scale degrees (e.g., D Minor contains C♯ in 4.7% of notes, B♭ in 8.4%). Across all 24 pieces, we find a strong positive correlation ($r = 0.90$, $p < 10^{-8}$) between accidental usage and distance from C Major in PC space, confirming that pieces with more "black piano keys" occupy more distant latent regions; which can be understood as a form of statistical orthogonality between pieces in different keys. That is, the cosine distance between F♯ and C is much bigger than that between C and G.

Yet close-tonic pairs remain closer than relative pairs despite having *more* accidentals. This cannot be explained by pitch overlap only. Instead, it suggests the latent space encodes harmonic function (i.e. the typical chord progressions, voice-leading patterns, and tonal relationships characteristic of each key) rather than just counting shared pitches. The autoencoder somehow learns that keys with tonics a whole step apart (C-D, F-G, etc.) share more functional harmonic structure than relative major-minor pairs, even when the latter share more raw pitches.

## 4.4. Classification Results on Unseen Bach Pieces

To evaluate generalization, we classified nine unseen Bach pieces from inventions, WTC Book II, and other keyboard works. Following the same preprocessing pipeline, pieces were converted to piano roll representations at 16 frames per second, normalized to 60 BPM, and segmented into 16-timestep sequences. Each test piece was projected into the 128-dimensional latent space and PCA-reduced to 2D. We computed sequence centroids and measured Euclidean distances to the 24 WTC Book I training centroids (see Figure 3).

Table 3 shows classification results. The model achieved 55.6% top-1 accuracy (5/9 correct) and 100% top-2 accuracy. Strong performance was observed on pieces with clear tonal centers: Invention No. 4 in D Minor (0.28 PC units), BWV 892 in B Major (0.54 PC units), and Invention No. 13 in A Minor (0.63 PC units).

Four pieces required the second prediction: Inventions No. 7 and No. 8, BWV 883, and Partita No. 1. Misclassifications follow interpretable patterns where predicted keys are closely related through circle-of-fifths relationships or relative major-minor pairs.

Table 3: Classification results for 9 unseen test pieces

| Piece | Truth | 1st Pred. | Dist | 2nd Pred. | Dist |
|---|---|---|---|---|---|
| Invention No. 4 | D Minor | **D Minor** | 0.28 | F Major | 1.10 |
| Partita No. 1 | B♭ Major | C Minor | 0.43 | **B♭ Major** | 0.75 |
| BWV 892 (Book II) | B Major | **B Major** | 0.54 | C# Minor | 1.36 |
| Invention No. 13 | A Minor | **A Minor** | 0.63 | G Major | 1.34 |
| French Suite No. 5 | G Major | **G Major** | 0.67 | A Minor | 0.93 |
| Invention No. 8 | F Major | C Major | 0.71 | **F Major** | 1.86 |
| BWV 883 (Book II) | F# Minor | E Major | 0.95 | **F# Minor** | 1.39 |
| Invention No. 7 | E Minor | D Major | 1.05 | **E Minor** | 1.86 |
| Invention No. 1 | C Major | **C Major** | 0.11 | F Major | 1.27 |

## 5. Conclusion

This study demonstrates that deep autoencoders trained on symbolic music can uncover fundamental structures of tonal harmony. Without any supervision or theoretical priors, our model rediscovered central principles of 12-TET music: hierarchical tonal organization, circle-of-fifths geometry, and relative-key proximity.
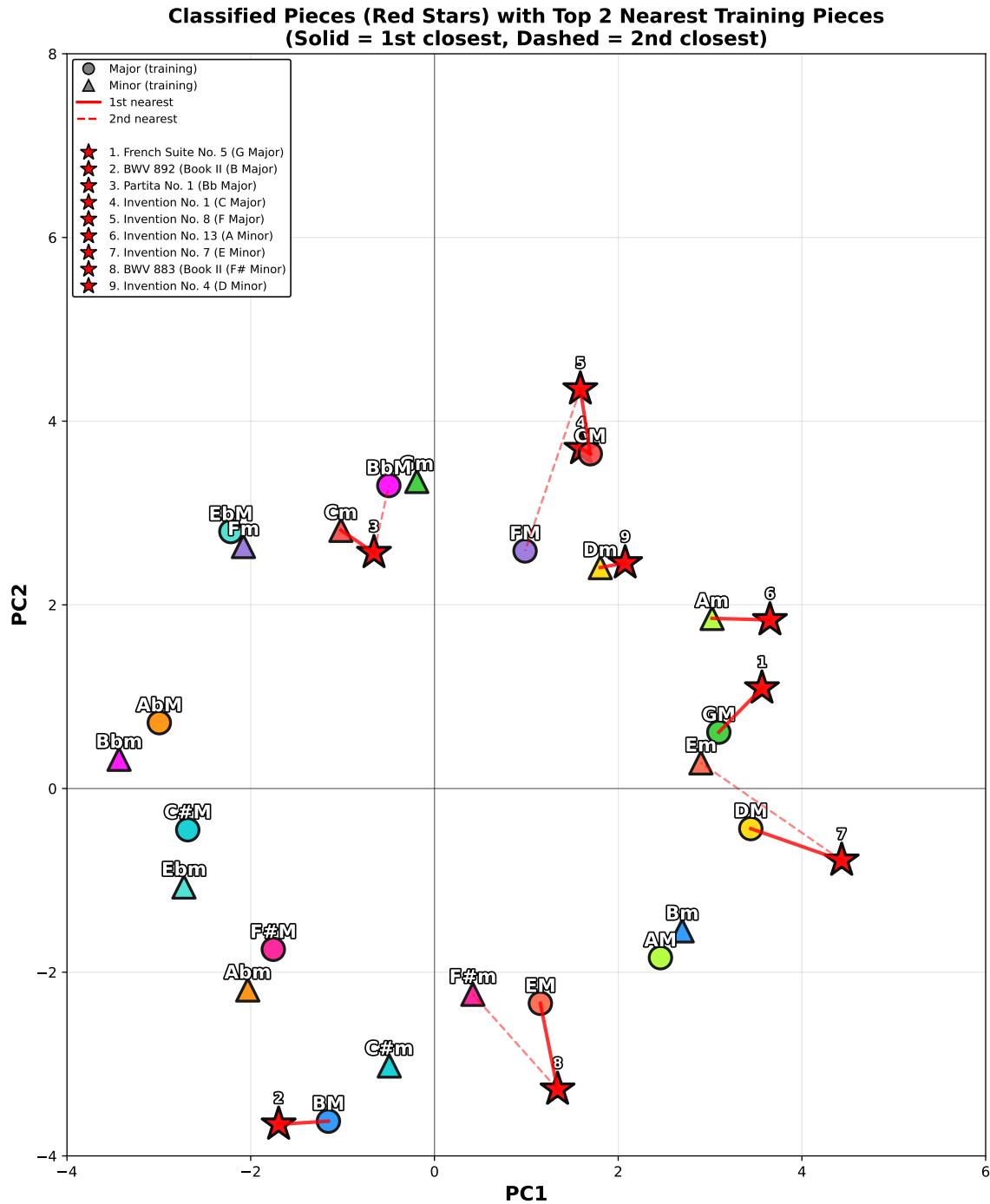
Figure 3: Classification of nine unseen Bach pieces in PC1-PC2 space. Red stars indicate test pieces (numbered 1-9, see legend), with solid lines connecting to the nearest training piece and dashed lines to the second-nearest. Training pieces from WTC Book I are shown as circles (major keys) and triangles (minor keys), colored by tonic.

Quantitatively, the first two principal components of the learned latent space aligned strongly with theoretical expectations, while relative major–minor pairs clustered over three times closer than non-relatives.

In addition, analysis of close-tonic pairs revealed that the latent space encodes a sort of **harmonic function** rather than merely pitch collections. Keys with tonics a whole step apart (e.g., C Major and D Minor) clustered closer than relative key pairs, despite having *more* accidentals (20.8% vs. 12.2%). This proximity, combined with the strong correlation between accidental usage and latent space distance ($r = 0.90$, $p < 10^{-8}$), demonstrates that the autoencoder learns functional harmonic relationships rather than just counting shared pitches.

Beyond quantitative validation, we develop an interactive visualization 4 that makes our results more accessible to users who want to explore the discovered latent space. We make the model's internal geometry interpretable; transforming an opaque network into a lens for musical understanding.

Finally, these findings invite a dialogue between machine learning and musicology. Extending this framework to other polyphonic compositions, style transfer, and other musical traditions may illuminate whether such geometric regularities generalize across cultures and genres. In this sense, the latent space becomes more than a technical means to an end (typically for music generation or classification). It becomes a map of how harmony itself organizes information.

## Acknowledgments

## References

Paul J Blazek, Kesavan Venkatesh, and Milo M Lin. Automated discovery of algorithms from data. *Nature Computational Science*, 4(2):110–118, 2024.

Jean-Pierre Briot. From artificial neural networks to deep learning for music generation: history, concepts and trends. *Neural Computing and Applications*, 33(1):39–65, 2021.

Nick Bryan-Kinns, Berker Banar, Corey Ford, Courtney N Reed, Yixiao Zhang, Simon Colton, and Jack Armitage. Exploring xai for the arts: Explaining latent space in generative music. *arXiv preprint arXiv:2308.05496*, 2023.

Gaëtan Hadjeres, François Pachet, and Frank Nielsen. Deepbach: a steerable model for bach chorales generation. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, pages 1362–1371. PMLR, 2017.

Cheng-Zhi Anna Huang, Ashish Vaswani, Jakob Uszkoreit, Noam Shazeer, Ian Simon, Curtis Hawthorne, Andrew M Dai, Matthew D Hoffman, Monica Dinculescu, and Douglas Eck. Music transformer. *arXiv preprint arXiv:1809.04281*, 2018.

Diederik P Kingma, Max Welling, et al. Auto-encoding variational bayes, 2013.

Eunjeong Stella Koh, Shlomo Dubnov, and Dustin Wright. Rethinking recurrent latent variable model for music composition. In *2018 IEEE 20th International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–6. IEEE, 2018.

Hyungui Lim, Seungyeon Rhyu, and Kyogu Lee. Chord generation from symbolic melody using BLSTM networks. In *Proceedings of the 18th International Society for Music Information Retrieval Conference (ISMIR)*, pages 621–627, 2017.

Dima Mrad, Sara Najem, Pablo Padilla, and Francis Knights. A network perspective on js bach's 6 violin sonatas and partitas, bwv 1001-1006. *Physica A: Statistical Mechanics and its Applications*, 654:130124, 2024.

Ashis Pati and Alexander Lerch. Attribute-based regularization of latent spaces for variational auto-encoders. *Neural Computing and Applications*, 33:4429–4444, 2021.

Ziyu Wang, Yiyi Zhang, Yixiao Zhang, Junyan Jiang, Ruihan Yang, Junbo Zhao, and Gus Xia. Pianotree vae: Structured representation learning for polyphonic music. *arXiv preprint arXiv:2008.07118*, 2020.

# Appendix A. Interactive User Interface for Latent Space Exploration

## A.1. Interactive interface

One of the goals of this study is to help both scientists and musicians interpret what deep unsupervised learning methods do. To facilitate this type of exploration, we developed an interactive user interface (UI) that provides real-time visual, auditory, and symbolic access to individual latent sequences. The system enables users to associate latent points with their corresponding musical content by selecting and listening to specific sequences. Currently this interface has to be run through github.

The UI displays musical sequences as points in a low-dimensional principal component analysis (PCA) space. Two modes are provided: i) A two-dimensional $(PC1, PC2)$ scatter plot, and ii) a three-dimensional $(PC1, PC2, \text{Time})$ scatter plot, where Time represents the sequence's relative position within its musical section.

Color is used to distinguish musical sections, helping users visually track structure across modes. The interface also includes interactive filtering, highlighting, and playback functionality to support intuitive exploration.
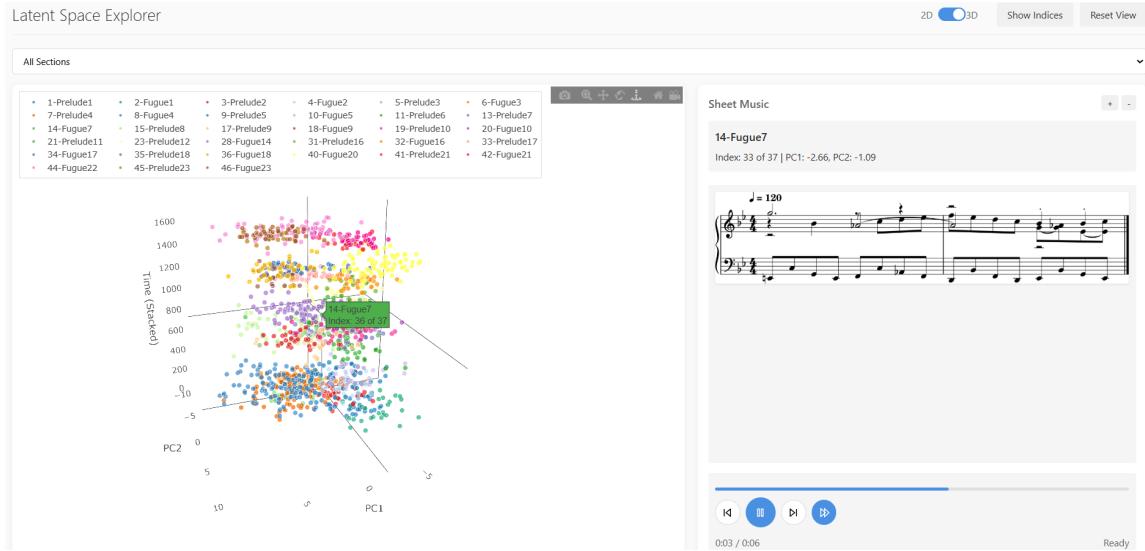


Figure 4: Full latent space view in 3D with all sections shown.

Clicking on any point in the visualization instantly triggers audio playback and, when available, displays the corresponding sheet music. This allows users to connect spatial position with musical meaning through both listening and visual score inspection. A dedicated control allows users to show or hide numeric indices over the points, indicating each sequence's position within its section. These labels help in orienting the user within the structural flow of the music. Furthermore, a toggle enables seamless switching between 2D and 3D visualizations.

The currently selected point is visually marked with a distinct color, and this highlight remains consistent across view switches and interactions. This ensures the user can maintain orientation while exploring different dimensions. Users can also filter the latent space to view

only the points from a selected musical section. This feature enables focused examination of specific parts of the dataset and aids in comparative analysis across sections.

## Appendix B.  Data and Training Details

### B.1.  Included Pieces from *The Well-Tempered Clavier*, Book I

The following table 4 lists the 24 Prelude–Fugue pairs from J.S. Bach's *The Well-Tempered Clavier*, Book I, that were included in the centroid-based latent space analysis. Each pair represents one of the 24 major and minor keys, providing a complete tonal coverage of the book. All pieces were preprocessed, tempo-normalized, and encoded prior to projection into PCA space.

| Section | Piece (Book I) | Key |
| --- | --- | --- |
| Section 1 | Prelude and Fugue 1 | C major |
| Section 2 | Prelude and Fugue 2 | C minor |
| Section 3 | Prelude and Fugue 3 | C# major |
| Section 4 | Prelude and Fugue 4 | C# minor |
| Section 5 | Prelude and Fugue 5 | D major |
| Section 6 | Prelude and Fugue 6 | D minor |
| Section 7 | Prelude and Fugue 7 | E♭ major |
| Section 8 | Prelude and Fugue 8 | D♯ minor |
| Section 9 | Prelude and Fugue 9 | E major |
| Section 10 | Prelude and Fugue 10 | E minor |
| Section 11 | Prelude and Fugue 11 | F major |
| Section 12 | Prelude and Fugue 12 | F minor |
| Section 13 | Prelude and Fugue 13 | F# major |
| Section 14 | Prelude and Fugue 14 | F# minor |
| Section 15 | Prelude and Fugue 15 | G major |
| Section 16 | Prelude and Fugue 16 | G minor |
| Section 17 | Prelude and Fugue 17 | A♭ major |
| Section 18 | Prelude and Fugue 18 | G# minor |
| Section 19 | Prelude and Fugue 19 | A major |
| Section 20 | Prelude and Fugue 20 | A minor |
| Section 21 | Prelude and Fugue 21 | B♭ major |
| Section 22 | Prelude and Fugue 22 | B♭ minor |
| Section 23 | Prelude and Fugue 23 | B major |
| Section 24 | Prelude and Fugue 24 | B minor |

Table 4: Included Prelude-Fugue pairs from *The Well-Tempered Clavier*, Book I by J.S. Bach.

### B.2.  Training Configuration

**Experiment 1: Sequence Length 16**
- **Latent Dimension:** 128
- **Encoder Layers:** [1024, 512, 256]
- **Activation Function:** Swish
- **Dropout Rate:** 0.1
- **Loss Function:** Binary Cross-Entropy

- **Learning Rate:** 0.0005
- **Batch Size:** 1024
- **Epochs:** 1000 (Early stopping patience: 50)
- **L2 Regularization:** 0.001
- **Validation Split:** 0.2
- **LR Scheduler:** Reduce on Plateau (patience = 5, factor = 0.5, min LR = 1e-5)