

Oregon State University

Probing for Advanced Music Theory Concepts in Generative Music Models

EvoMUSART: AI Across Music, Image, Language, and Sound

Derek Kwan and Patrick J. Donnelly

Session 3A: April 8, 2026

soundbendor lab

Contributions and Motivation

Contributions and Motivation

Contributions

Contributions and Motivation

Contributions

- Created **SynTheory-Plus** dataset

Contributions and Motivation

Contributions

- Created **SynTheory-Plus** dataset
  - 153,876 raw audio files of five advanced music theory (MT) concepts (rhythm, loudness, harmony)

Contributions and Motivation

Contributions

- Created **SynTheory-Plus** dataset
  - 153,876 raw audio files of five advanced music theory (MT) concepts (rhythm, loudness, harmony)
- Probe OpenAI Jukebox [1], Meta MusicGen [2], Encodec [3] for dataset concepts

Contributions and Motivation

Contributions

- Created **SynTheory-Plus** dataset
  - 153,876 raw audio files of five advanced music theory (MT) concepts (rhythm, loudness, harmony)
- Probe OpenAI Jukebox [1], Meta MusicGen [2], Encodec [3] for dataset concepts
  - Train a concept property classifier [4] on frozen models

Contributions and Motivation

Contributions

- Created **SynTheory-Plus** dataset
  - 153,876 raw audio files of five advanced music theory (MT) concepts (rhythm, loudness, harmony)
- Probe OpenAI Jukebox [1], Meta MusicGen [2], Encodec [3] for dataset concepts
  - Train a concept property classifier [4] on frozen models
  - High accuracy = encoded

Contributions and Motivation

Contributions

- Created **SynTheory-Plus** dataset
  - 153,876 raw audio files of five advanced music theory (MT) concepts (rhythm, loudness, harmony)
- Probe OpenAI Jukebox [1], Meta MusicGen [2], Encodec [3] for dataset concepts
  - Train a concept property classifier [4] on frozen models
  - High accuracy = encoded

Contributions and Motivation

Contributions

- Created **SynTheory-Plus** dataset
  - 153,876 raw audio files of five advanced music theory (MT) concepts (rhythm, loudness, harmony)
- Probe OpenAI Jukebox [1], Meta MusicGen [2], Encodec [3] for dataset concepts
  - Train a concept property classifier [4] on frozen models
  - High accuracy = encoded

Motivation

Contributions and Motivation

Contributions

- Created **SynTheory-Plus** dataset
  - 153,876 raw audio files of five advanced music theory (MT) concepts (rhythm, loudness, harmony)
- Probe OpenAI Jukebox [1], Meta MusicGen [2], Encodec [3] for dataset concepts
  - Train a concept property classifier [4] on frozen models
  - High accuracy = encoded

Motivation

- Audio = indirect representation of music

Contributions and Motivation

Contributions

- Created **SynTheory-Plus** dataset
  - 153,876 raw audio files of five advanced music theory (MT) concepts (rhythm, loudness, harmony)
- Probe OpenAI Jukebox [1], Meta MusicGen [2], Encodec [3] for dataset concepts
  - Train a concept property classifier [4] on frozen models
  - High accuracy = encoded

Motivation

- Audio = indirect representation of music
- Encoding → concept representable by model

Contributions and Motivation

Contributions

- Created **SynTheory-Plus** dataset
  - 153,876 raw audio files of five advanced music theory (MT) concepts (rhythm, loudness, harmony)
- Probe OpenAI Jukebox [1], Meta MusicGen [2], Encodec [3] for dataset concepts
  - Train a concept property classifier [4] on frozen models
  - High accuracy = encoded

Motivation

- Audio = indirect representation of music
- Encoding → concept representable by model
  - Need representation for generation (future work)

Contributions and Motivation

Contributions

- Created **SynTheory-Plus** dataset
  - 153,876 raw audio files of five advanced music theory (MT) concepts (rhythm, loudness, harmony)
- Probe OpenAI Jukebox [1], Meta MusicGen [2], Encodec [3] for dataset concepts
  - Train a concept property classifier [4] on frozen models
  - High accuracy = encoded

Motivation

- Audio = indirect representation of music
- Encoding → concept representable by model
  - Need representation for generation (future work)
- MT concepts underlie compositional practice, syntax

Dataset Creation

Dataset Creation

SynTheory-Plus

Dataset Creation

SynTheory-Plus

- extends Wei et al. (SynTheory) [5]

Dataset Creation

SynTheory-Plus

- extends Wei et al. (SynTheory) [5]
  - basic MT concepts: notes, intervals, scales, triad chords, triadic chord progressions, tempi, time signatures

Dataset Creation

SynTheory-Plus

- extends Wei et al. (SynTheory) [5]
  - basic MT concepts: notes, intervals, scales, triad chords, triadic chord progressions, tempi, time signatures

Dataset Creation

SynTheory-Plus

- extends Wei et al. (SynTheory) [5]
  - basic MT concepts: notes, intervals, scales, triad chords, triadic chord progressions, tempi, time signatures

Rhythm and loudness

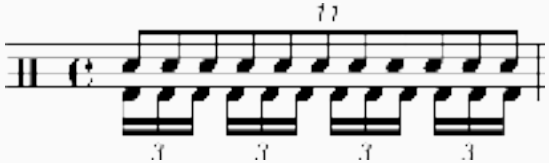
Dataset Creation

SynTheory-Plus

- extends Wei et al. (SynTheory) [5]
  - basic MT concepts: notes, intervals, scales, triad chords, triadic chord progressions, tempi, time signatures

Rhythm and loudness

- Polyrhythms



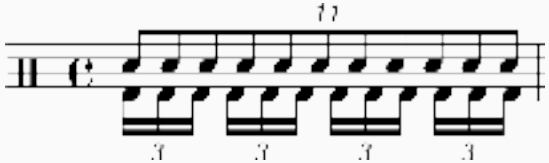
Dataset Creation

SynTheory-Plus

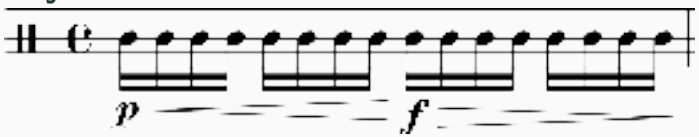
- extends Wei et al. (SynTheory) [5]
  - basic MT concepts: notes, intervals, scales, triad chords, triadic chord progressions, tempi, time signatures

Rhythm and loudness

- Polyrhythms



- Dynamics



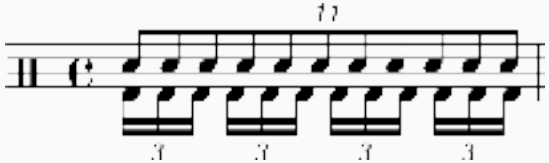
Dataset Creation

SynTheory-Plus

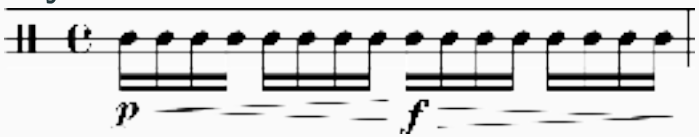
- extends Wei et al. (SynTheory) [5]
  - basic MT concepts: notes, intervals, scales, triad chords, triadic chord progressions, tempi, time signatures

Rhythm and loudness

- Polyrhythms



- Dynamics



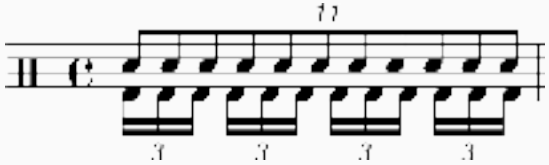
Dataset Creation

SynTheory-Plus

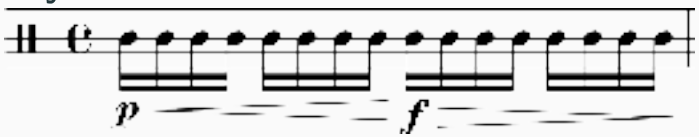
- extends Wei et al. (SynTheory) [5]
  - basic MT concepts: notes, intervals, scales, triad chords, triadic chord progressions, tempi, time signatures

Rhythm and loudness

- Polyrhythms



- Dynamics



Harmony

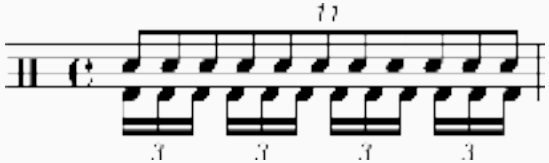
Dataset Creation

SynTheory-Plus

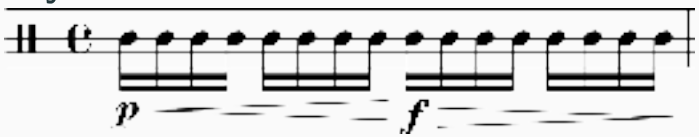
- extends Wei et al. (SynTheory) [5]
  - basic MT concepts: notes, intervals, scales, triad chords, triadic chord progressions, tempi, time signatures

Rhythm and loudness

- Polyrhythms

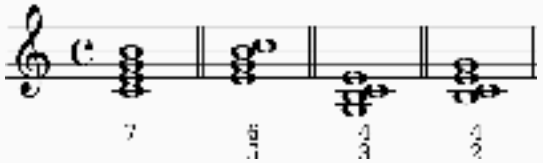


- Dynamics



Harmony

- Seventh chords



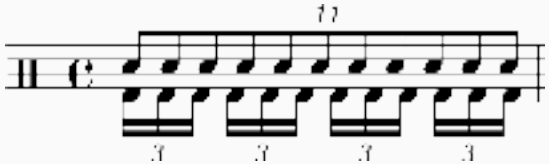
Dataset Creation

SynTheory-Plus

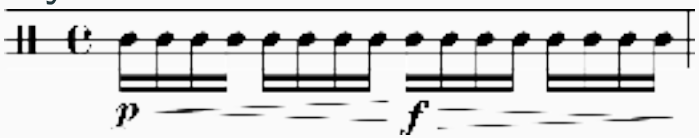
- extends Wei et al. (SynTheory) [5]
  - basic MT concepts: notes, intervals, scales, triad chords, triadic chord progressions, tempi, time signatures

Rhythm and loudness

- Polyrhythms

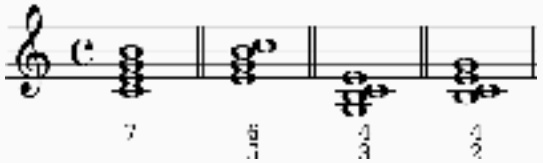


- Dynamics

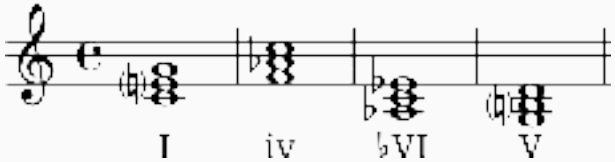


Harmony

- Seventh chords



- Mode mixture



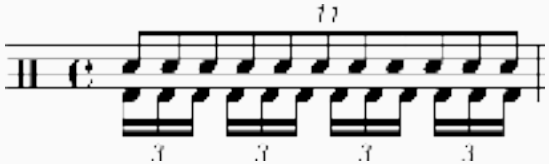
Dataset Creation

SynTheory-Plus

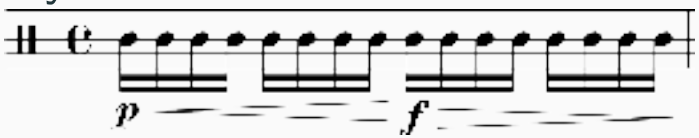
- extends Wei et al. (SynTheory) [5]
  - basic MT concepts: notes, intervals, scales, triad chords, triadic chord progressions, tempi, time signatures

Rhythm and loudness

- Polyrhythms

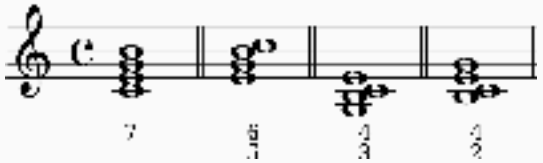


- Dynamics

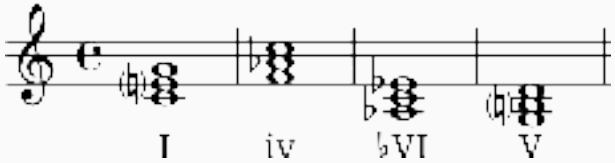


Harmony

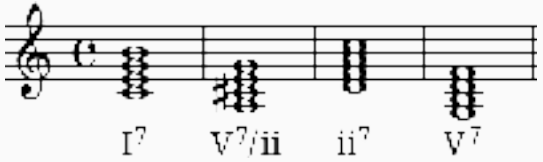
- Seventh chords



- Mode mixture



- Secondary dominants



Methods

Methods

Dataset preparation

Methods

Dataset preparation

- Generate MIDI files dataset examples

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs
- Input WAVs, extract decoder layer outputs from Jukebox, Encodec (MusicGen encoder), and MusicGen-(small/medium/large)

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs
- Input WAVs, extract decoder layer outputs from Jukebox, Encodec (MusicGen encoder), and MusicGen-(small/medium/large)

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs
- Input WAVs, extract decoder layer outputs from Jukebox, Encodec (MusicGen encoder), and MusicGen-(small/medium/large)

Classification

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs
- Input WAVs, extract decoder layer outputs from Jukebox, Encodec (MusicGen encoder), and MusicGen-(small/medium/large)

Classification

- Train linear layer and MLP classifiers on outputs

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs
- Input WAVs, extract decoder layer outputs from Jukebox, Encodec (MusicGen encoder), and MusicGen-(small/medium/large)

Classification

- Train linear layer and MLP classifiers on outputs
  - Adam optimizer, CE loss

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs
- Input WAVs, extract decoder layer outputs from Jukebox, Encodec (MusicGen encoder), and MusicGen-(small/medium/large)

Classification

- Train linear layer and MLP classifiers on outputs
  - Adam optimizer, CE loss
- Experiments per-layer, per-model, per-dataset

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs
- Input WAVs, extract decoder layer outputs from Jukebox, Encodec (MusicGen encoder), and MusicGen-(small/medium/large)

Classification

- Train linear layer and MLP classifiers on outputs
  - Adam optimizer, CE loss
- Experiments per-layer, per-model, per-dataset
  - Polyrhythms - 34 classes (polyrhythms)

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs
- Input WAVs, extract decoder layer outputs from Jukebox, Encodec (MusicGen encoder), and MusicGen-(small/medium/large)

Classification

- Train linear layer and MLP classifiers on outputs
  - Adam optimizer, CE loss
- Experiments per-layer, per-model, per-dataset
  - Polyrhythms - 34 classes (polyrhythms)
  - Dynamics - 7 classes (dynamic expression)

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs
- Input WAVs, extract decoder layer outputs from Jukebox, Encodec (MusicGen encoder), and MusicGen-(small/medium/large)

Classification

- Train linear layer and MLP classifiers on outputs
  - Adam optimizer, CE loss
- Experiments per-layer, per-model, per-dataset
  - Polyrhythms - 34 classes (polyrhythms)
  - Dynamics - 7 classes (dynamic expression)
  - Seventh chords - 8 classes (chord quality)

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs
- Input WAVs, extract decoder layer outputs from Jukebox, Encodec (MusicGen encoder), and MusicGen-(small/medium/large)

Classification

- Train linear layer and MLP classifiers on outputs
  - Adam optimizer, CE loss
- Experiments per-layer, per-model, per-dataset
  - Polyrhythms - 34 classes (polyrhythms)
  - Dynamics - 7 classes (dynamic expression)
  - Seventh chords - 8 classes (chord quality)
  - Mode mixture - 2 classes (has mode-mixture)

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs
- Input WAVs, extract decoder layer outputs from Jukebox, Encodec (MusicGen encoder), and MusicGen-(small/medium/large)

Classification

- Train linear layer and MLP classifiers on outputs
  - Adam optimizer, CE loss
- Experiments per-layer, per-model, per-dataset
  - Polyrhythms - 34 classes (polyrhythms)
  - Dynamics - 7 classes (dynamic expression)
  - Seventh chords - 8 classes (chord quality)
  - Mode mixture - 2 classes (has mode-mixture)
  - Secondary Dominants - 3 classes (normal, 2ndary dom., tritone sub.)

Methods

Dataset preparation

- Generate MIDI files dataset examples
- Synthesize MIDIs as WAVs
- Input WAVs, extract decoder layer outputs from Jukebox, Encodec (MusicGen encoder), and MusicGen-(small/medium/large)

Classification

- Train linear layer and MLP classifiers on outputs
  - Adam optimizer, CE loss
- Experiments per-layer, per-model, per-dataset
  - Polyrhythms - 34 classes (polyrhythms)
  - Dynamics - 7 classes (dynamic expression)
  - Seventh chords - 8 classes (chord quality)
  - Mode mixture - 2 classes (has mode-mixture)
  - Secondary Dominants - 3 classes (normal, 2ndary dom., tritone sub.)
- Compare with baselines (audio features)

Results

Our Linear Probe Results

Model	Poly-rhythms	Dyn.	Seventh Chords	Mode Mixture	2nd ary Dom.
Jukebox	1.0000	0.9992	0.9564	0.9970	0.9227
MGL	0.9847	0.9936	0.6863	0.8916	0.6684
MGM	0.9503	0.9825	0.5649	0.8701	0.6735
MGS	0.8481	0.9530	0.5915	0.7909	0.6153
EnC	0.2617	0.3850	0.2114	0.5532	0.3784
Mel.	0.2928	0.6484	0.2914	0.6600	0.4031
MFCC	0.1776	0.5964	0.2510	0.5955	0.4230
Chroma	0.1251	0.2258	0.1665	0.6945	0.5846
Concat.	0.4257	0.7879	0.3656	0.7966	0.6365

Our MLP Probe Results

Model	Poly-rhythms	Dyn.	Seventh Chords	Mode Mixture	2nd ary Dom.
Jukebox	1.0000	0.9990	0.9958	0.9980	0.9814
MGL	0.9998	0.9684	0.9764	0.9919	0.7560
MGM	0.9976	0.9527	0.9751	0.9802	0.6574
MGS	0.9935	0.9432	0.9724	0.9936	0.7137
EnC	0.7155	0.5376	0.6738	0.8916	0.5279
Mel.	0.4871	0.6572	0.8399	0.9671	0.6203
MFCC	0.7240	0.7720	0.5040	0.8100	0.5783
Chroma	0.3695	0.3933	0.9560	0.9973	0.8708
Concat.	0.7608	0.8600	0.9322	0.9879	0.7825

$\geq$ .90, $\geq$ 2·chance, $\geq$ chance

Takeaways

Takeaways

- Showed advanced music theory concepts encoded in the latent spaces of OpenAI's Jukebox and Meta's MusicGen and EnCodec

Takeaways

- Showed advanced music theory concepts encoded in the latent spaces of OpenAI's Jukebox and Meta's MusicGen and EnCodec
  - with MLP and linear classifiers

Takeaways

- Showed advanced music theory concepts encoded in the latent spaces of OpenAI's Jukebox and Meta's MusicGen and EnCodec
  - with MLP and linear classifiers
  - advances on Wei et al. [5] demonstrating basic MT concept encoding (just MLP)

Takeaways

- Showed advanced music theory concepts encoded in the latent spaces of OpenAI's Jukebox and Meta's MusicGen and EnCodec
  - with MLP and linear classifiers
  - advances on Wei et al. [5] demonstrating basic MT concept encoding (just MLP)
  - have not shown generation possibility, but can be used for generation steering [6], [7], [8]

Takeaways

- Showed advanced music theory concepts encoded in the latent spaces of OpenAI's Jukebox and Meta's MusicGen and EnCodec
  - with MLP and linear classifiers
  - advances on Wei et al. [5] demonstrating basic MT concept encoding (just MLP)
  - have not shown generation possibility, but can be used for generation steering [6], [7], [8]
- Created SynTheory-Plus isolating advanced music theory concepts

Takeaways

- Showed advanced music theory concepts encoded in the latent spaces of OpenAI's Jukebox and Meta's MusicGen and EnCodec
  - with MLP and linear classifiers
  - advances on Wei et al. [5] demonstrating basic MT concept encoding (just MLP)
  - have not shown generation possibility, but can be used for generation steering [6], [7], [8]
- Created SynTheory-Plus isolating advanced music theory concepts

Takeaways

- Showed advanced music theory concepts encoded in the latent spaces of OpenAI's Jukebox and Meta's MusicGen and EnCodec
  - with MLP and linear classifiers
  - advances on Wei et al. [5] demonstrating basic MT concept encoding (just MLP)
  - have not shown generation possibility, but can be used for generation steering [6], [7], [8]
- Created SynTheory-Plus isolating advanced music theory concepts

huggingface.co/datasets/derekxkwan/syntheory_plus

References

1. P. Dhariwal, H. Jun, C. Payne, J. W. Kim, A. Radford, and I. Sutskever, **Jukebox: A generative model for music,** last accessed 2025/10/21, 2020. DOI: 10.48550/arXiv.2005.00341.

2. J. Copet et al., **"Simple and controllable music generation,"** in *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Eds., last accessed 2025/10/21, vol. 36, Curran Associates, Inc., 2023, pp. 47 704–47 720. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2023/hash/94b472a1842cd7c56dcb125fb2765fbd-Abstract-Conference.html.

3. A. Défossez, J. Copet, G. Synnaeve, and Y. Adi, **"High fidelity neural audio compression,"** *Transactions on Machine Learning Research*, 2023, last accessed 2025/10/21, ISSN: 2835-8856. [Online]. Available: https://openreview.net/forum?id=ivCd8z8zR2.

4. Y. Belinkov, **"Probing classifiers: Promises, shortcomings, and advances,"** *Computational Linguistics*, vol. 48, no. 1, pp. 207–219, Mar. 2022. DOI: 10.1162/coli_a_00422.

5. M. Wei, M. Freeman, C. Donahue, and C. Sun, **"Do music generation models encode music theory?"** In *Proceedings of the 25th International Society for Music Information Retrieval Conference*, (San Francisco, California, USA and Online), last accessed 2025/08/19, ISMIR, Nov. 2024, pp. 680–687. DOI: 10.5281/zenodo.14877427.

6. D. Panda et al., **Fine-grained control over music generation with activation steering,** last accessed 2025/10/05, 2025. DOI: 10.48550/arXiv.2506.10225.

7. D. Zhao, D. Beaglehole, T. Berg-Kirkpatrick, J. McAuley, and Z. Novack, **Steering autoregressive music generation with recursive feature machines,** last accessed 2025/10/23, Oct. 21, 2025. DOI: 10.48550/arXiv.2510.19127. arXiv: 2510.19127[cs]. Accessed: Oct. 23, 2025.

8. S. Facchiano et al., **Activation patching for interpretable steering in music generation,** last accessed 2025/10/05, 2025. DOI: 10.48550/arXiv.2504.04479.