IMPORTANT POINTS
from
Fundamentals of Music Processing

Audio, Analysis, Algorithms, Applications

Meinard Müller

CHAPTER -1
Music Representations

1. **Score**: Music symbols called scores are used to represent notes
2. **Sheet music**: Printed form of musical score
3. **MIDI**: Musical Instrumental Digital Interface protocol where event messages specify pitches, velocities, and other important parameters
4. **Notes**: May refer to both a musical symbol as well as a pitched sound.
5. **Pitch**: Refers to a perceptual property that allows a listener to order a sound on a frequency- related scale.
   For example, pitch of a note in keyboard tells the player which key to press
6. **Same pitch class** : Two notes with fundamental frequencies in the ratio of two are very similar and hence grouped under the same pitch class. => leads to the notion of **octaves.** For example, the pitches A3 (220 Hz), A4 (440 Hz), and A5 (880 Hz) sound similar.
7. The note **A4** is determined to have a **fundamental frequency of 440 Hz** and serves as a **reference.**
8. All notes that have the same chroma value belong to the same pitch class.
9. Notes with different chroma values have a different "sound color."
10. Music is typically organized into temporal units, referred to as **beats**. Repeating sequences of stressed and unstressed beats, in turn, form higher temporal patterns, which are related to what is called the **rhythm** of music and is expressed in terms of the musical meter.
11. **Piano-roll representation** : The horizontal axis of this two-dimensional representation encodes time, whereas the vertical axis encodes pitch. Every note is described by an axis-parallel rectangle coding three parameters. The first parameter is the onset time, given by the leftmost horizontal coordinate of the rectangle, and the second is the pitch, given by the lower vertical coordinate of the rectangle. Finally, the third parameter is the duration of the note, encoded by the width of the rectangle.
12. **MIDI Representations:** MIDI does not represent musical sound directly, but only represents performance information encoding the instructions about how an instrument has been played or how music is to be produced.
    The most important MIDI messages are the note-on and the note-off commands, which correspond to the start and the end of a note, respectively. Each note-on and

note-off message is, among others, equipped with a MIDI note number, a value for the key velocity, a channel specification, as well as a timestamp.

The MIDI note number is an integer between 0 and 127 and encodes a **note's Pitch.** MIDI note numbers encode, in increasing order, the musical pitches C0 to G#9. For example, note C4 has the MIDI note number 60, whereas the concert pitch A4 has the MIDI note number 69.

The key velocity is again an integer between 0 and 127, which controls the intensity of the sound—in the case of a note-on event it determines the volume, whereas in the case of a note-off event it controls the decay during the release phase of the tone. The exact interpretation of the key velocity, however, depends on the respective instrument or synthesizer. The MIDI channel is an integer between 0 and 15. Intuitively speaking, this number prompts the synthesizer to use the instrument that has been previously assigned to the respective channel number. Note that each channel, in turn, supports polyphony, i.e., multiple simultaneous notes. Finally, the time stamp is an integer value that represents how many clock pulses or ticks to wait before the respective note-on or note-off command is executed. Before

(a)

(b)

Time (Ticks)	Message	Channel	Note Number	Velocity
60	NOTE ON	1	67	100
0	NOTE ON	1	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	1	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	67	100
0	NOTE ON	1	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	1	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	67	100
0	NOTE ON	1	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	1	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	63	100
0	NOTE ON	2	51	100
0	NOTE ON	2	39	100
240	NOTE OFF	1	63	0
0	NOTE OFF	2	51	0
0	NOTE OFF	2	39	0

(c)

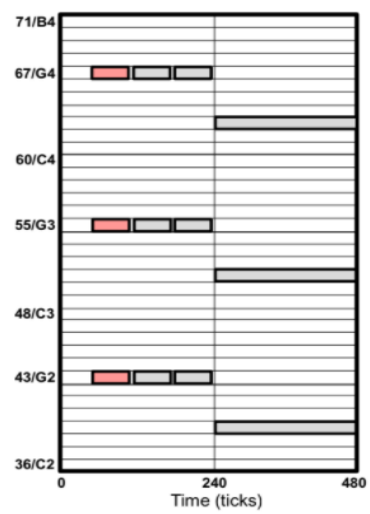


Fig. 1.13 Various symbolic music representations of the first twelve notes of Beethoven's Fifth. **(a)** Sheet music representation. **(b)** MIDI representation (in a simplified, tabular form). **(c)** Piano-roll representation.

13. **MusicXML :** MusicXML is a textual data format that defines a set of rules for encoding documents in a way that is both human and machine readable.

Fig. 1.15 Textual description in the MusicXML format of a half note $E^{\flat}4$. The clef, key signature, and time signature are defined at the beginning of the MusicXML file.

```
<note>
  <pitch>
    <step>E</step>
    <alter>-1</alter>
    <octave>4</octave>
  </pitch>
  <duration>2</duration>
  <type>half</type>
</note>
```

14. **Audio Signal :**
    An audio signal is a representation of sound. As opposed to sheet music and symbolic representations, an audio representation encodes all information needed to reproduce an acoustic realization of a piece of music. This includes the temporal, dynamic, and tonal micro deviations that make up the specific performance style of a musician.
15. **Wave :** a (mechanical) wave can be described as an oscillation that travels through space, where energy is transferred from one point to another.
16. The **period** of the wave is defined as the time required to complete a cycle.
17. The **frequency**, measured in Hertz (Hz), is the reciprocal of the period.
18. A sinusoid is completely specified by its frequency, its amplitude (the peak deviation of the sinusoid from its mean), and its phase (determining where in its cycle the sinusoid is at time zero).
19. Sometimes the sound resulting from a sinusoid is called a **harmonic sound or pure tone.**
20. **Pitch** is a subjective attribute of sound. In the case of complex sound mixtures its relation to frequency can be especially ambiguous. In the case of pure tones, however, the relation between frequency and pitch is clear. For example, a sinusoid having a frequency of 440 Hz corresponds to the pitch A4.
    This particular pitch is known as concert pitch, and it is used as the reference pitch to which a group of musical instruments are tuned for a performance. Since a slight change in frequency does not necessarily lead to a perceived change, one usually associates an entire range of frequencies with a single pitch.
    We can associate to each pitch $p \in [0:127]$ a centre frequency $F_{pitch}(p)$ (measures in Hz) by

$$F_{\mathrm{pitch}}(p) = 2^{(p-69)/12} \cdot 440.$$

21. An octave is divided into 1200 cents, so that each semitone corresponds to 100 cents. Again the ratio of frequencies one cent apart is constant, yielding the value

$$2^{1/1200} \approx 1.0005777895.$$

22. REAL LIFE SITUATION:

Real-world sounds are far from being a simple pure tone with a well-defined frequency. Playing a single note on an instrument may result in a complex sound that contains a mixture of different frequencies changing over time. Intuitively, such a musical tone can be described as a superposition of pure tones or sinusoids, each with its own frequency of vibration, amplitude, and phase. A partial is any of the sinusoids by which a musical tone is described. The frequency of the lowest partial present is called the **fundamental frequency of the sound.**

A harmonic (or a harmonic partial) is a partial that is an integer multiple of the fundamental frequency. Partials, as well as harmonics, are counted upwards along the frequency axis. This convention implies that the fundamental frequency is the first partial, as well as the first harmonic of a musical tone. The term inharmonicity is used to denote a measure of the deviation of a partial from the closest ideal harmonic, typically measured in cents for each partial. Finally, another term that is often used in music theory is the overtone, which is any partial except the lowest. This can lead to numbering confusion when comparing overtones with partials, since the first overtone is the second partial.

23. Nonpitched, or indefinite-pitched instruments, such as cymbals, gongs, or tam-tams, make sounds rich in inharmonic partials.

24. Dynamics, Intensity, and Loudness

On the audio side, dynamics correlate with a perceptual property called loudness, by which sounds can be ordered on a scale extending from quiet to loud.

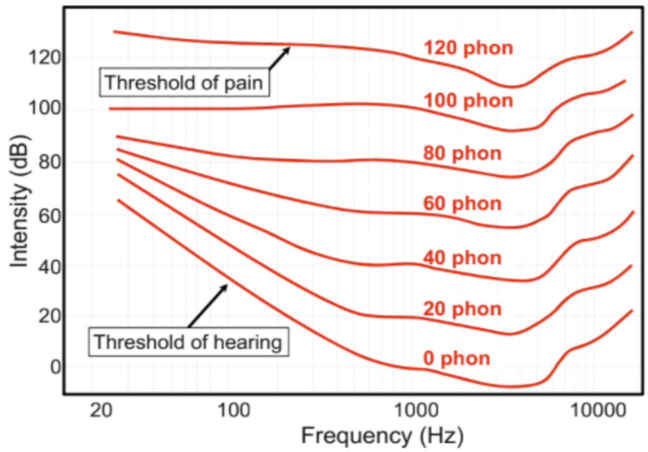
Loudness is a subjective measure which correlates to objective measures of sound intensity and sound power.

A **decibel (dB)** scale is a logarithmic unit expressing the ratio between two values.

$$\mathrm{dB}(I) := 10 \cdot \log_{10}\left(\frac{I}{I_{\mathrm{TOH}}}\right).$$

Two sounds with the same intensity but different frequencies are generally not perceived to have the same loudness.

The perceived loudness of pure tones depending on the frequency has been determined and expressed by the unit **phon.**



25. Timbre or tone color:

Timbre allows a listener to distinguish the musical tone of a violin, an oboe, or a trumpet even if the tone is played at the same pitch and with the same loudness. As with pitch and loudness, timbre is a perceptual property of sound.

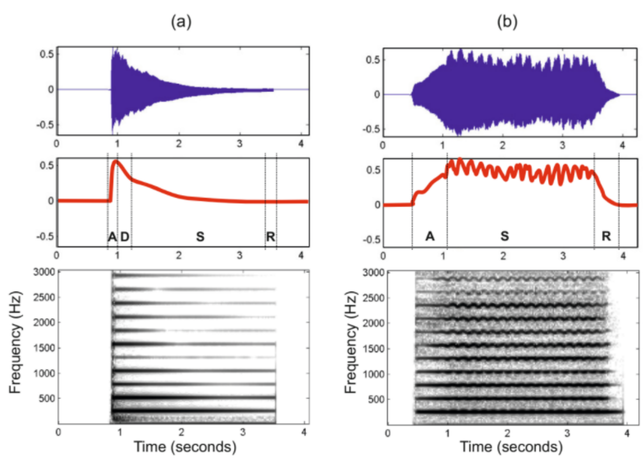


Fig. 1.23 Waveform, amplitude envelope, and spectrogram representation for different instruments playing the same note C4 (261.6 Hz). **(a)** Piano. **(b)** Violin.

Timbre is a multidimensional phenomenon that is hard to measure. It is the irregularities and variations that make a musical tone sound interesting and that give it a particular and natural quality.

26. Spectrogram :

The composition of a sound in terms of its partials can be visualized by a so-called spectrogram, which shows the intensity of the occurring frequencies over time.

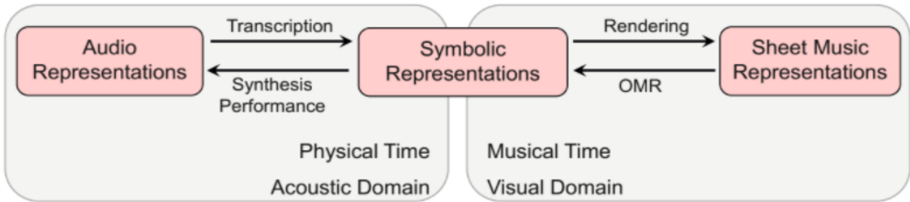


Fig. 1.24 Illustration of three classes of music representation and their relations.

All the images and the text have been taken from the chapter-1 of the book itself.

Happy Music Learning!