Computer Music - MIDI

Sound waves, whether occurred natural or man-made, are often very complex, i.e., they consist of many frequencies. Digital sound is relatively straight forward to record complex sound. However, it is quite difficult to generate (or synthesize) complex sound.

There is a better way to generate high quality music. This is known as MIDI — Musical Instrument Digital Interface.

It is a communication standard developed in the early 1980s for electronic instruments and computers. It specifies the hardware connection between equipments as well as the format in which the data are transfered between the equipments.

Common MIDI devices include electronic music synthesisers, modules, and MIDI devices in common sound cards.

General MIDI is a standard specified by MIDI Manufacturers Association. To be GM compatible, a sound generating device must meet the General MIDI system level 1 performance requirement.

3.1 MIDI Hardware

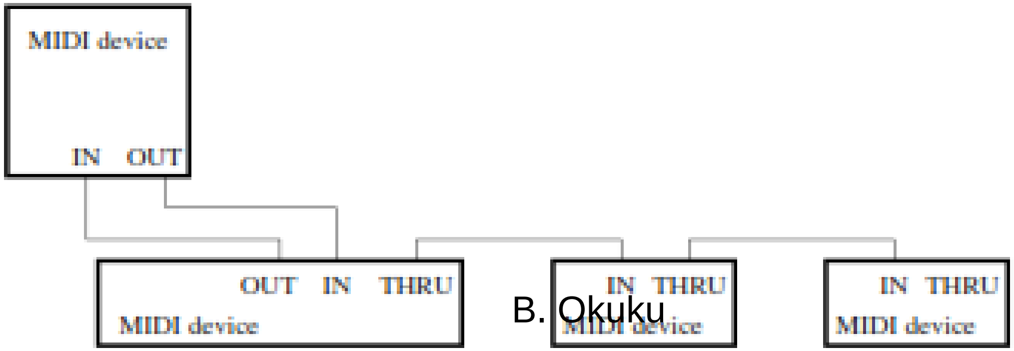
An electronic musical instrument or a computer which has MIDI interface should has one or more MIDI ports. The MIDI ports on musical instruments are usually labelled with:

IN — for receiving MIDI data;

OUT — for outputting MIDI data that are generated by the instrument;

THRU — for passing MIDI data to the next instrument.

MIDI devices can be daisy-chained together.



3.2 MIDI Data

Unlike digital sound, MIDI data does not encode individual samples. MIDI data encode musical events and commands to control instruments.
MIDI data are grouped into MIDI messages. Each MIDI message represents a musical event, e.g., pressing a key, setting a switch or adjusting foot pedals. A sequence of MIDI messages is grouped into a track.
An instrument or a computer satisfies both the hardware interface and the data format is known as a MIDI device.

3.3 MIDI Channels and Modes

MIDI devices communicate with each other through channels. The MIDI standard specifies each MIDI connection has 16 channels. Each instrument can be mapped to a single channel (omni Off), or it can use all 16 channels (Omni On).

Some instruments are capable of playing more than one note at the same time, e.g., organs and piano. This is known as polyphony. Other instruments, such as flute, is monophony since they can only play one note at a time.

Each MIDI device must be set to one of the modes for receiving MIDI data:

3.4 Instrument Patch

Each MIDI device is usually capable of producing sound resembling several real instruments and/or noise effects (e.g., telephone, aircraft). Each instrument or noise effect is known as a patch, or preset.

The general MIDI standard specifies 128 patches (ranges from 0 to 127).

ID	Sound	ID	Sound
0	Acoustic grand piano	35	Acoustic bass drum
7	Clarinet	45	Low tom
10	Music box	76	High wood block
19	Church organ	80	Mute triangle
40	Violin		
48	String ensemble I		
124	Telephone ring		
125	Helicopter		
126	Applause		

3.5 MIDI files

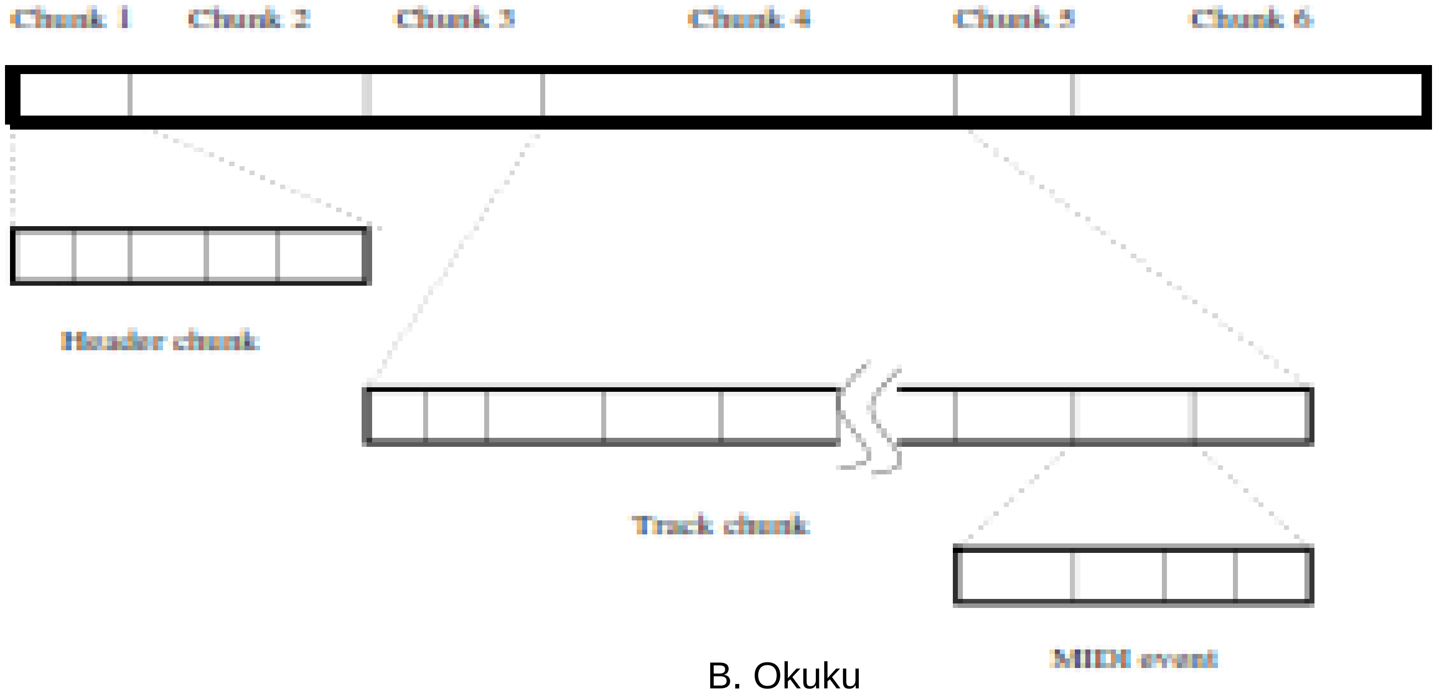
When using computers to play MIDI music, the MIDI data are often stored in MIDI files. Each MIDI files contains a number of chunks. There are two types of chunks:

- Header chunk — contains information about the entire file: the type of MIDI file, number of tracks and the timing.

- Track chunk — the actual data of MIDI track.

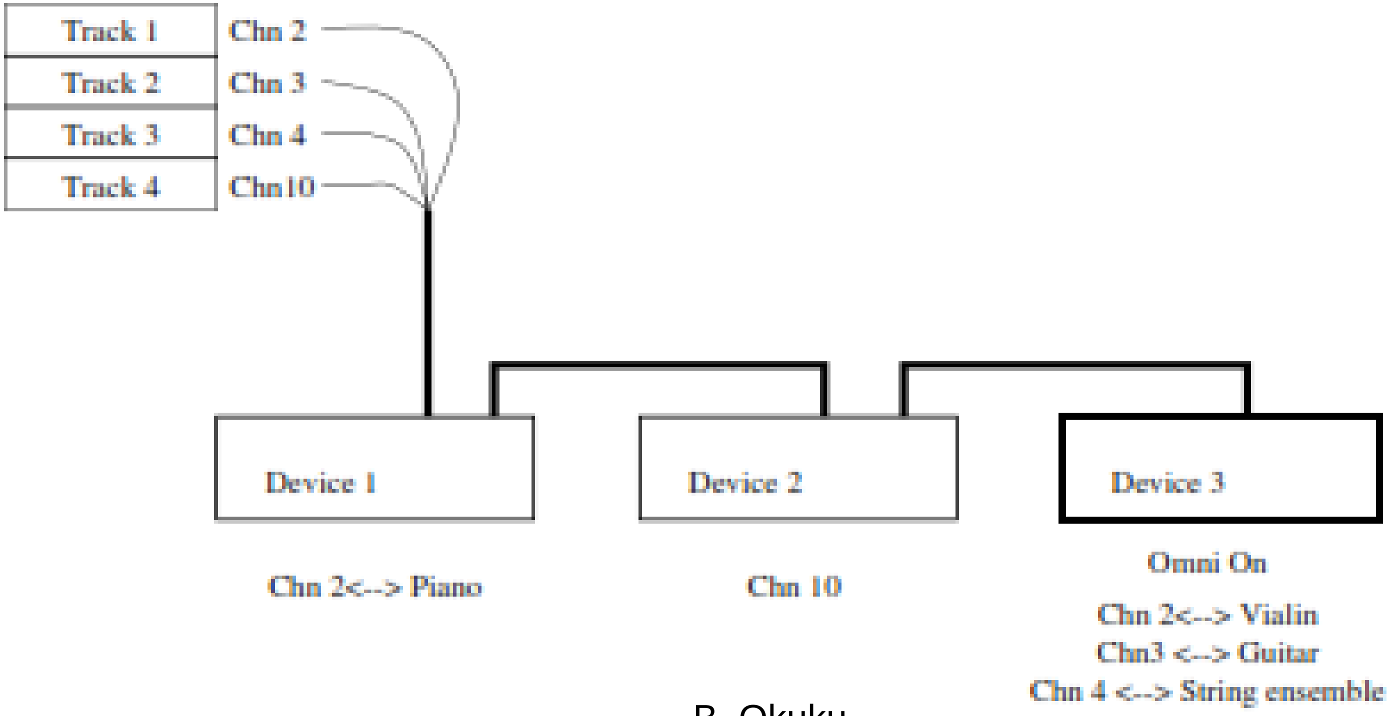
There are three types of MIDI file:

- •0 single multichannel track
- •1 one or more simultaneous track of a
- •sequence
- •2 one or more sequentially independent single-track patterns



Tracks, channels and patches

- Multiple tracks can be played at the same time.
- Each track can be assigned to a different channel.
- Each channel can accept more than one track.
- Each channel is assigned a patch, therefore generates sound of a particular instrument



3.6 How MIDI Sounds Are Synthesized

A simplistic view is that:

• the MIDI device stores the characteristics of sounds produced by different sound sources;

• the MIDI messages tell the device which kind of sound, at which pitch is to be generated, how long the sound is played and other attributes the note should have.

There are two ways of synthesizing sounds:

- FM Synthesis (Frequency Modulation) —Using one sine wave to modulate another sine wave, thus generating a new wave which is rich in timbre. It consists of the two original waves, their sum and difference and harmonics.

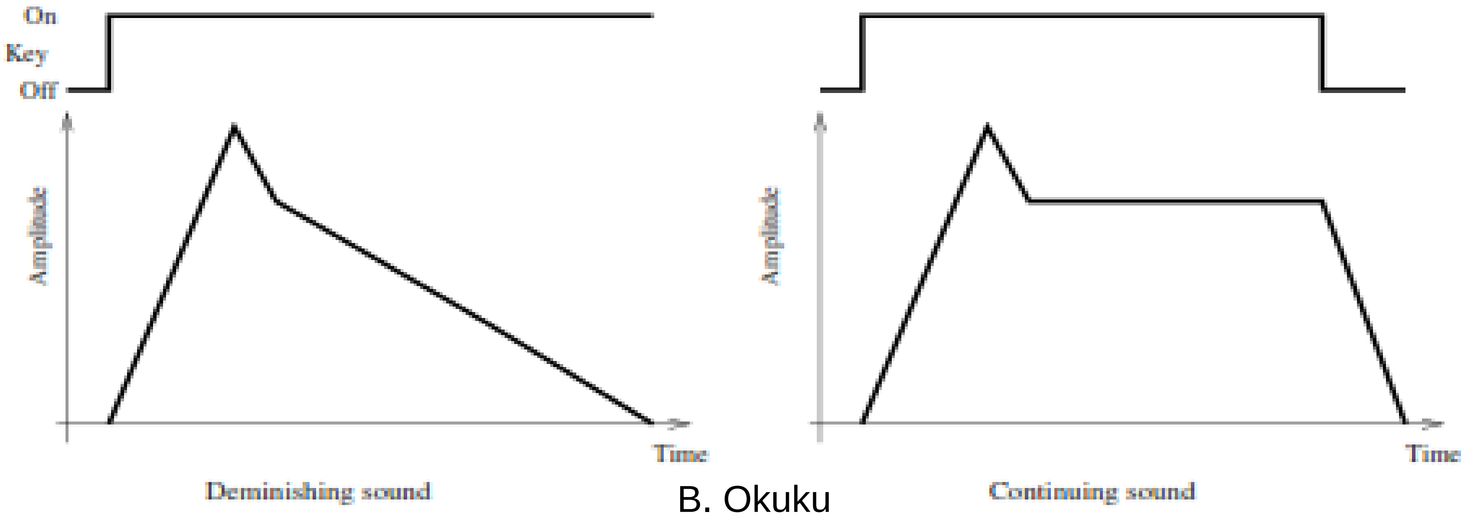
The drawbacks of FM synthesis are: the generated sound is not real; there is no exact formula for generating a particular sound.

- Wave-table synthesis — It stores representative digital sound samples. It manipulates these sample, e.g., by changing the pitch, to create the complete range of notes.

MIDI Sound Attributes

The shape of the amplitude envelop has great influence on the resulting character of sound. There are two different types of envelop:

- Diminishing sound — gradually die out;
- Continuing sound — sustain until turned off.

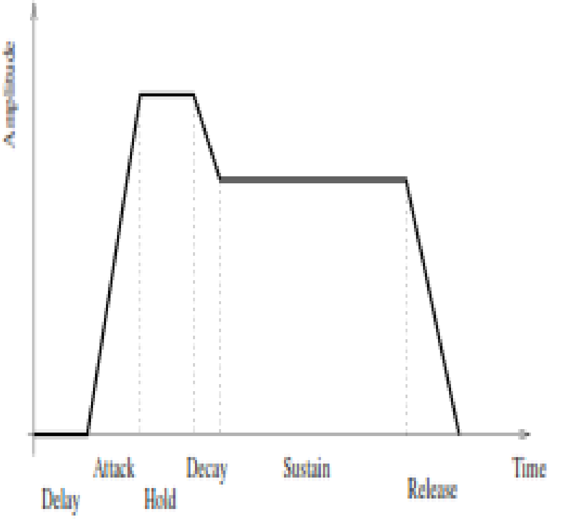


The Amplitude Envelop

•Delay — the time between when a key is played and when the attack phase begins

•Attack — the time from no sound to maximum amplitude

•Hold —the time envelop will stay at the peak level beform starting the decay phase



•Decay — the time it takes the envelop to go from the peak level to the sustain level

• Sustain — the level at which the envelop remains as long as a key is held down

• Release — the time is takes for the sound to fade to nothing

3.7 MIDI software

MIDI player for playing MIDI music. This includes:

- Windows media player can play MIDI files
- Player come with sound card — Creative Midi player
- Freeware and shareware players and plug-ins— Midigate, Yamaha Midplug, etc.

MIDI sequencer for recording, editing and playing MIDI

- Cakewalk Express, Home Studio, Professional
- Cubasis
- Encore
- Voyetra MIDI Orchestrator Plus

- Configuration — Like audio devices, MIDI devices require a driver. Select and configure MIDI devices from the control panel.

4 Summary – MIDI versus digital audio

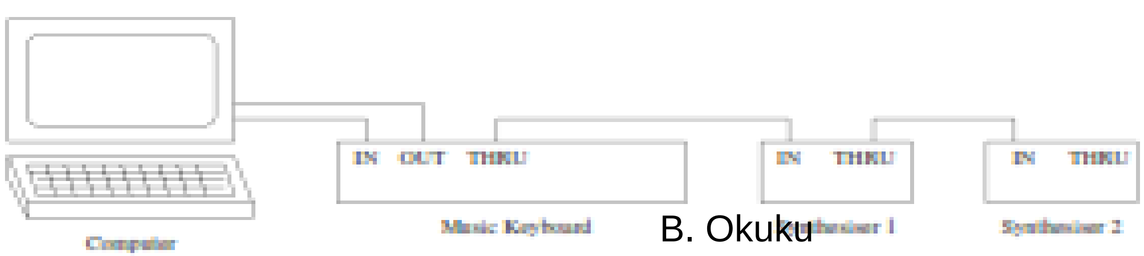
Digital Audio	MIDI
• Digital representation of physical sound waves	• Abstract representation of musical sounds and sound effects
• File size is large if without compression	• MIDI files are much more compact
• Quality is in proportion to file size	• File size is independent to the quality
• More software available	• Much better sound if the sound source is of high quality
• Play back quality less dependent on the sound sources	• Need some music theory
• Can record and play back any sound including speech	• Cannot generate speech

5 Exercises

1. Your hard disk has 256Mbytes of free space. You are going to record a speech with a sampling rate of 11KHz, 8-bit resolution and a single channel. What is the length of the recording that can be stored in the hard disk? (Answer in secs)
2. A multimedia presentation has 30 minutes of CD-quality digital audio in .wav files. What is the storage required for these files?
3. In a teleconferencing system, a network connection having bandwidth of 100Kbytes/sec is allocated to duplex audio link. What is the maximum sampling rate and resolution in which uncompressed audio data can be transmitted in real-time?
4. You are developing a network voice communication program. It uses the Internet to connect two remote users and allows them to talk to each other in real time. What is the most appropriate sampling rate for recording their voice?
5. In the configuration shown below, the music

keyboard is a general MIDI device which is set to ‘Omni On’ mode, Synthesiser 1 is a drum machine which accepts signals on the drum channel only, and Synthesiser 2 is a guitar accepting signals from all channels but can only synthesise guitar sound. A MIDI sequencer is playing a MIDI file on the computer. The table below lists tracks with their
patch and channel settings. What kind of sound will
be heard from each device?

Track	Patch	Channel	Track	Patch	Channel
1	2	3	2	8	4
3	28	5	4	49	6
5	57	7	6	73	2



5.1 Solution to Exercises

1. The size of one second of recording is 11,050 bytes.
the length of recording that can be stored in 256MBytes is

$$(256 \times 1024 \times 1024)/11,025 = 24,347(seconds)$$

2. CD-quality digital audio means 44,100Hz Sampling rate, 16-bit resolution and stereo.

$$44100 \times 2 \times 2 \times (30 \times 60)$$
$$= 317,510,000 \text{ bytes}$$
$$= 302.8 \text{ MBytes}$$

3. Since the link is duplex, we will allocate half the bandwidth to each direction, i.e., 50Kbytes/sec. Using the formula below $S = R \times (b/8) \times C \times D,$

and assuming we record in mono, we have

$50,000 = R \times (b/8)$ Since it is a teleconferencing application, the sound will mostly be speech. We can use lower sampling rate and higher resolution. Therefore, we can use 22.05KHz sampling rate and 16-bit resolution. The most important limit is that the data can not be larger than the bandwidth, otherwise, we will not be able to transmit them.

4. Because the data we want to handle are voice, a sampling rate of 11KHz and mono will be enough.
This creates a data rate of 11K/sec. If the users are connected to the Internet via a LAN, this data rate should be handled by the system adequately.
However, if they are connected via modems, the data rate will be too high. Compression technique has to be used.

5. Since the music keyboard can accept signal from all channels, the general MIDI patch map, we will hear the following sounds:

Track	Patch	Channel	Sound
1	2	3	Bright Acoustic Piano
2	8	4	Clavinet Chromatic
3	28	5	Electric Guitar
4	49	6	String Ensemble 1
5	57	7	Trumpet
6	73	2	Piccolo

Synthesiser 2 will produce no sound because no signal is sent to Channel 10. Synthesiser 3 will produce sound for signals from all channels but they all sound like guitar.