MUSICAL SOUNDS

Properties, analysis
and additive resynthesis

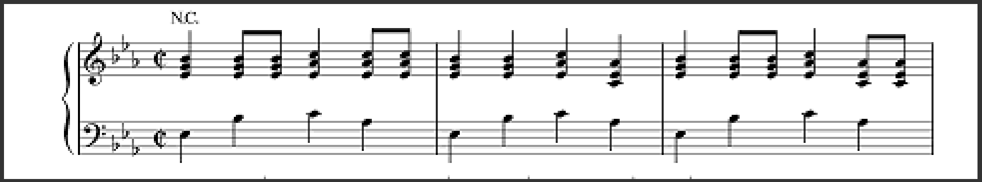
Musical sound

Definition: a musical sound is a sound produced by a musical instrument.

Musical sounds have the following features:

- pitch – placement on a musical scale
- timbre – how it “sounds”
- loudness
- duration

Loudness and timbre are usually dynamic – they change as the sound plays.



Musical sounds

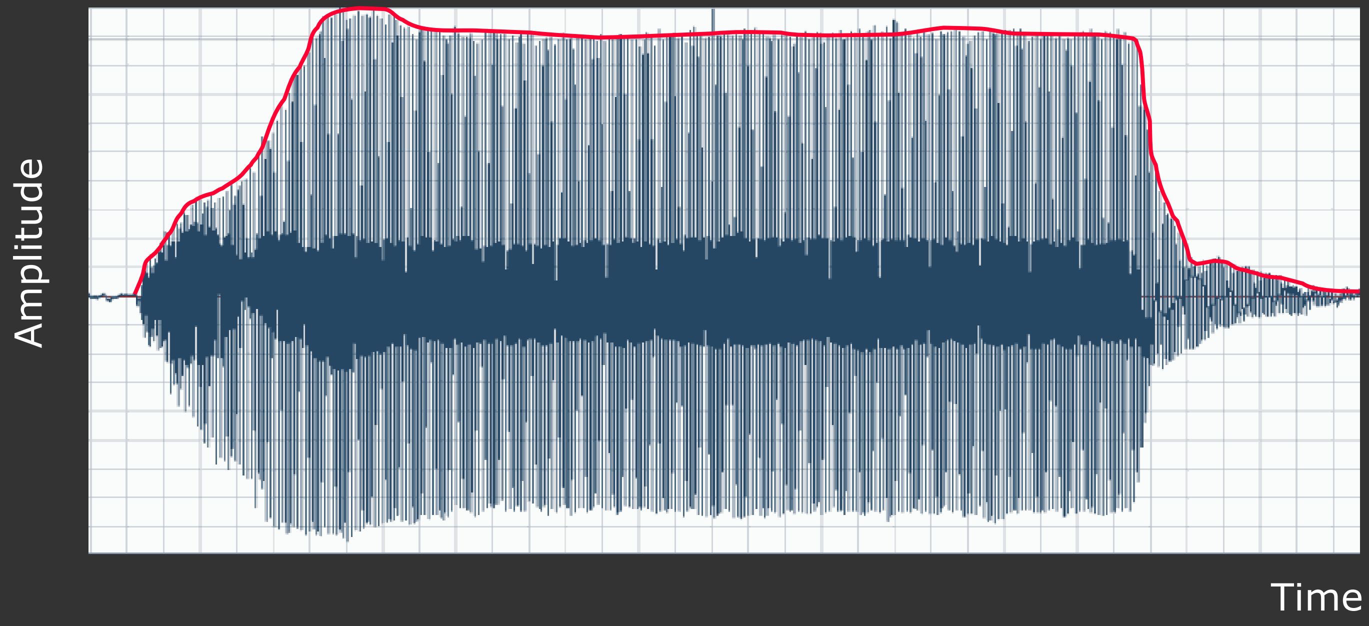
- Melodic sounds:
  - have a defined pitch,
  - they have a harmonic structure,
  - instruments: strings, winds, a few percussives
- Rhythmic (non-melodic) sounds:
  - pitch is undefined,
  - they have noise-like character,
  - most percussive instruments (e.g. a drum kit).

Synthetic musical sounds

- The aim of sound synthesis is to create a signal that has properties of a musical sound.
- That does not mean that we have to recreate sounds of existing instruments.
- We can create synthetic signals that sound the way we like, but they (usually) need to have features of a musical sound.
- The most important factors we need to control:
  - amplitude envelope,
  - spectral structure,
  - variability in time.

Temporal analysis and envelope

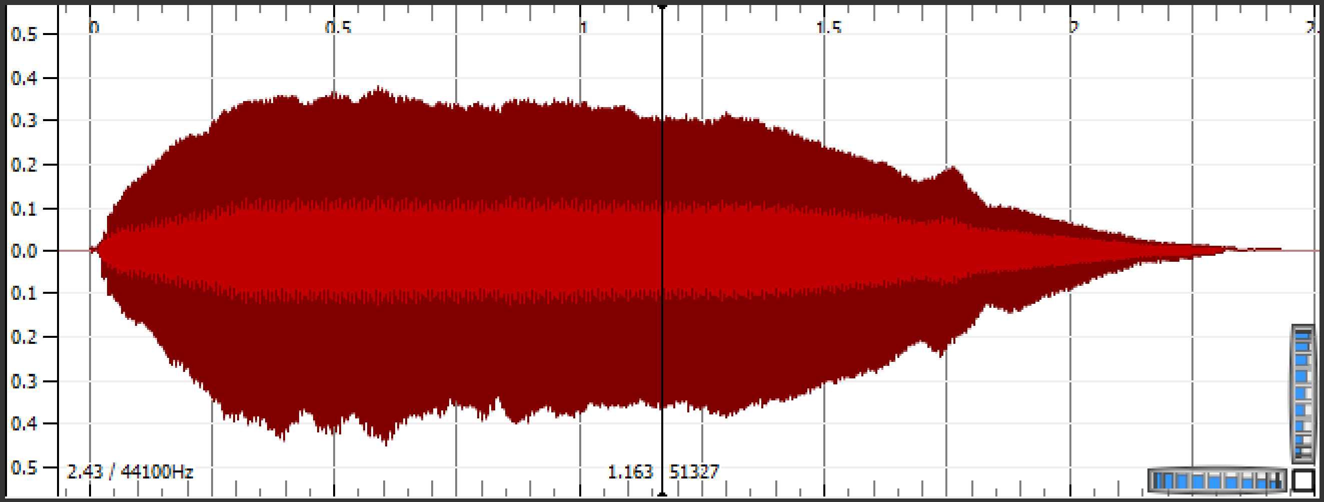
- Temporal analysis shows how the amplitude changes in time.
- We can also determine sound duration.
- Envelope follows the “edge” of the time plot.
- Envelope represents loudness variations in sound.



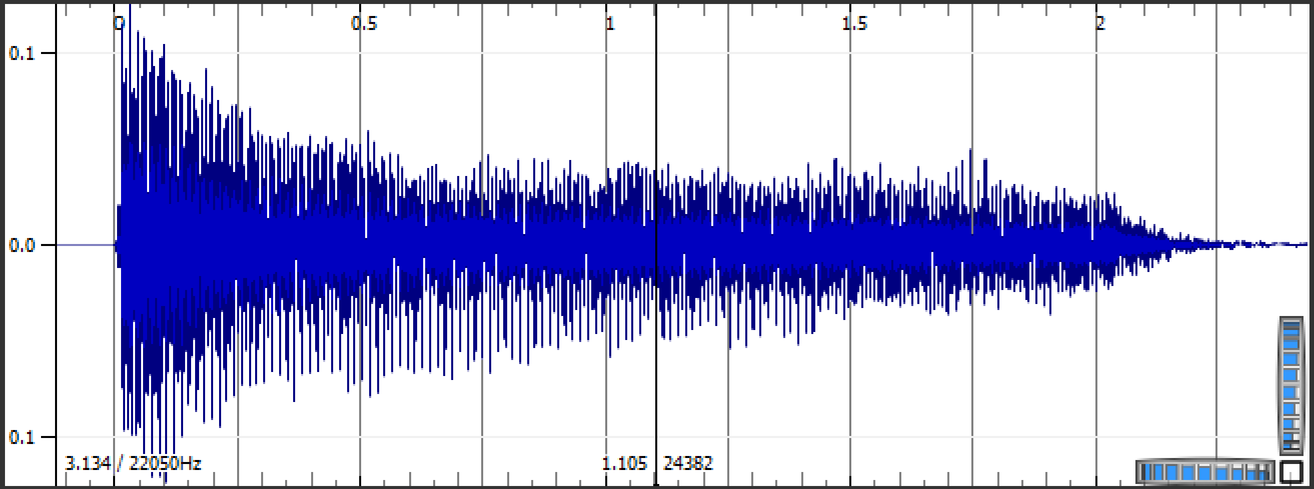
Phases of the sound envelope

- Attack (A) + decay (D)
  - the sound builds up, its envelope rises up
  - the initial transient (unsteady state)
  - large changes in timbre – attack defines the sound
- Sustain (S)
  - steady state, the sound continues to play, but its loudness and timbre may change (e.g. vibrato).
  - some instruments do not have the sustain phase.
- Release (R)
  - the sound naturally fades out.

Examples of sound envelope



Trumpet



Piano

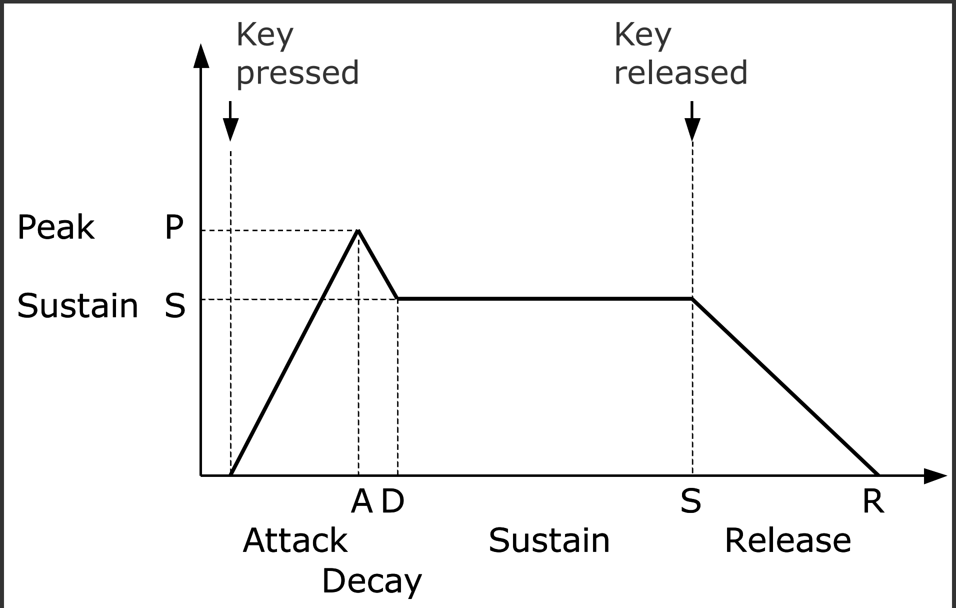
Envelope and the sound

- The shape of envelope depends on the instrument.
- The envelope also changes due to articulation
  - the way a musician plays on the instrument.
  For example, picking the guitar string harder shortens the attack and lengthens the release phase.
- How do we create the envelope in a synthesizer?
- Envelope generator creates a control signal
  - an ADSR envelope.
- This signal is used to control the gain of the output amplifier, so that the loudness changes according to the envelope shape.

ADSR envelope

Parameters of the classic ADSR envelope:

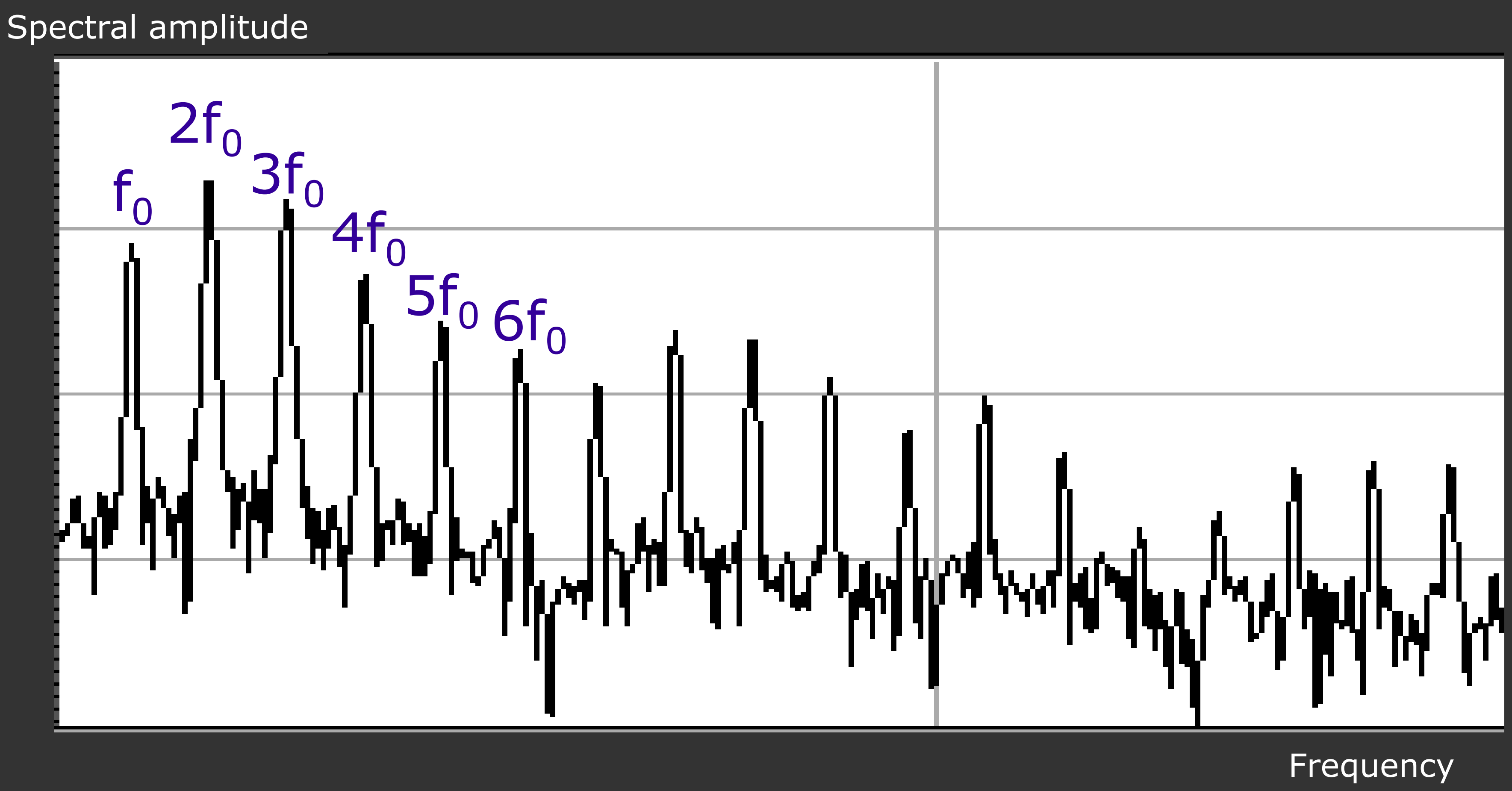
- A: attack phase duration
- D: decay phase duration
- S: sustain phase level (not duration!)
- R: release phase duration



Spectral analysis

- Determining sound properties in the frequency domain.
- Spectrum of the musical sound defines its pitch and timbre.
- Fourier analysis: any periodic signal can be decomposed into a sum of harmonic (sine) tones (partials) with different amplitudes and frequencies.
- How to compute a spectrum:
  - cut a part of the sound (preferably: a period) with a window,
  - compute the Fourier transform (FFT),
  - the result: spectral amplitude vs frequency.

Spectrum of a typical musical sound



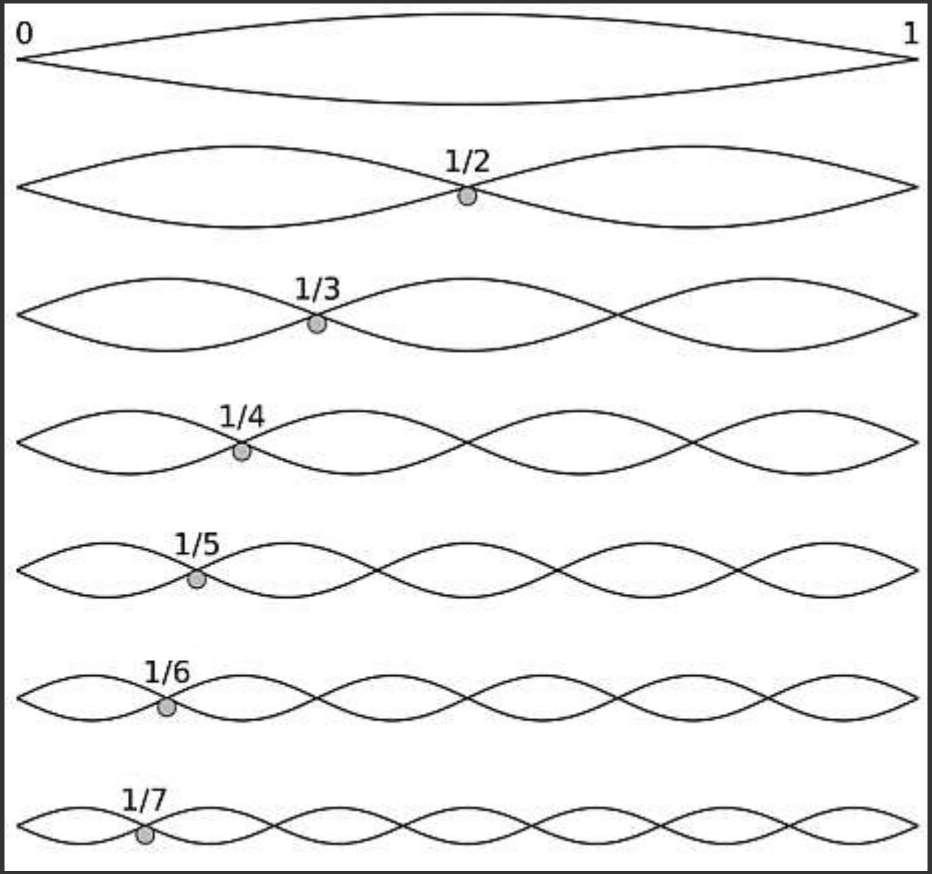
Spectrum of a musical sound

What do we observe in the plot?

- The spectrum has peaks – strong maxima.
- Peaks are placed in equal distances from each other, they form harmonic series.
- The first peak is found at the fundamental frequency of the sound (f_0).
- Higher peaks are harmonics: first ($2f_0$), second ($3f_0$), third ($4f_0$), and so on.
- A spectrum can also contain small non-harmonic peaks and a noise floor.

Spectrum of a musical sound

Why does the spectrum look like this? Because the sound is composed from standing waves, with different amplitudes.



Spectrum of a musical sound

REMEMBER THIS!!!

- **The fundamental frequency of a sound defines its pitch.**
- **Structure of all spectral components defines the timbre of the sound.**

Spectrum of a musical sound

It’s not that easy in practice.

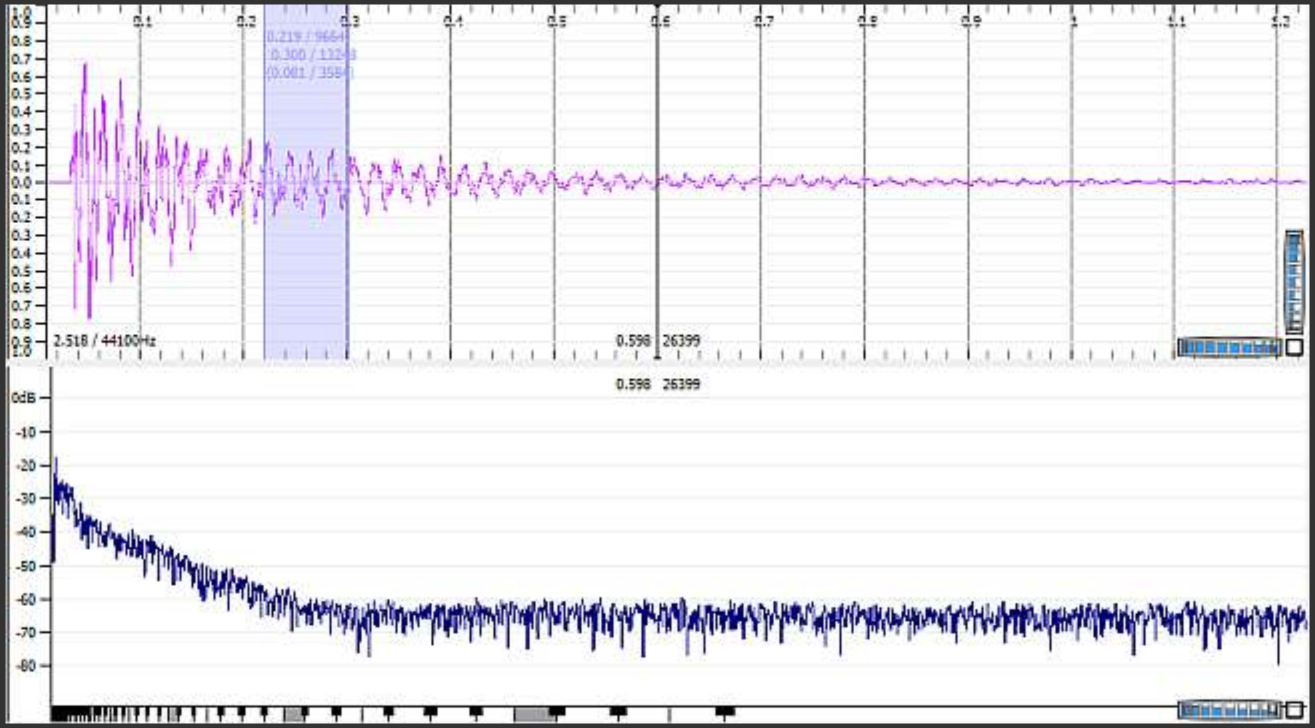
- Sometimes the first spectral peak is not the fundamental. It has to be the first peak in the harmonic series.
- In some sounds, even peaks (f1, f3, f5, …) are missing from the spectrum – but it’s still a harmonic sound.
- A few bell-like instruments produce inharmonic sounds.
- Most percussive sounds have a noise-like spectrum; there are no peaks, so no fundamental frequency and therefore, no pitch.

Percussive sounds

- Most percussive sounds have undefined pitch, they cannot be positioned on a musical scale.
- They usually are a band-limited noise with a specific envelope (very short attack, no sustain, long release).
- Depending on the spectral structure, we can say that the sound is higher or lower.
- But we cannot define a pitch (e.g. A1), because there is no fundamental frequency.

Percussive sounds

Synthetic percussive sounds are created by filtering a wideband noise and adding an envelope (very easy to do).



Describing a sound

- Low / high:
  - describes a pitch – position on a musical scale,
  - does not depend on a timbre.
- Dark / bright:
  - describes a timbre
  - higher bandwidth (larger number of partials) means that the sound is brighter
  - does not depend on the pitch.
- Quiet / loud:
  - depends only on the amplitude.

Additive synthesis

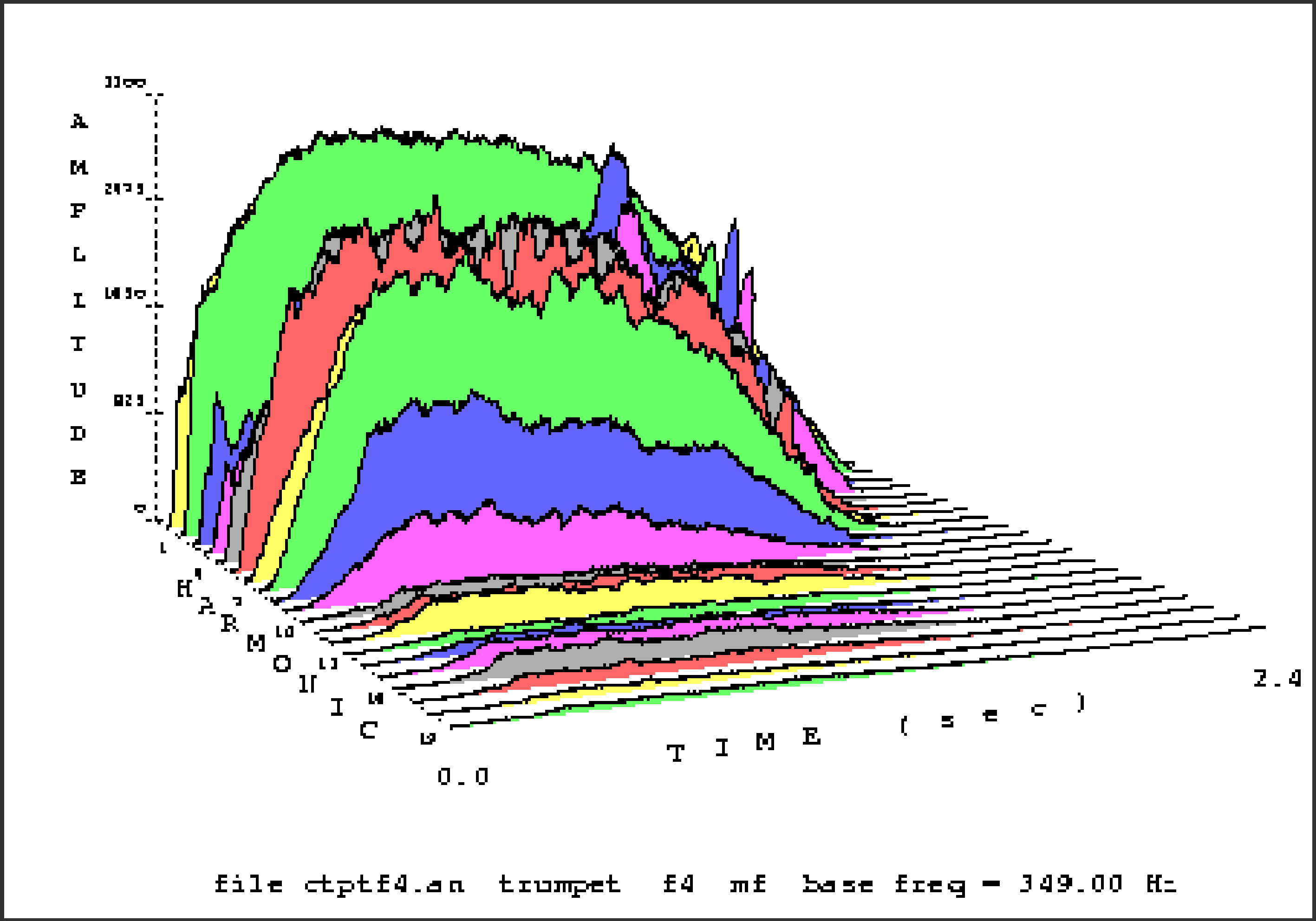
- We already know that a harmonic sound may be decomposed into partials.
- This process can be reversed: we can sum partials:
  - with frequencies in a harmonic order defined by a fundamental,
  - with amplitudes selected in order to obtain the desired spectral shape.
- This is the additive sound synthesis (from Latin: *additio*).
- This method was rarely used in commercial synthesizers.

Spectral changes in musical sounds

- The example of a spectrum presented earlier was captured at an arbitral moment.
- If the spectrum remains the same for the whole sound duration, the sound is dull, dead, uninteresting.
- The spectrum (and hence the timbre) of musical instruments sounds is variable, dynamic.
- Articulation (the way a musician plays the instrument) has a very large impact on the timbre changes, especially in the attack phase.
- In order to get alive, interesting sounds, we need to introduce the timbre changes into the synthesis process.

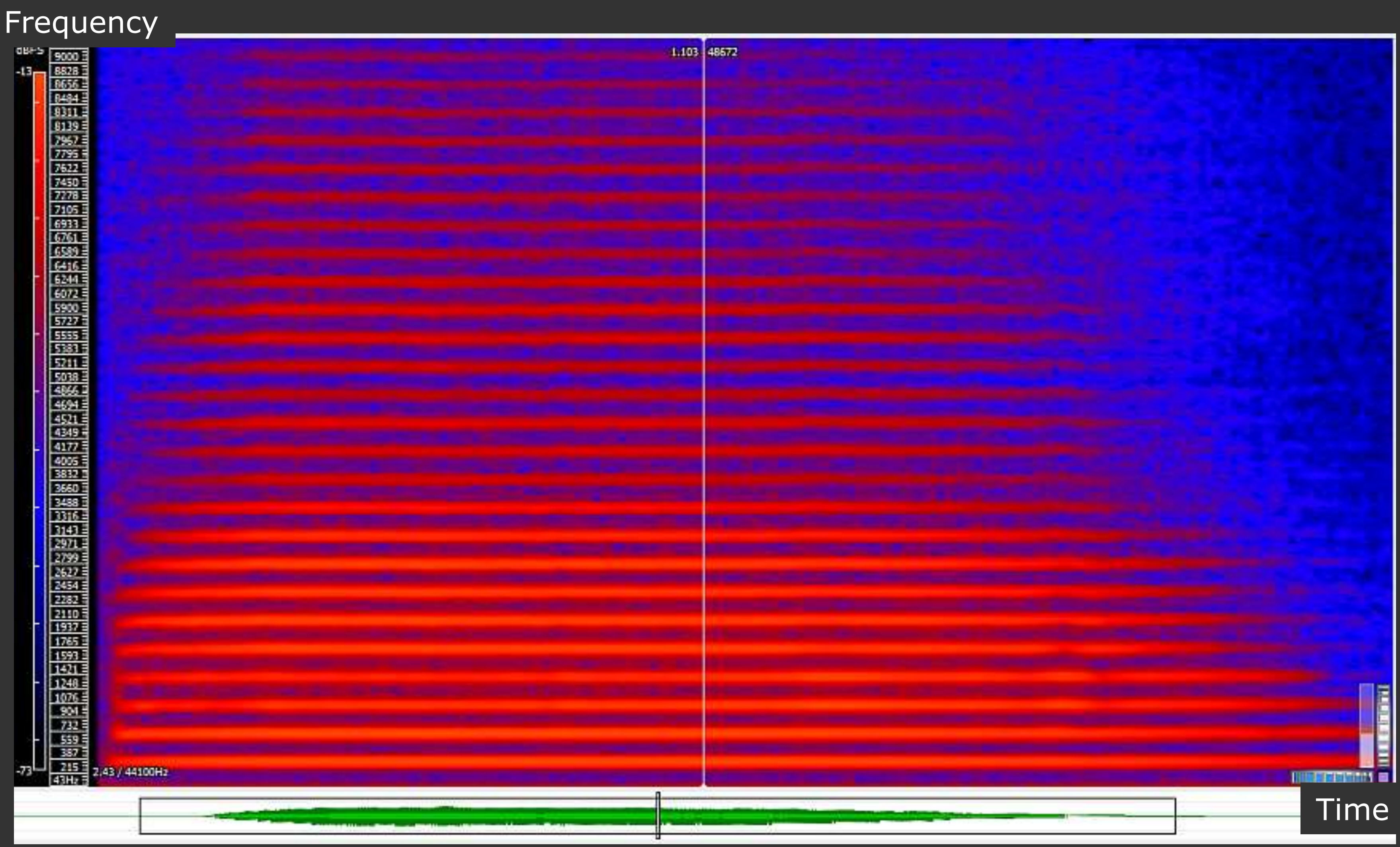
Waterfall spectrum plot

Observe how the partials change in time.



Spectrogram – a plot of spectral changes

3D plot: time vs. frequency vs. spectral amplitude (color)



Describing the sound, Part Two

- Alive, warm, dynamic sound:
  - – the sound changes as it plays,
  - – changes in timbre, pitch (e.g. vibrato), loudness,
  - – analogue oscillators were not perfect, but they introduced a desired variability in sounds.
- Dead, cold, static, „synthetic” sound:
  - – no changes as the sound plays,
  - – digital oscillators – the sound remains the same,
  - – dull, boring results,
  - – in order to make the synthetic sounds more alive, we need a modulation.

Sound parameters and its timbre

So, why two musical instruments produce sounds with different timbre, while their pitch is the same?

- Different envelopes.
- Different spectral structure.
- Different changes in spectrum, especially during the attack phase.

How to produce synthetic sounds with different timbre?

- Set the desired envelope (easy).
- Shape the static spectrum (easy).
- Ensure spectral changes in time, so that the sound is alive (this is much more difficult).

Spectral changes in the additive synthesis

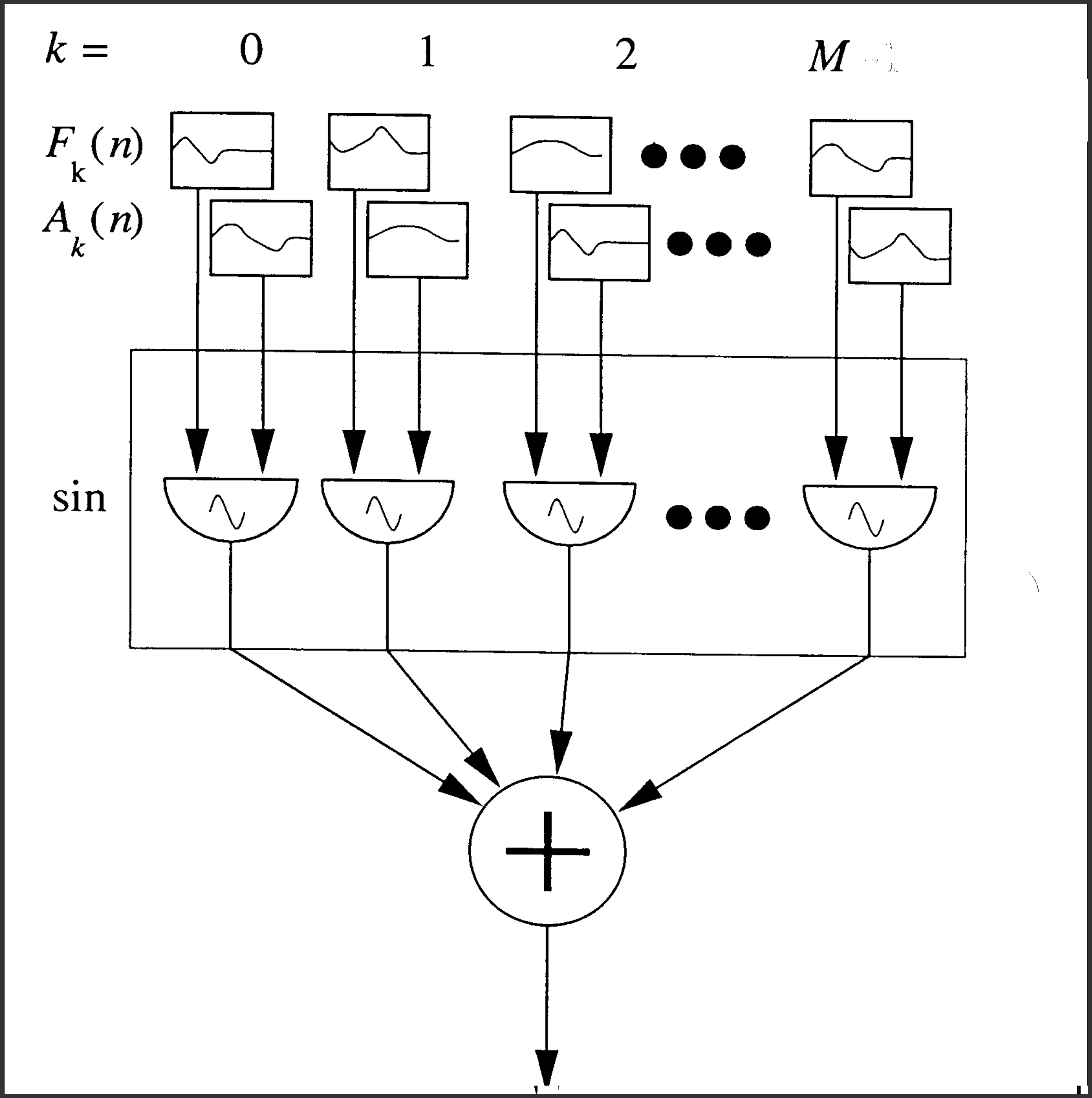
Back to the additive synthesis: how do we introduce dynamic spectral changes? The parameters can't be constant, they must be a function of time. We need to control:

- amplitude of each partial: $A_k(t)$
- frequency deviation (from the harmonic frequencies) of each partial: $\Delta f_k(t)$

If you *really* need a formula:

$$y(n) = \sum_{k=1}^{M} A_k(n) \sin\left(2\pi n\left(k \cdot f_0 + \Delta f_k(n)\right)\right)$$

A block diagram of additive synthesis



The easiest synthesis you can imagine.

Practical problem

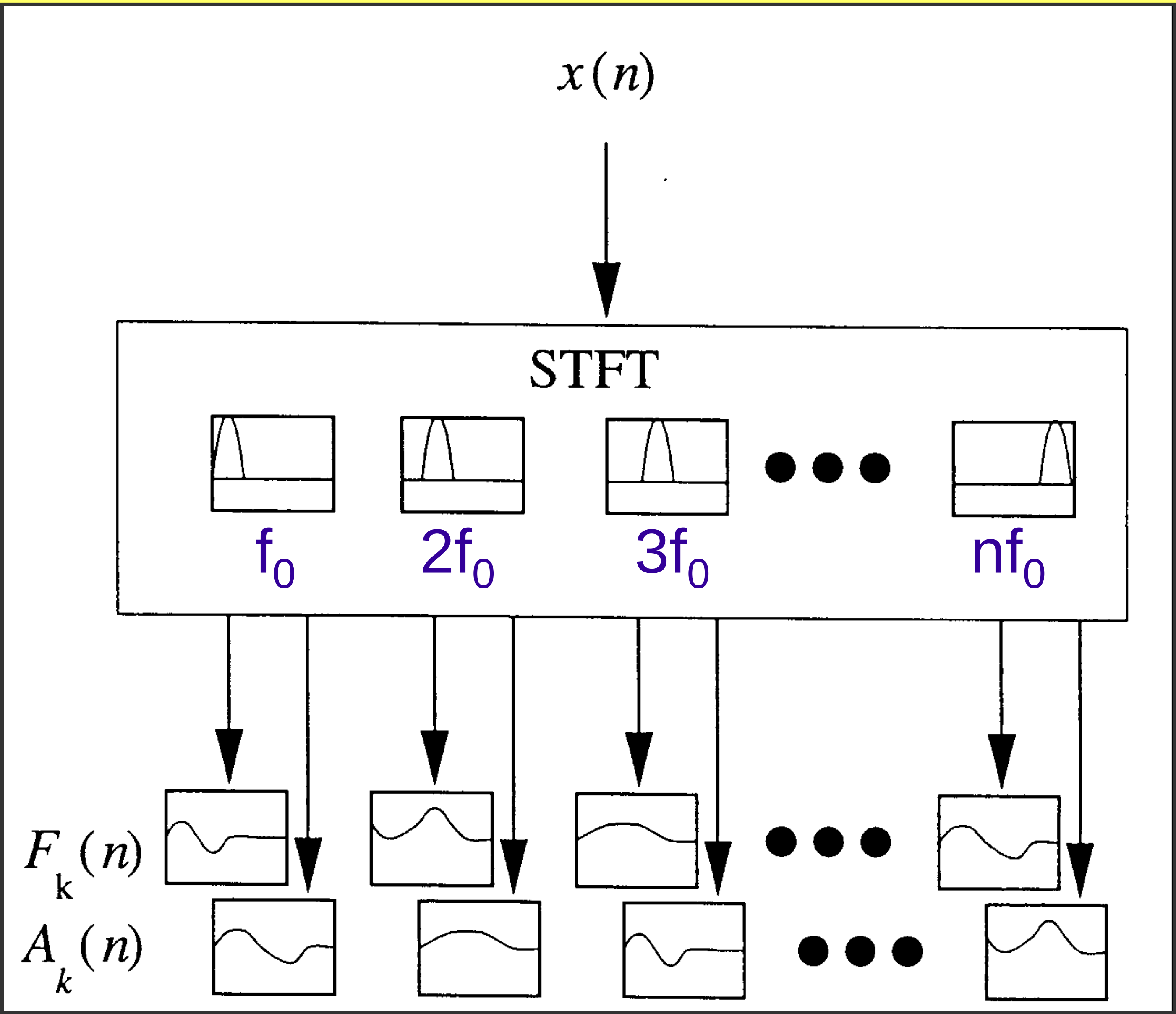
How do we create the control functions?

- They can be created "by hand". Fairlight tried this, it didn't work (too cumbersome).
- We can extract the parameters from the analysis of recorded musical sounds.
- We can then build a sound by the additive resynthesis.
- Instead of generating and summing the partials, we can use IFFT (inverse Fourier transform), it's easier.
- In practice, samplers do (almost) the same thing much easier.

PV analysis

- PV – phase vocoder analysis
- A bank of narrowband filters tuned to the harmonic frequencies.
- The filters measure energy in each band and produce the control functions.
- This method is inaccurate and coarse, the results are not very good sounding.
- We need to know the fundamental frequency.
- This method fails if the partials go outside their band, especially in the attack phase.

PV analysis



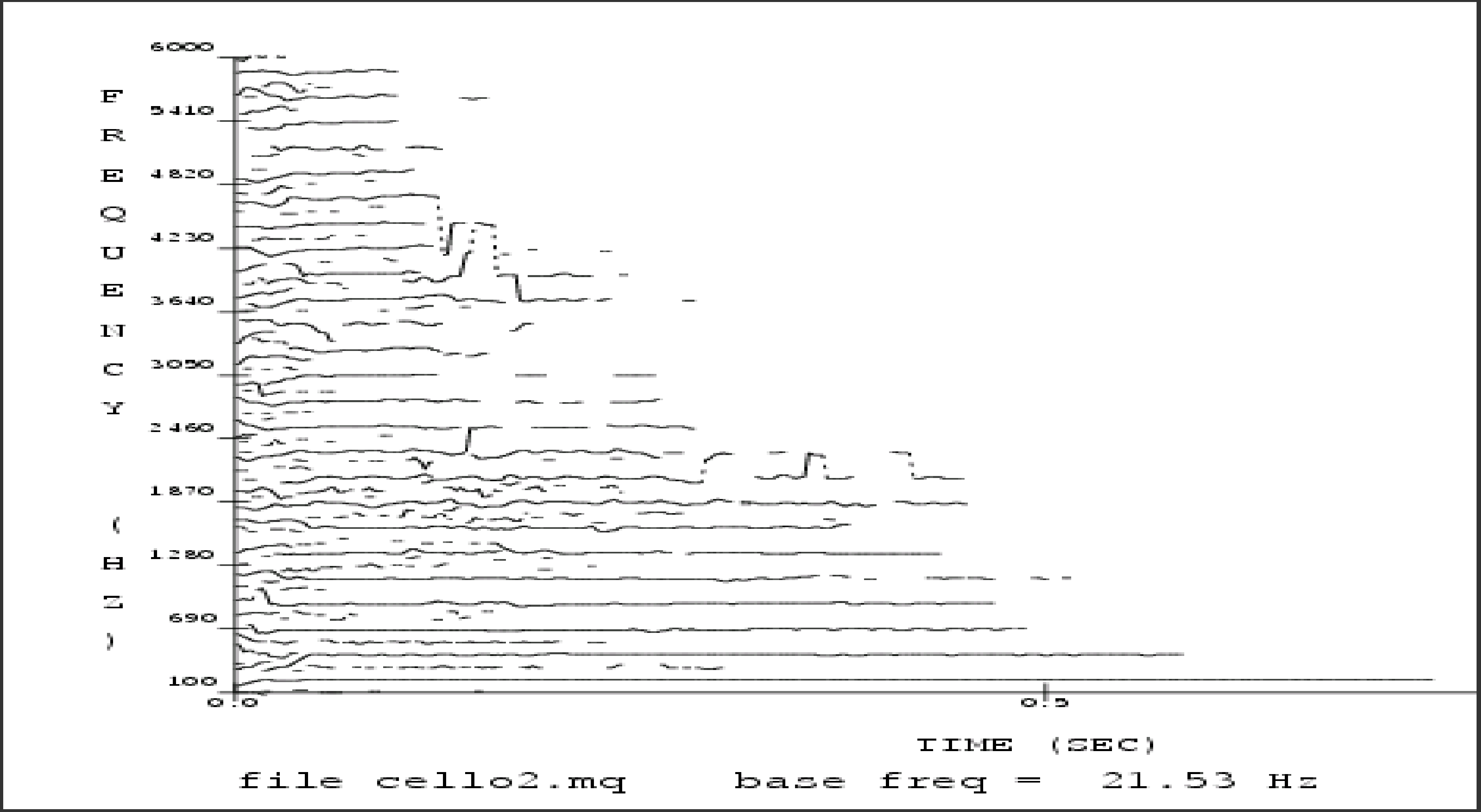
MQ analysis

MQ – an analysis method proposed by McAulay & Quatieri

- Digital FFT analysis in short signal frames.
- Local spectral maxima are found in each frame.
- Maxima that occur in the consecutive frames form spectral paths.
- Paths with sufficient duration and level are selected as the control functions.
- Sound resynthesis by IFFT.
- This method is much more accurate than PV.

MQ analysis

The analysis result – the extracted paths



Advantages of the additive synthesis

But, why would we decompose the sound and then rebuild it back? The answer: because we can modify the extracted parameters. For example, we can:

- transpose the sound – change its pitch, without changing its duration (sampling can't do that!),
- time stretch/compress the sound, without changing its pitch,
- remove unwanted spectral components and noise,
- modify the sound, add/remove partials, mix sounds, add effects, etc. (again: sampling can't do that).

Examples of additive synthesizers

Kurzweil K150 (1986)

- digital additive synthesis
- 240 oscillators
- each partial can be controlled with a computer

Kawai K5000 (1996)

- advanced sound workstation
- additive + sampling
- difficult to use

Multitone generators

- Multitone generators (also called additive generators) are used in some synthesizers.
- They generate tones in harmonic series and sum them in a defined proportion.
- This is not an additive synthesis: amplitudes are not individually controlled, the composed sound is further processed as a whole.
- Sometimes, two separate multitone generators for odd and even partials are used – more realistic effects.

Additive synthesis – a summary

Pros:

- it can recreate sounds of musical instruments,
- sound spectrum can be modified directly,
- it’s easy to do pitch shifting and time stretching,
- conceptually easy algorithm.

Cons:

- not for creating new sounds,
- complicated control of the synthesis process – each partial must be controlled separately,
- a sampler gives a similar result and it’s much easier to use.

Musical sound synthesis

To conclude the lecture: how do EMIs create sounds?

- We can compose the sound from the partials (additive synthesis – we have already learned that).
- We can combine and shape raw initial sounds (subtractive and wavetable synthesis)
- We can create the sound with a computer algorithm, by a modulation (FM synthesis)
- We can simply process and play back recorded sounds (sampling)
- We can build a computer model of an instrument that will create synthetic sounds (waveguide synthesis).

Bibliography

- SPEAR - Sinusoidal Partial Editing Analysis and Resynthesis: http://www.klingbeil.com/spear/
- M.K. Klingbeil: *Spectral Analysis, Editing and Resynthesis: Methods and Applications*. Columbia Univ. 2009 (dostępne ze strony SPEAR)
- R.J. McAulay, T.F. Quatieri: *Speech Analysis/Synthesis Based on A Sinusoidal Representation*. IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. ASSP-34, no. 4, Aug. 1986, pp. 744-754.
- M. Russ: Sound Synthesis and Sampling. Focal Press, Oxford 1996.
- Vintage Synth Explorer: www.vintagesynth.com
- Wikipedia