MULTIBAND PERCEPTUAL MODULATION ANALYSIS, PROCESSING AND SYNTHESIS OF AUDIO SIGNALS

Sascha Disch

Laboratorium für Informationstechnologie
Leibniz Universität Hannover

Overview

- Motivation
- Multiband modulation analysis
  - Principle
  - System
  - Realization
- Multiband modulation synthesis
- Exemplary application:
  - Music transposition (or time stretching)
  - Key mode conversion
- Summary

Motivation

*From a **perceptual** view, audio signals are composed of low bandwidth and low frequency sub-processes, which **modulate** much higher carrier frequencies.*

- Appropriate multiband modulation decomposition?
- Goal: direct access to perceptually relevant parameters (pitch, roughness, timbre)
- Ill-defined problem: Many possible decompositions
- Properties of a suitable decomposition:
  1. Carriers provide a ‚sketch' representation => **signal adaptivity**
  2. Signal modification at minimal artifacts => **perceptual criteria**
  3. No signal modification => **transparent audio quality**

Modulation analysis: principle

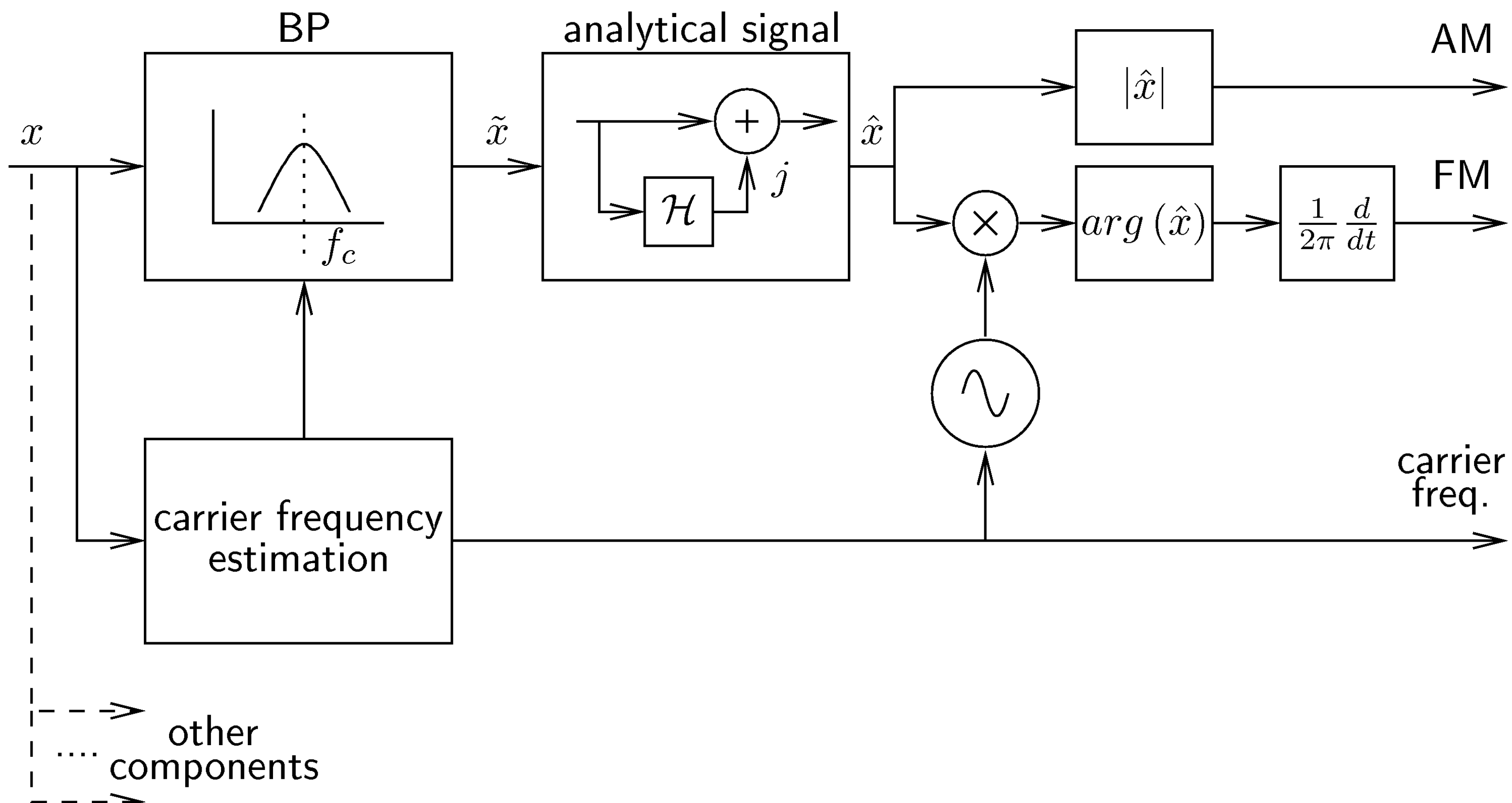
=> Task of spectral bandpass (BP) segmentation:

1. Signal adaptive **alignment** of sub-bands
2. Sub-bands match a **perceptual scale**
3. Sum of all sub-bands provides **seamless coverage** of spectrum

- Sub-bands are aligned to <u>local</u> **centers of gravity (COG)**
  - Local accumulation of energy in spectrum
  - Associated sound fused by human auditory system
- Further decomposition of sub-bands (components) into
  - Carrier frequency
  - Amplitude modulation (AM)
  - Frequency modulation (FM)

Modulation analysis: system

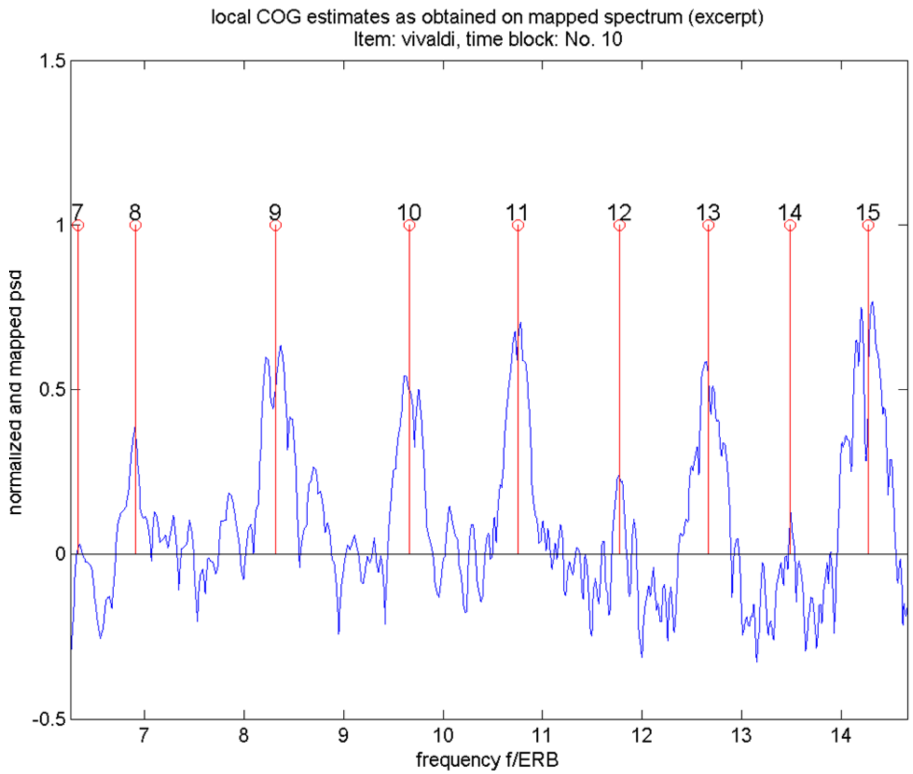
- Analysis: analytical sub-band signals
  - AM: absolute value (envelope)
  - FM: instantaneous frequency (IF), derivative of heterodyned phase

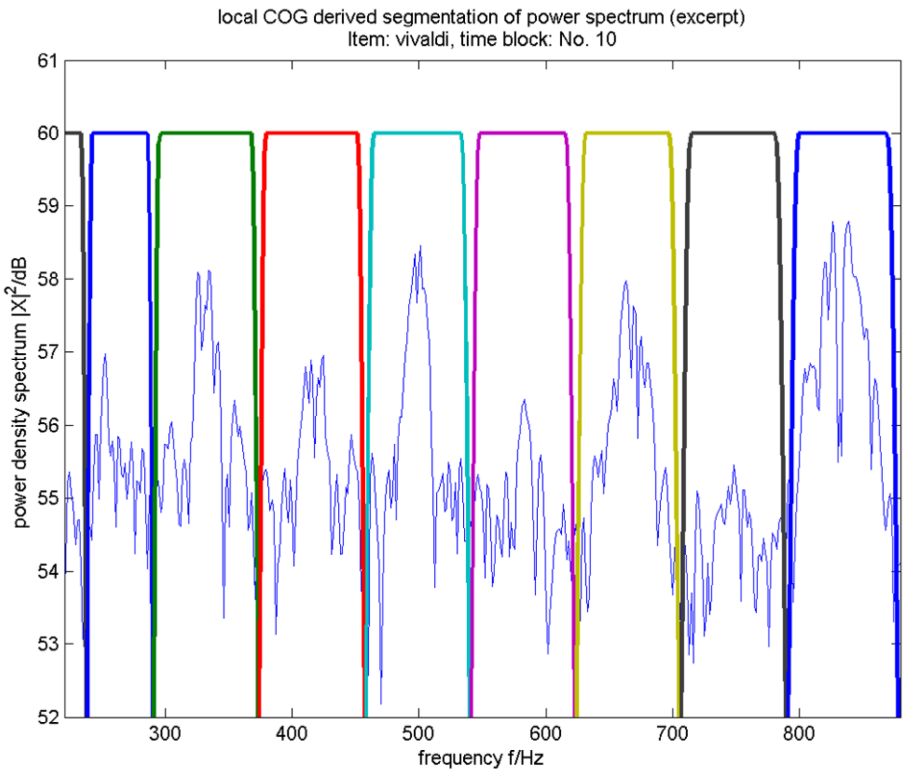


Modulation analysis: realization (1/2)

- Time-blockwise processing
- Initial spectra obtained via DFT
- Efficient spectral segmentation
  - Mapping of spectrum onto <u>perceptual ERB scale</u>
  - COG computation using a <u>constant</u> mean bandwidth
- Bandpass filter implementation (BP)
  - Spectral weighting
  - Inverse mapping of BP weights onto linear scale
- Analytic sub-band signals
  - Application of BP weights
  - Single-sided IDFT (positive frequencies only)

Modulation analysis: realization (2/2)

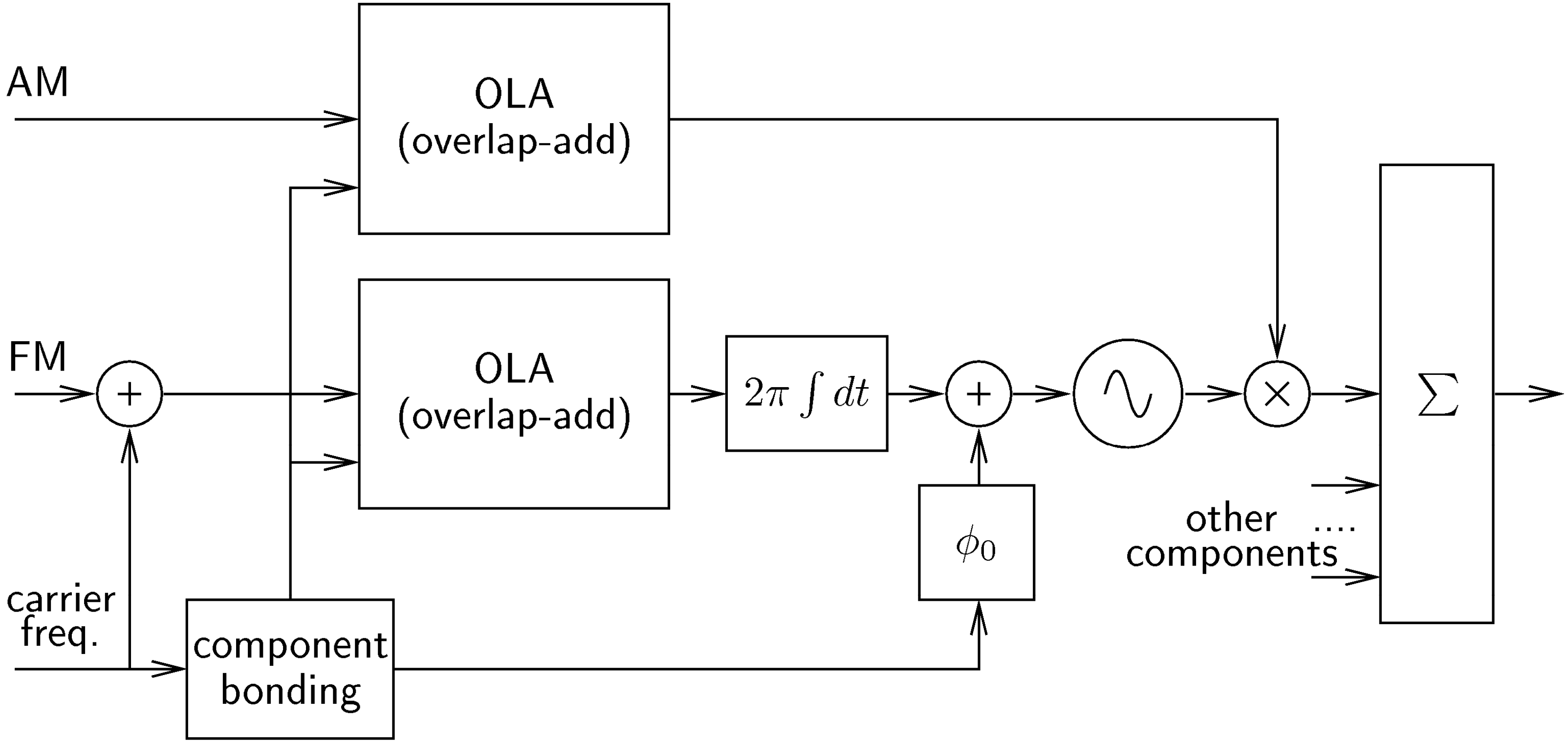




- Parameters
  - Block length ~340 ms, temporal stride ~85 ms, sub-band spectral width 0.5-1 ERB

Modulation synthesis: system

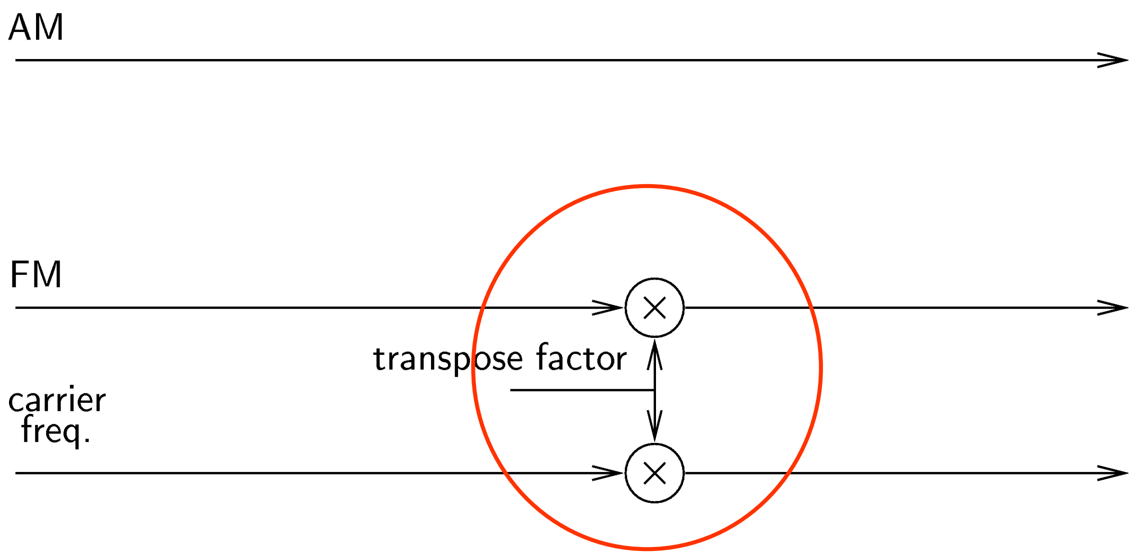
- Synthesis
  - Bonding: pairing of components between subsequent time blocks according to least spectral distance
  - Pairwise overlap-add (OLA) in modulation decomposition domain



Application: music transposition

- Changed pitch at unchanged playback speed
- Stretch/squeeze of <u>global</u> spectrum
  - Scaling of <u>all</u> carrier frequencies and FM by constant factor
  - Temporal structure of signal is maintained in the (unchanged) AM

	Original
Bocch erini	



Play all

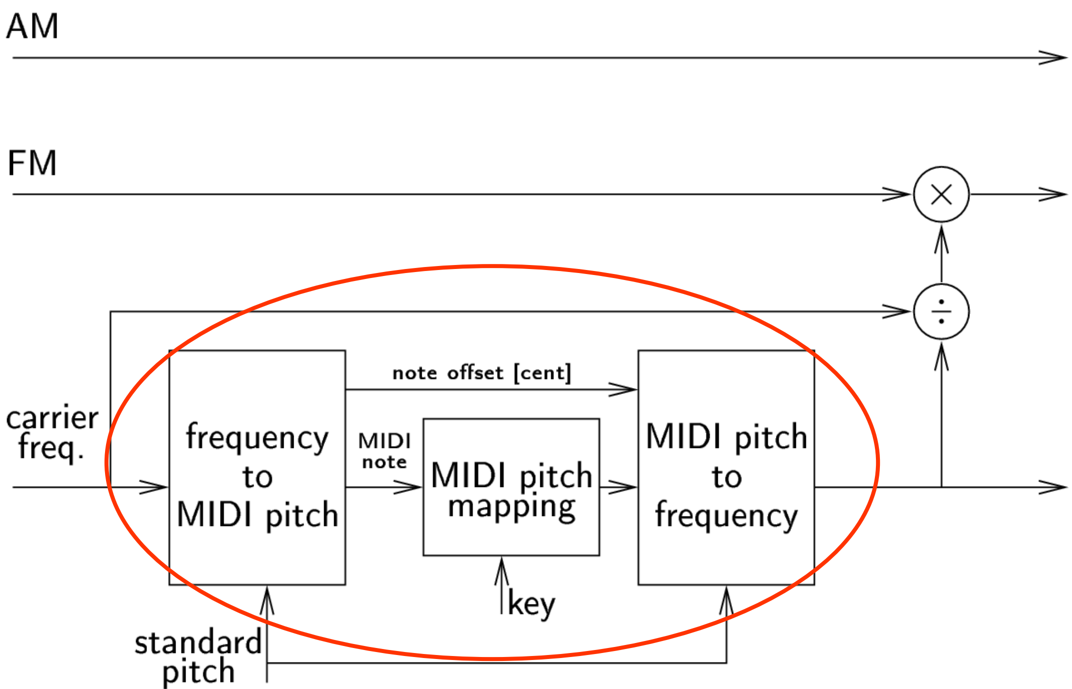
third down	third up

Half speed

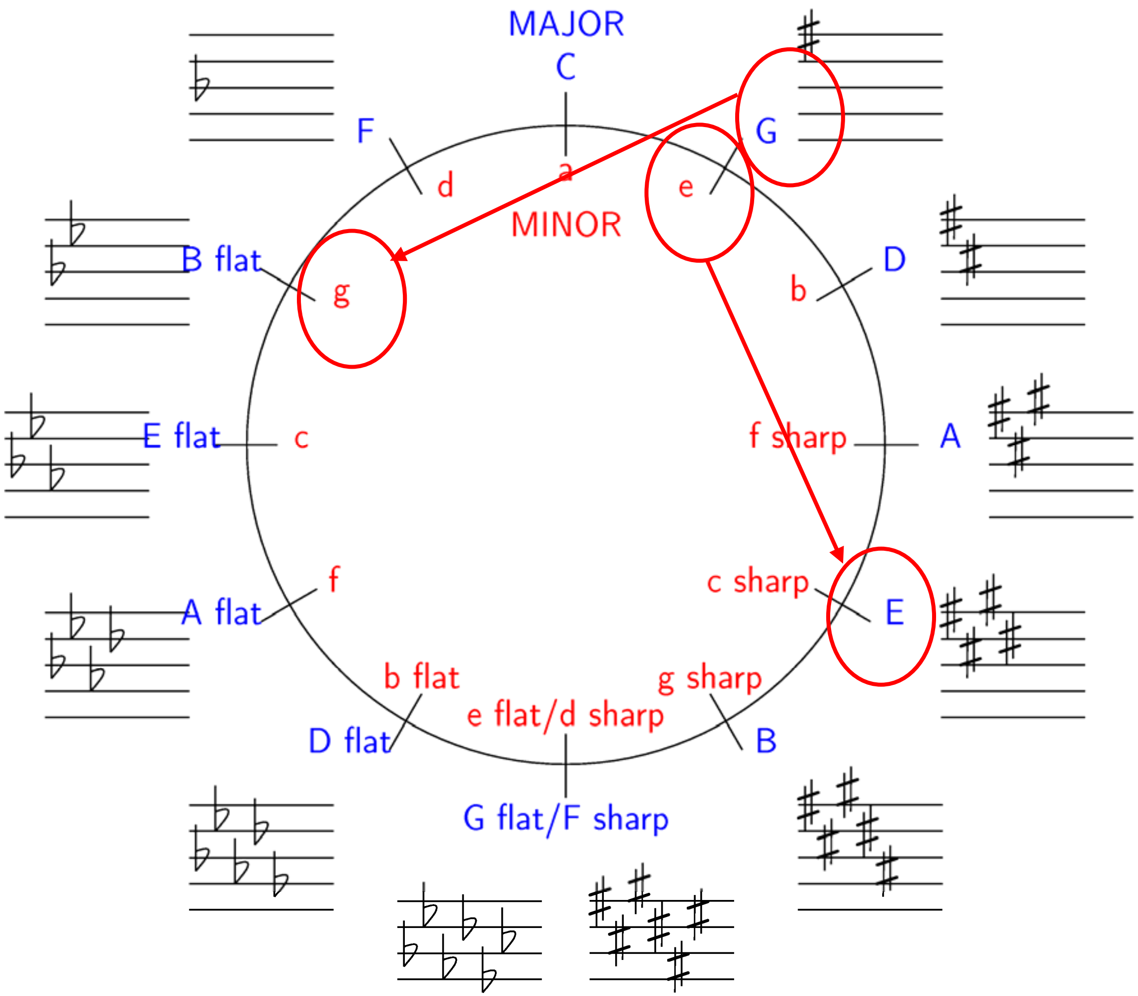
- Changed speed at unchanged pitch by decimation

Application: change of key mode (1/3)

- Key conversion of polyphonic music, e.g. major <-> minor
- Spectral mapping of selected carrier frequencies
  - Mapping of carriers onto MIDI-pitch representation
  - **A-priori** information: standard pitch (e.g. 440 Hz) and key
  - Shift of selected pitch values according to change of scale



Application: change of key mode (2/3)



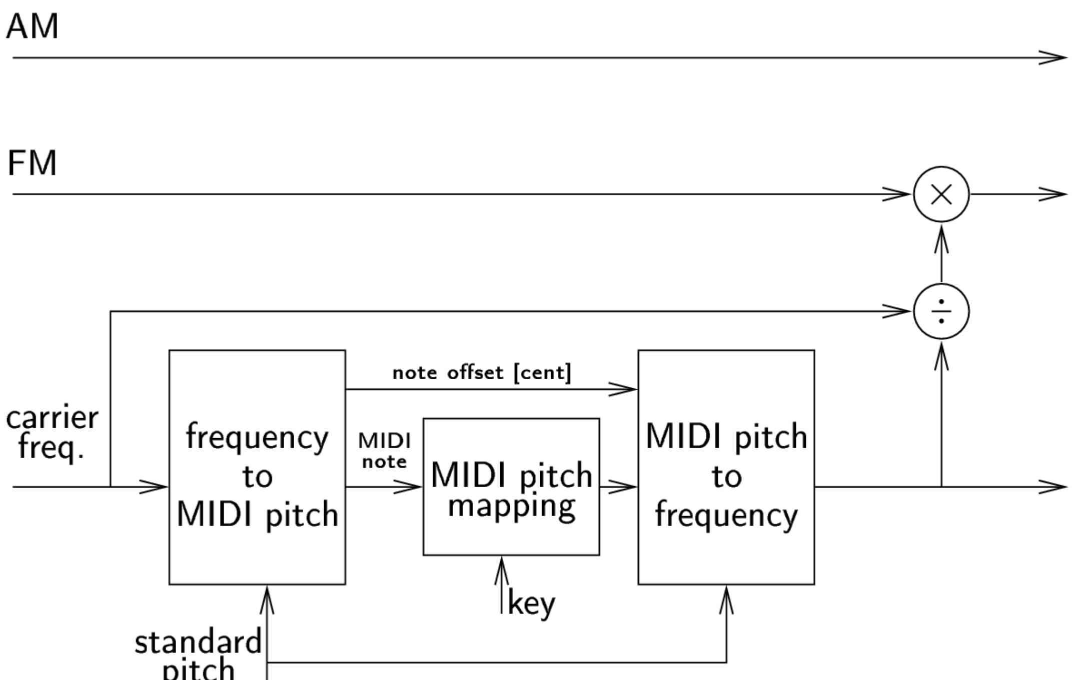
- Mapping of tones according to scale change
- Circle of fifth
- Rules (natural minor):
  - **Major to minor**: 3 steps counter clockwise
  - **Minor to major**: 3 steps clockwise
- 1 Step == mapping of one tone

Application: change of key mode (3/3)

- Key conversion of polyphonic music, e.g. major <-> minor
- Spectral mapping of <u>selected</u> carrier frequencies
  - Mapping of carriers onto MIDI-pitch representation
  - **A-priori** information: standard pitch (e.g. 440 Hz) and key
  - Shift of selected pitch values according to scale change

Play all

	Original
Vibra-phone	major
Vivaldi	major
Bach	minor



Processed
minor
minor
major

Summary

- A novel method for perceptual multiband modulation analysis, processing and synthesis of audio signals
- Suitable spectral decomposition of paramount importance for perceptual quality
- High quality synthesis
- Provides a direct, selective handle on pitch
- Application: Music transposition/key mode change
  - Audio post-processing
  - Music production
  - New exciting audio effects

End

questions/discussion