Time-Frequency Analysis for Audio Workshop

Matching the Analysis Scheme to the Signal

Fritz Menzer (fritz.menzer@epfl.ch)
Communication Systems, 5th year
Ecole Polytechnique Fédérale de Lausanne

15th April, 2004

Overview

1 Introduction **3**

2 Perfect Reconstruction - who cares? **4**

2.1 Definition of perfect reconstruction 4

2.2 Do we need perfect reconstruction? 5

3 Harmonic Band Wavelet Transform **7**

3.1 Coefficient modeling 10

3.2 Advantages / Drawbacks 11

4 From HBWT to inharmonic sound modeling **12**

4.1 Taking filters from different PR filterbanks 13

4.2 Why aliasing is not a problem 14

4.3 Method Overview 17

4.4 Sounds 18

5 Time-Frequency Analysis and Granular Synthesis **19**

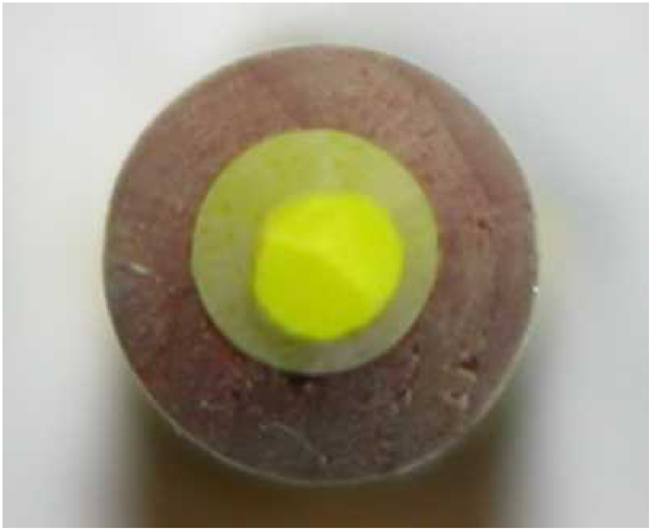
5.1 Time-domain effects 25

5.2 Scale of all grains in a 1024-band full-tree wavelet decomposition . . 26

A References **27**

1 Introduction

- If you know what you're looking at, you can examine it more precisely.



2 Perfect Reconstruction - who cares?

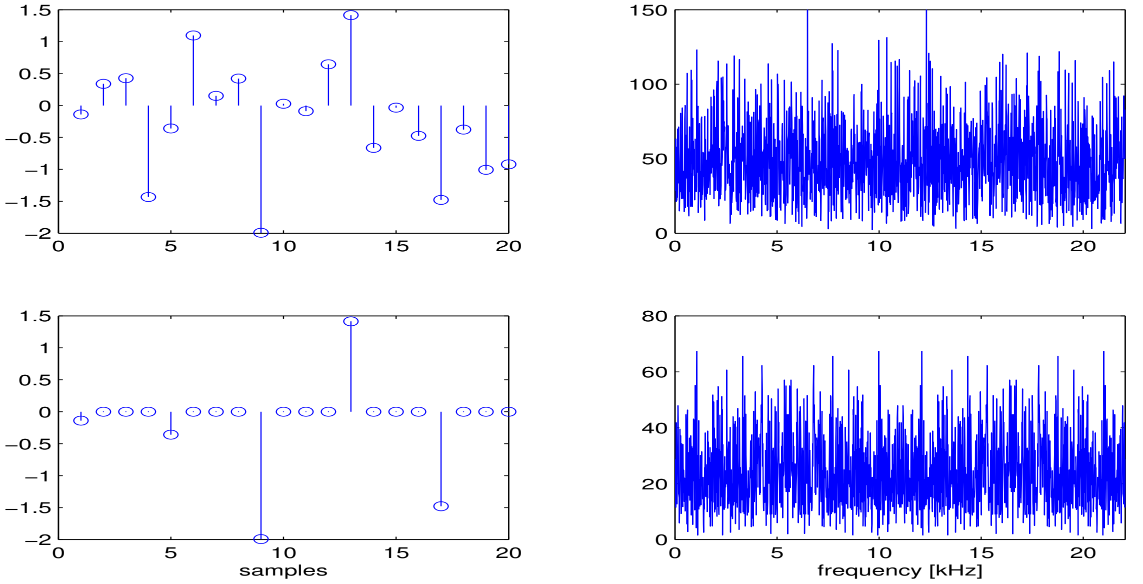
2.1 Definition of perfect reconstruction

- Definition: **Perfect Reconstruction (PR)** method: method providing direct and inverse transforms $\mathcal{T}$ and $\mathcal{T}^{-1}$ such that for any signal s,

$$\mathcal{T}^{-1}(\mathcal{T}(s)) = s$$

- FFT based methods, Cosine Modulated Filterbanks and Wavelet transforms are usually PR methods.
- Simple operations like filtering or distortion do not necessarily allow PR (i.e. it may be impossible to find $\mathcal{T}^{-1}$).
  Example: Quantisation obviously does not allow to reconstuct the original signal perfectly.

2.2 Do we need perfect reconstruction?



Noise

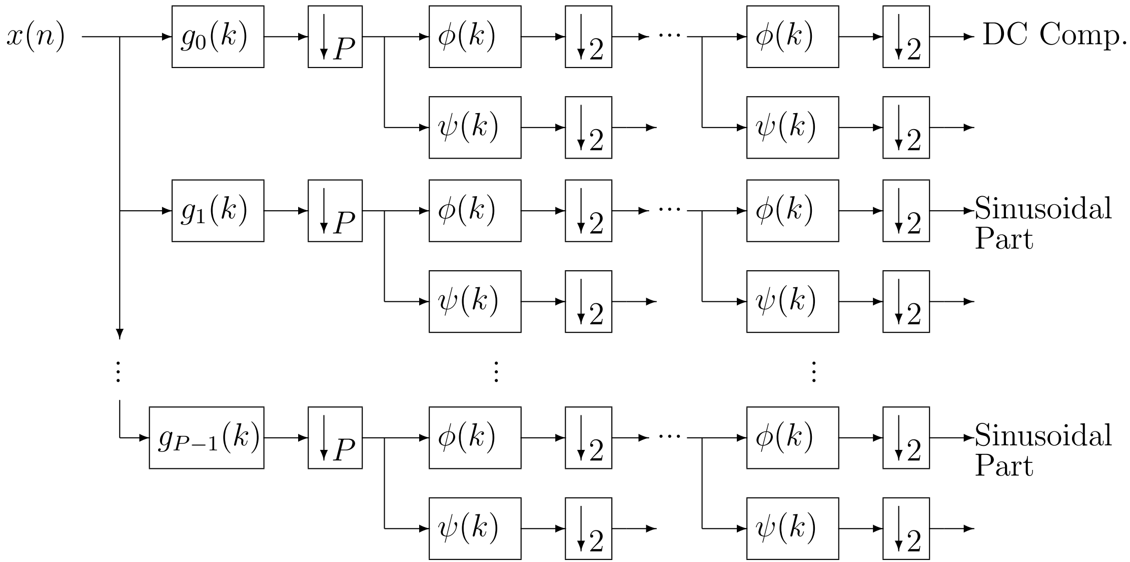
Noise, down- and upsampled by 4

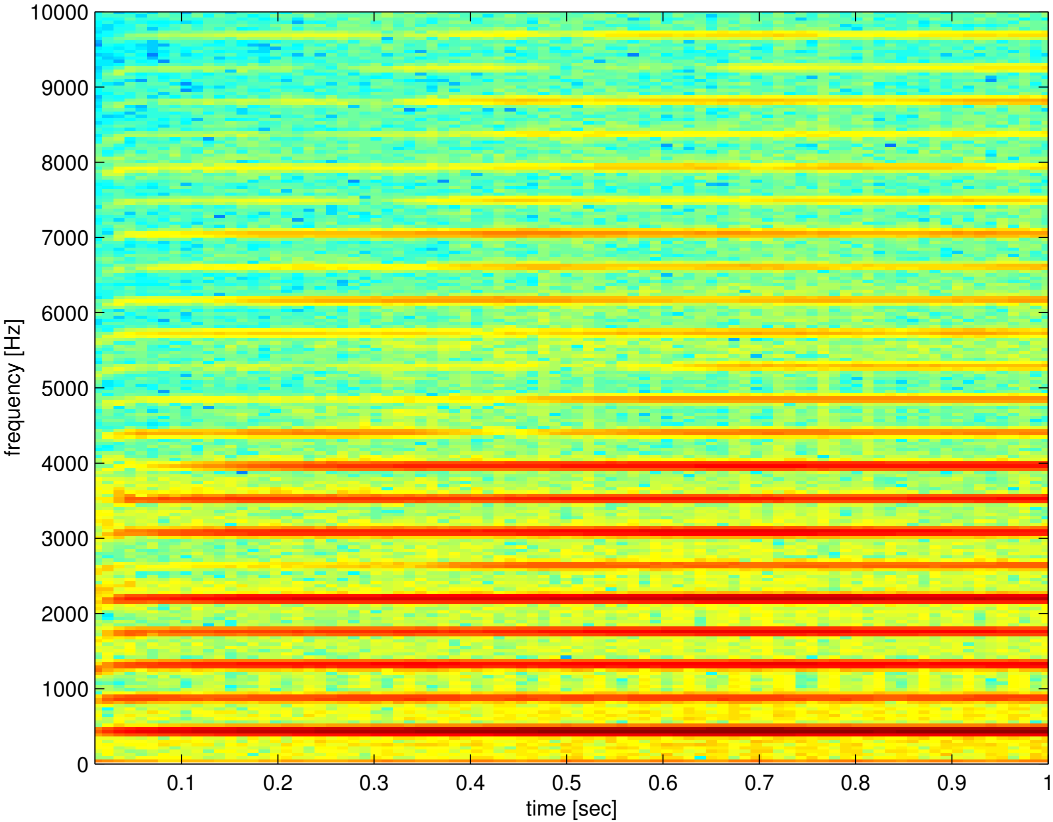
Do we need perfect reconstruction?

- Not needed for:
  - Modifying a signal
  - Handling noise
  - If the nature of the signal is known
- Why use PR methods for compression?
  - Generality (ideally any signal can be treated)
  - Localising the source of errors!

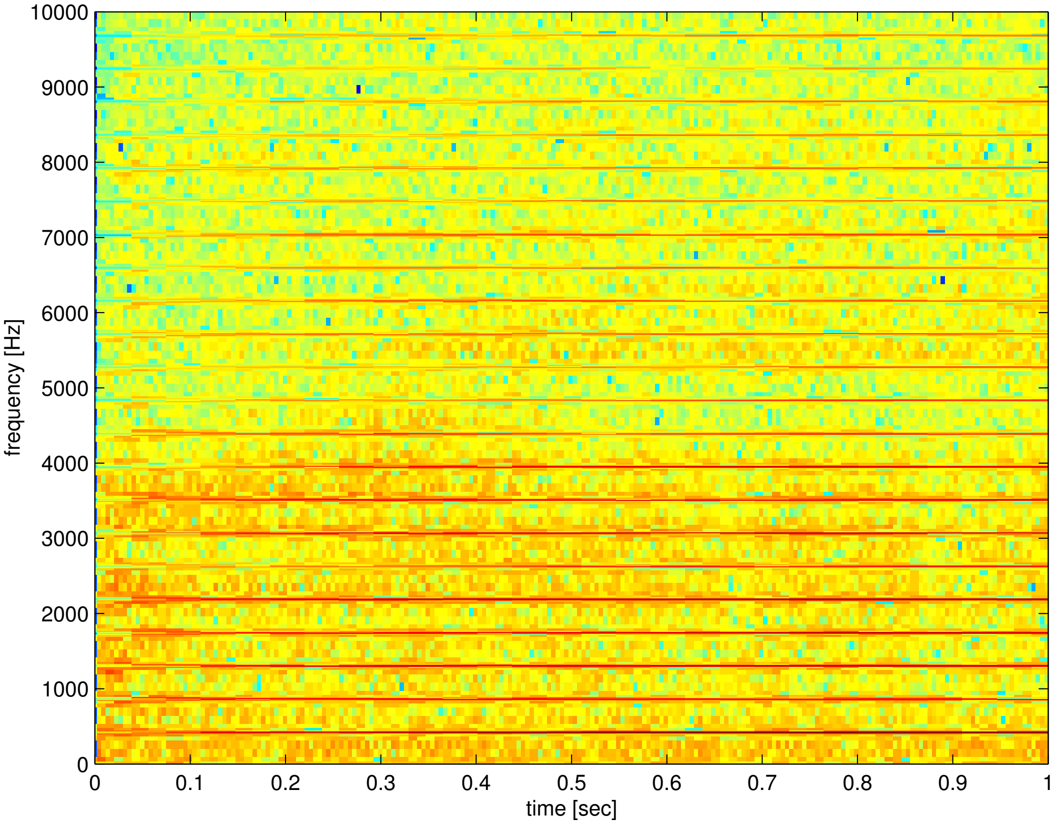
3 Harmonic Band Wavelet Transform

(Polotti and Evangelista, 2000)





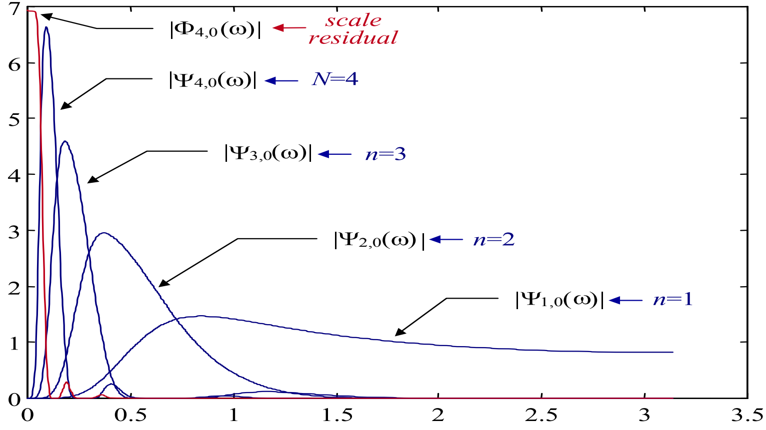
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
frequency [Hz]
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
time [sec]



10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
frequency [Hz]
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
time [sec]

3.1 Coefficient modeling

- Wavelet Transform
- Model the scale residual sinusoidally
- Model the wavelet coefficients using LPC



3.2 Advantages / Drawbacks

+ Meaningful adaptation of frequency and time resolution

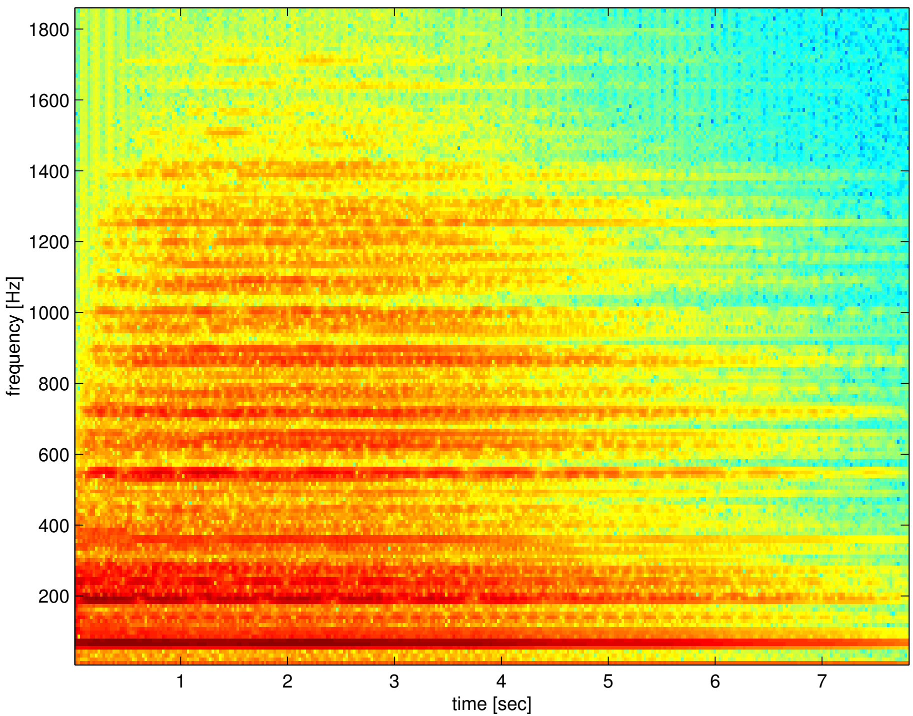
  $\Longrightarrow$ Visually better resolution

+ Reasonable model for the coefficients

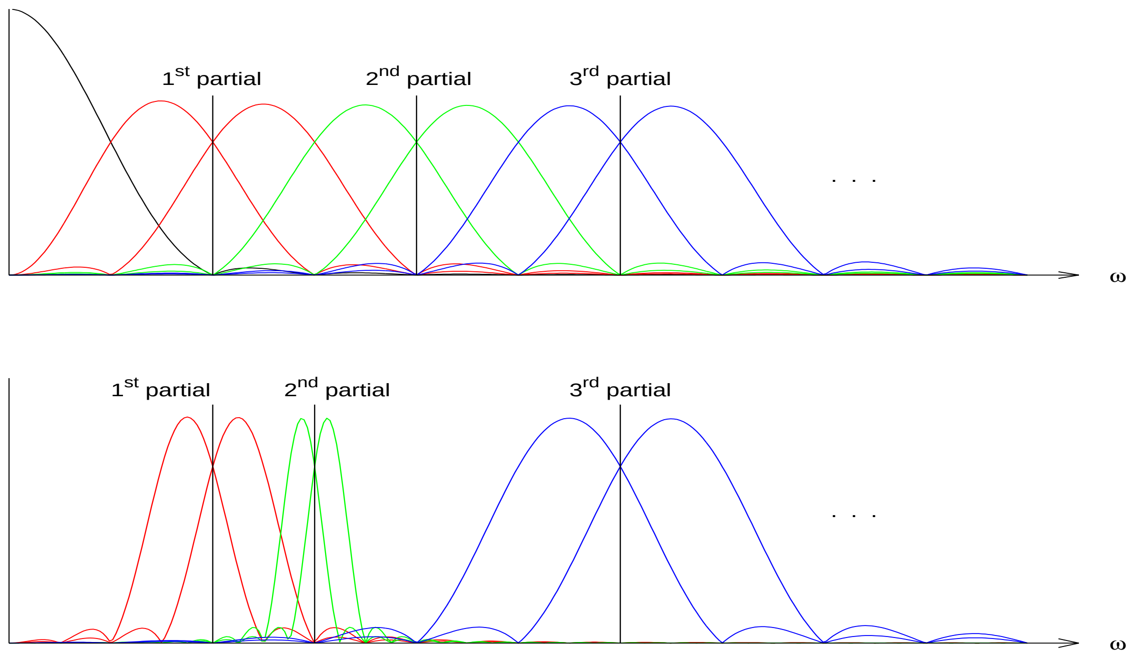
– Works only for monophonic, harmonic sounds

– No model for the transients

4 From HBWT to inharmonic sound modeling



4.1 Taking filters from different PR filterbanks



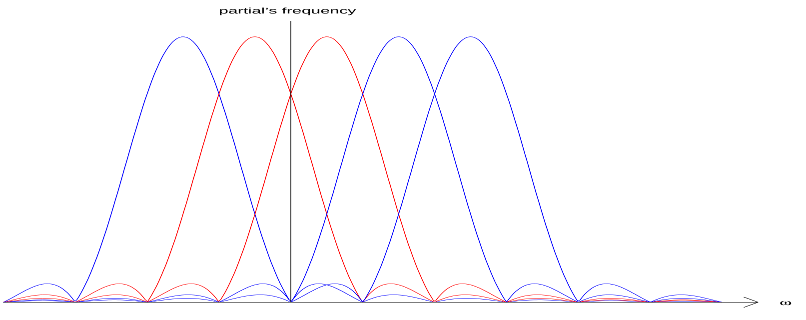
4.2 Why aliasing is not a problem

If a sinusoid of the form

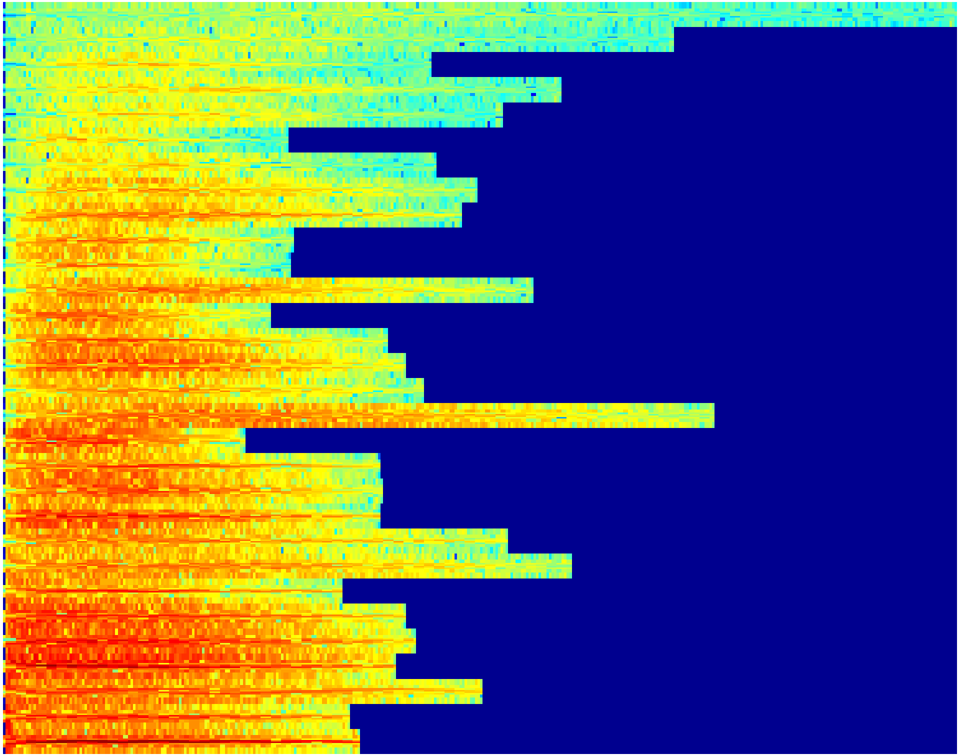
$$\sin\left(\frac{\hat{k}\pi}{P}t + \varphi\right)$$

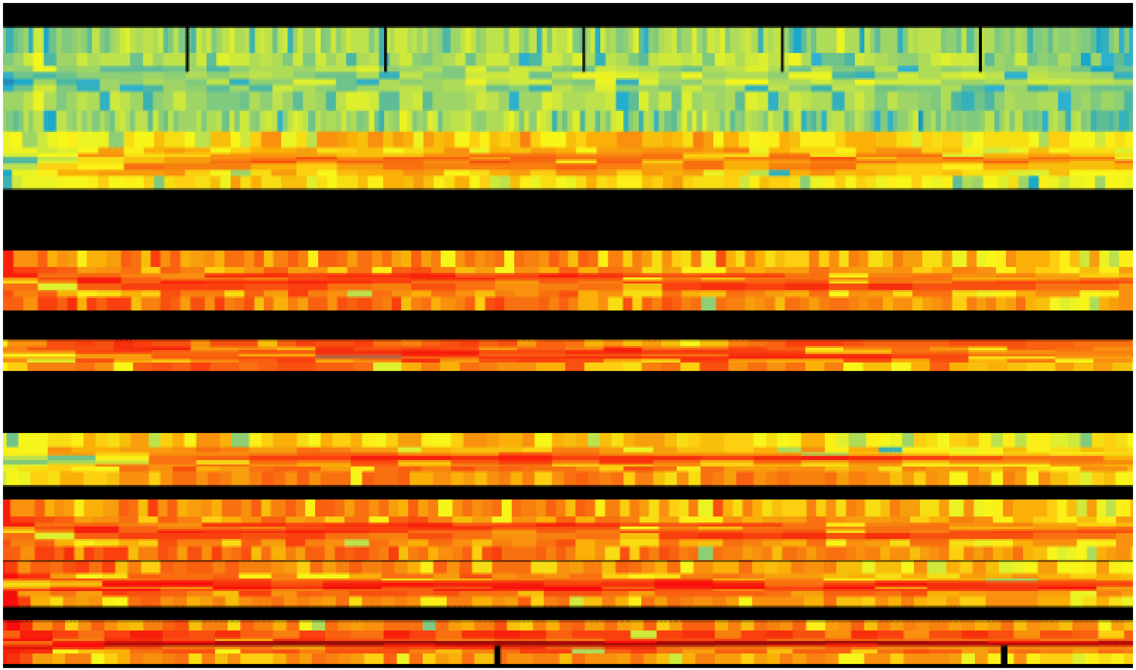
is the input to a P-channel cosine modulated filterbank, only two bands will output nonzero coefficients:

$$|H_k(e^{j\frac{\hat{k}\pi}{P}})| \neq 0 \Leftrightarrow k \in \{\hat{k}-1, \hat{k}\}$$



$\Longrightarrow$ there is no aliasing of the sinusoidal part, but only of the part that we model as noise!





4.3 Method Overview

Analysis

$$\begin{array}{c}\text{analyse signal} \rightarrow \text{find } N \text{ partials} \rightarrow \text{determine filterbank}\\ \downarrow \\ \text{calculate } 2N \text{ sets of filterbank coefficients + residual}\\ \downarrow \\ \text{calculate wavelet transform (WT) of filterbank coefficients}\\ \downarrow \\ \text{model the WT coefficients sinusoidally and with LPC}\end{array}$$

Synthesis

$$\begin{array}{c}\text{reconstruct WT coefficients}\\ \downarrow \\ \text{perform inverse wavelet transform} \rightarrow \text{get filterbank coefficients}\\ \downarrow \\ \text{inverse filterbank}\\ \downarrow \\ \text{add residual (or not)}\end{array}$$

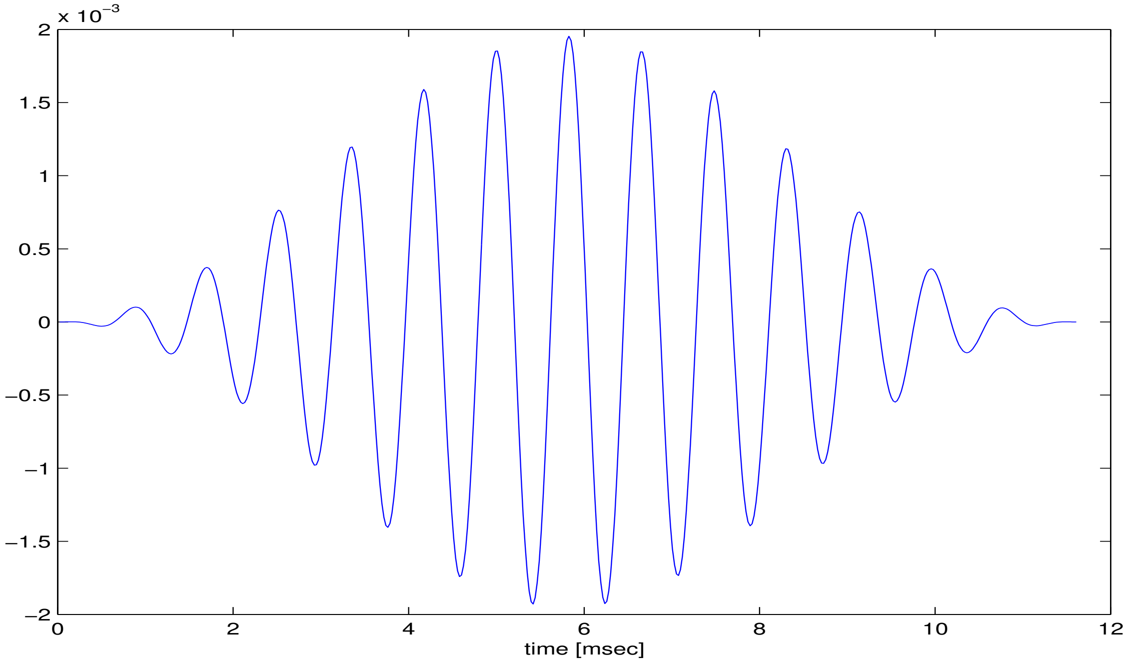
4.4 Sounds

- Original Gong
- Reconstructed from the Filterbank Coefficients
- Synthesized from model parameters
- 1 octave pitch-shifted Gong
- Time-stretched Gong
- Sinusoidal-only Gong
- First wavelet scale only
- Harmonic Gong

5 Time-Frequency Analysis and Granular Synthesis

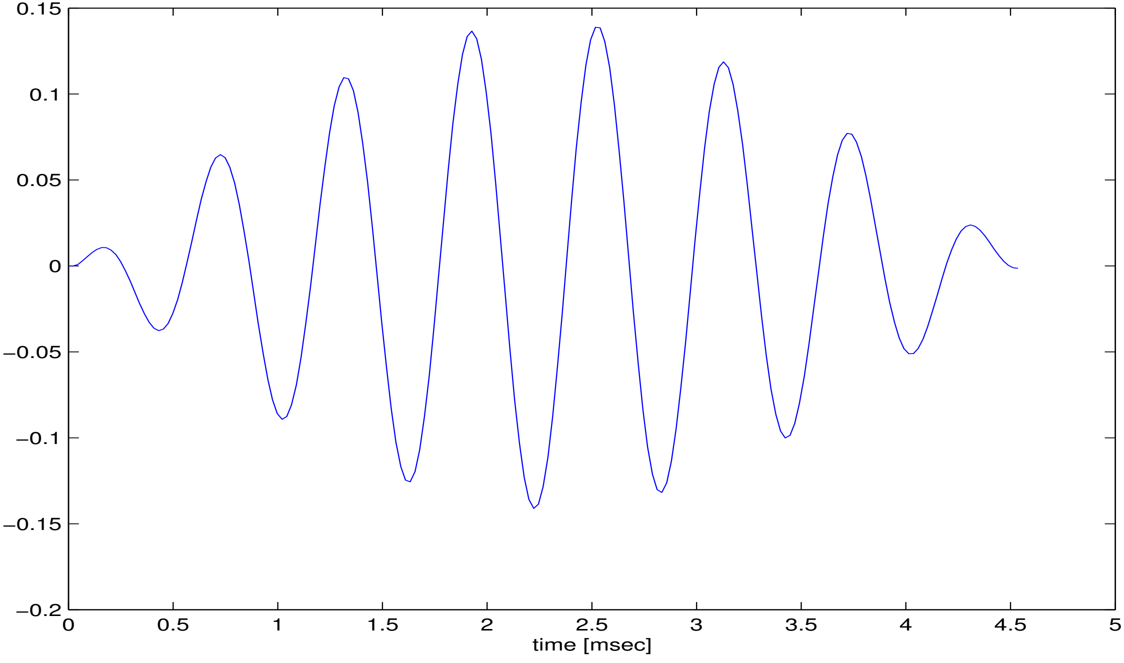
- Any Time-Frequency Transform implements a sort of Granular Synthesis.
- Each coefficient corresponds to a grain
- Grains are played at precise instants (instead of randomly)
- To produce a grain, set all coefficients to zero, except one that will be set to one. Then perform the inverse transform.

Windowed FFT (STFT) grain



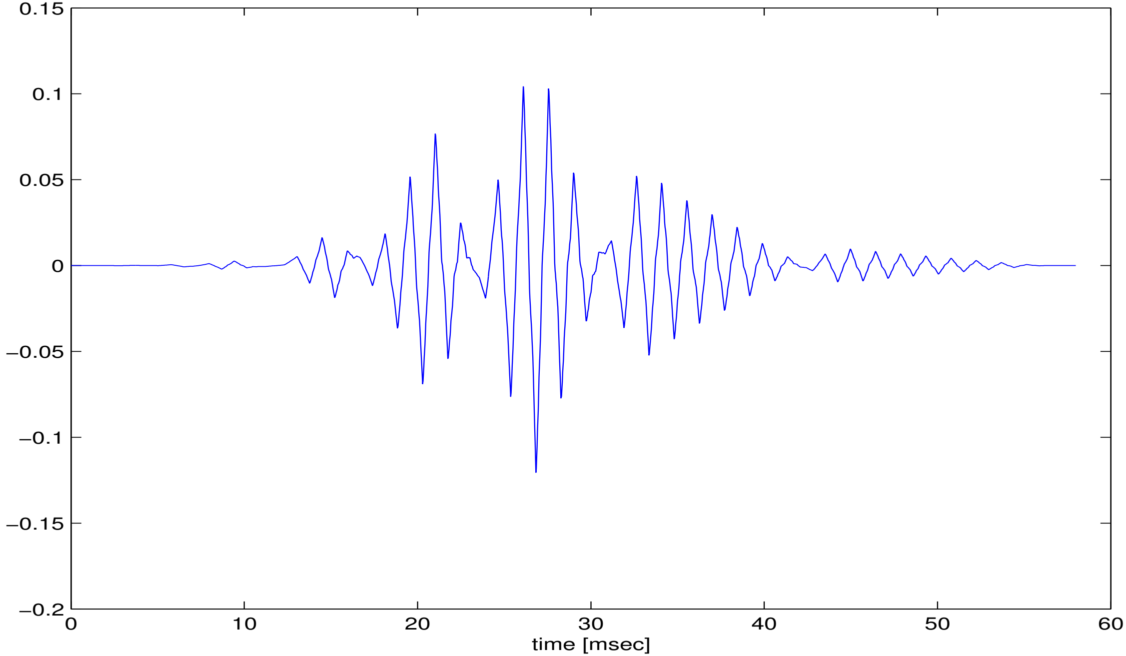
play

Cosine Modulated Filterbank grain



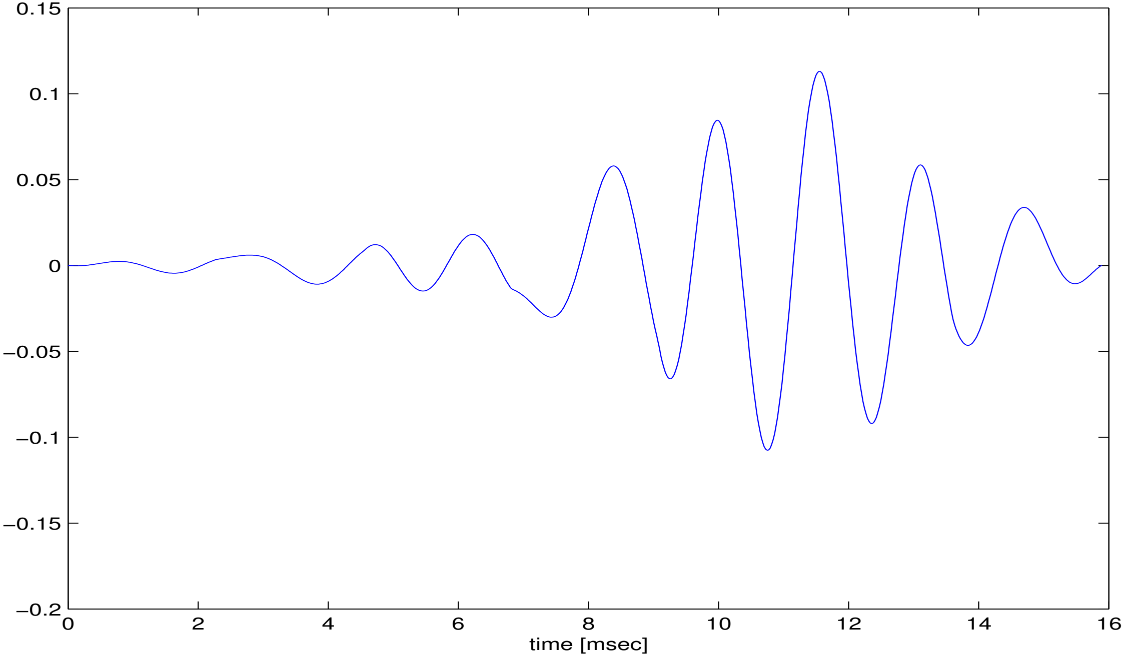
play

Full-tree wavelet "grain"



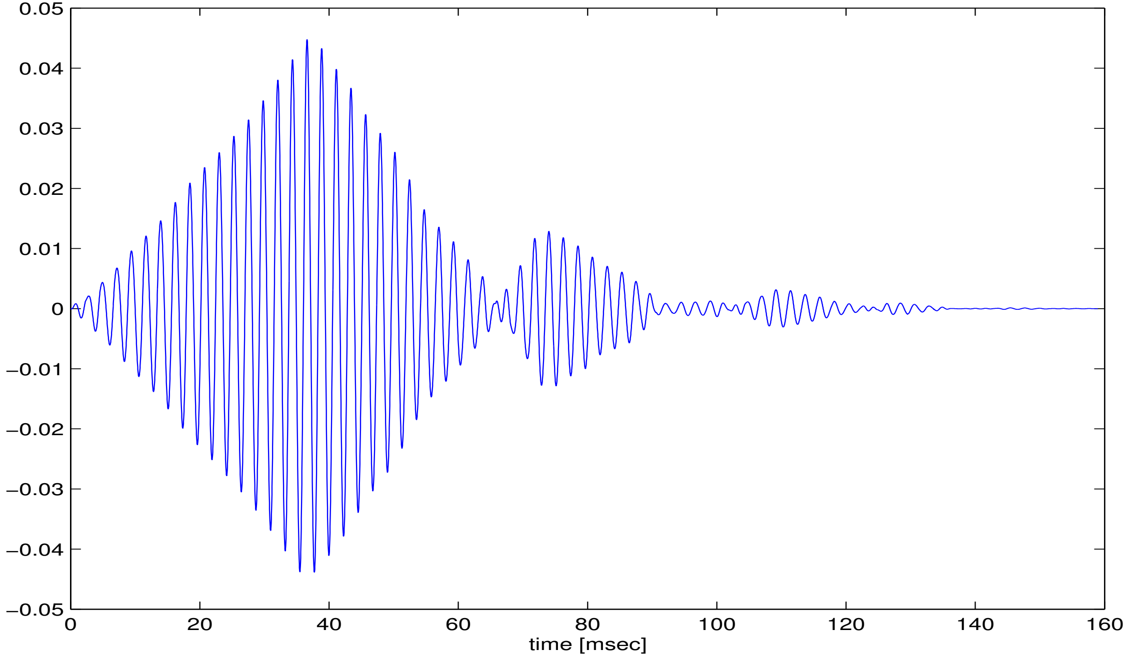
play

HBWT grain (noise part)



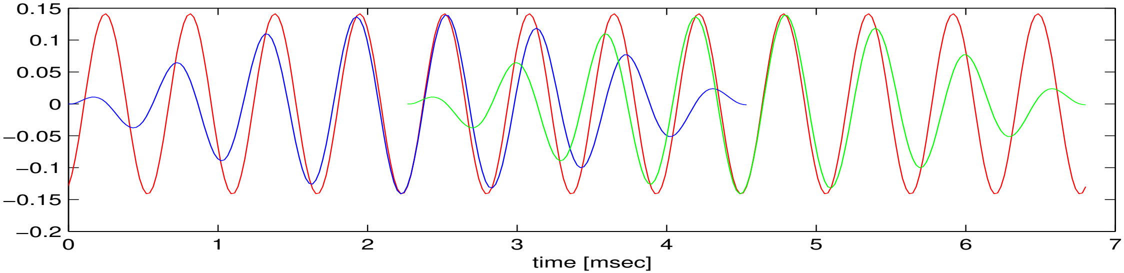
play

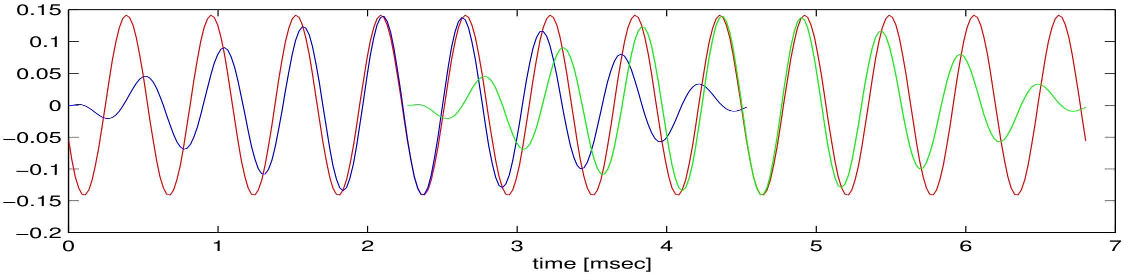
HBWT grain (sinusoidal part)



play

5.1 Time-domain effects

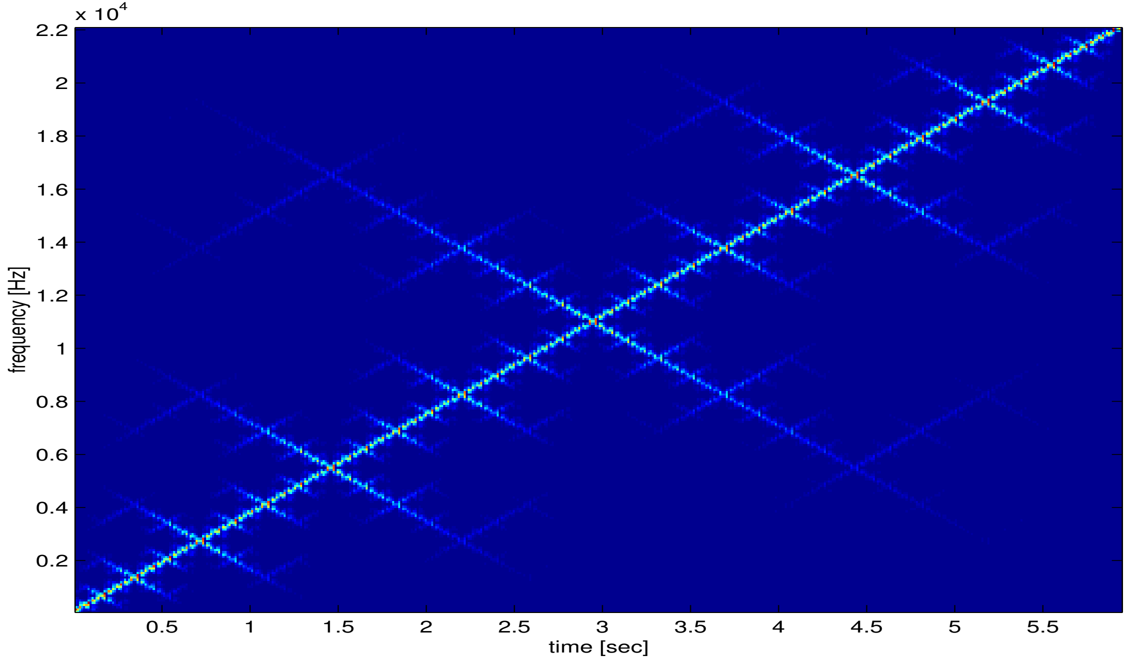




Channel 8: one grain played continuously

Channel 9: one grain played continuously

5.2 Scale of all grains in a 1024-band full-tree wavelet decomposition



play

A References

- Article on Harmonic Band Wavelet Transform by Polotti and Evangelista
  http://lcavwww.epfl.ch/publications/publications/2000/PolottiE00b.pdf
- DAFx 2002 paper on adaptation to inharmonic sounds
  http://lcavwww.epfl.ch/publications/publications/2002/PolottiE02.pdf
- Some material (presentation slides, Matlab functions and `pure data` objects)
  http://www.xsmusic.ch/