Tempo Adjustment with Waveform Similarity based Overlap-Add (WSOLA)

by Mitja Schmakeit

Outline

- Motivation
- OLA
- WSOLA
- Algorithmic Complexity
- Recent R&D

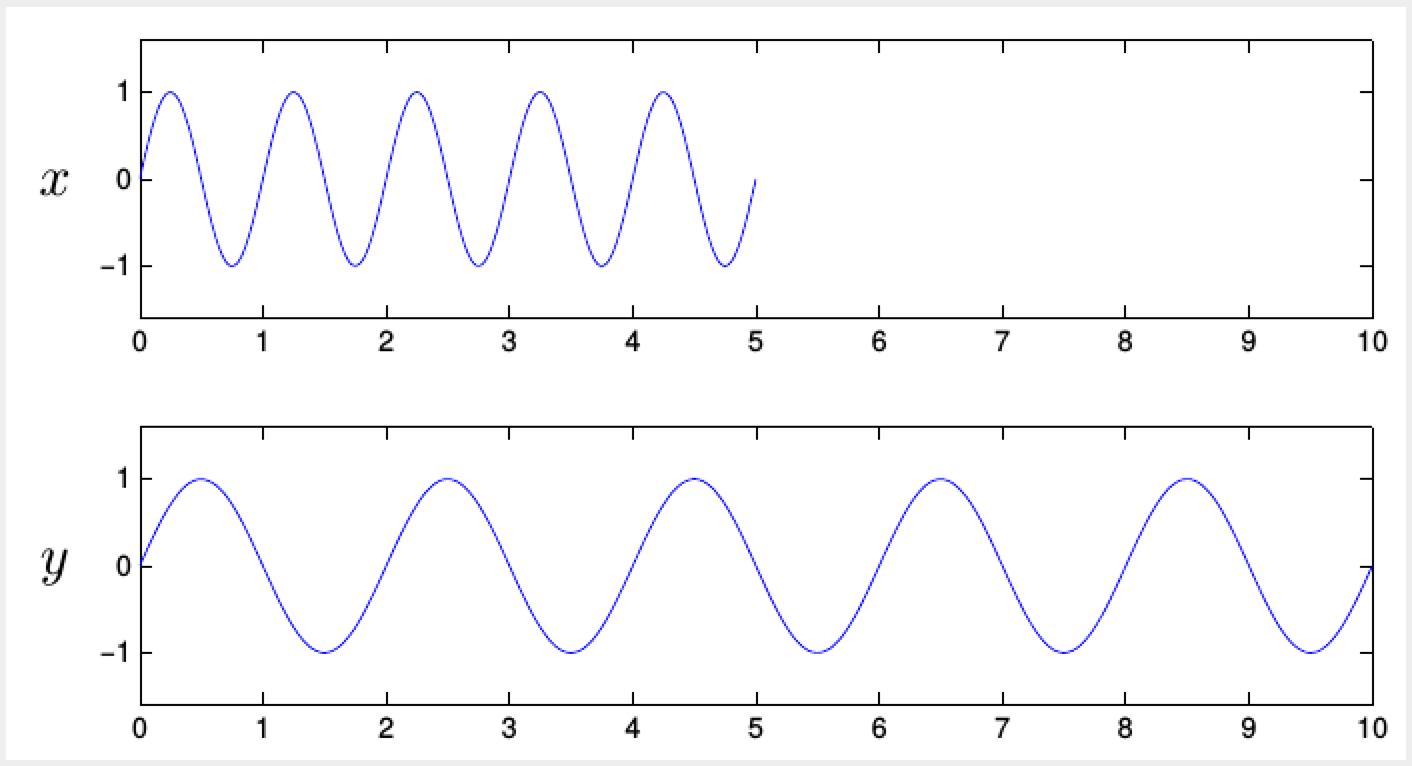
Motivation

Music Mixing requires two songs to be at the same tempo.

If the tempo differs, it needs to be adjusted.

For that task, there exist many different algorithms.

Resampling



[D11]

Algorithms

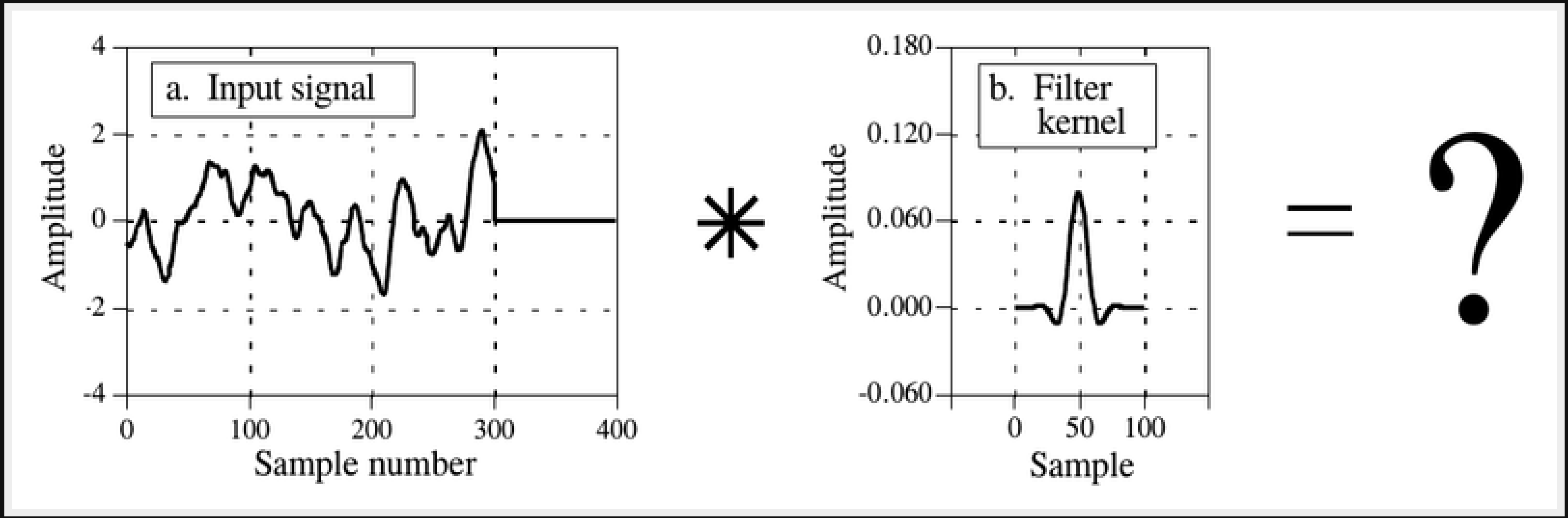
- OLA (Overlap and Add)
- WSOLA (Waveform-similarity based OLA)
- Phase Vocoder

OLA (Overlap and Add)

Basic algorithm in digital signal processing

All more specialized algorithms utilize OLA

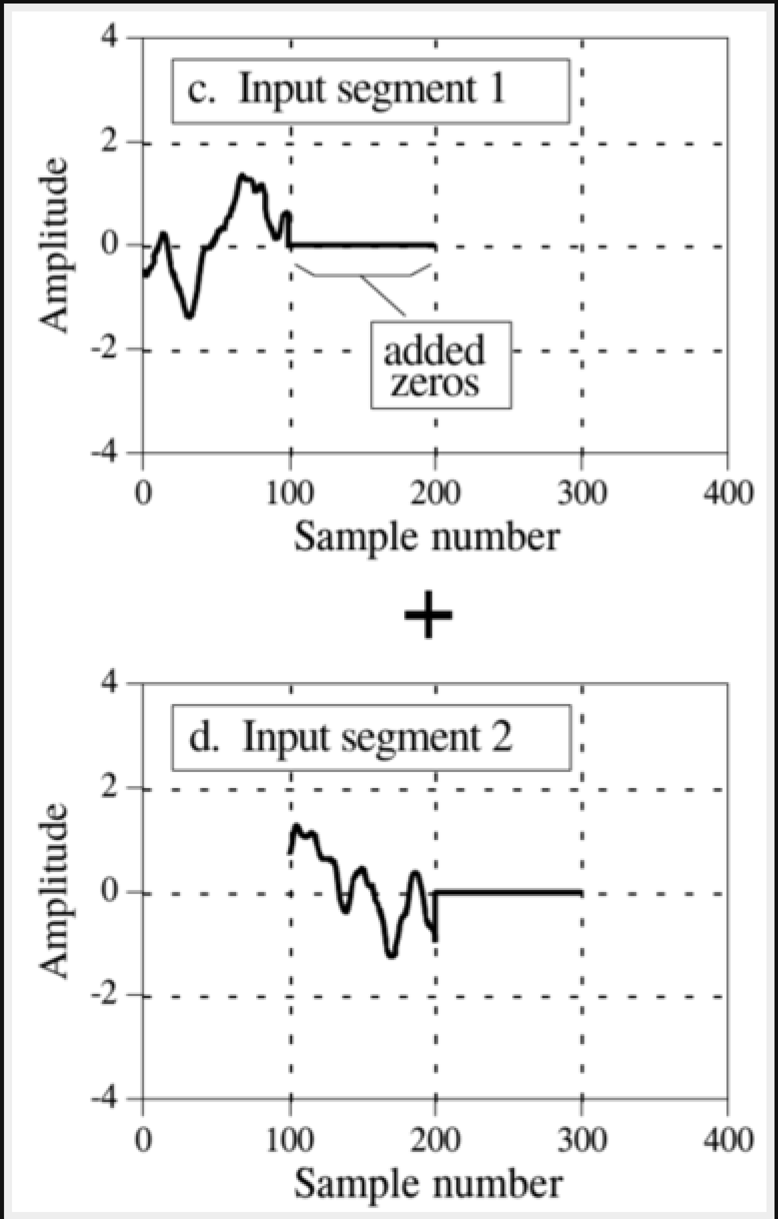
OLA — Input



$x \in \mathbb{R}^M$: input signal of size M

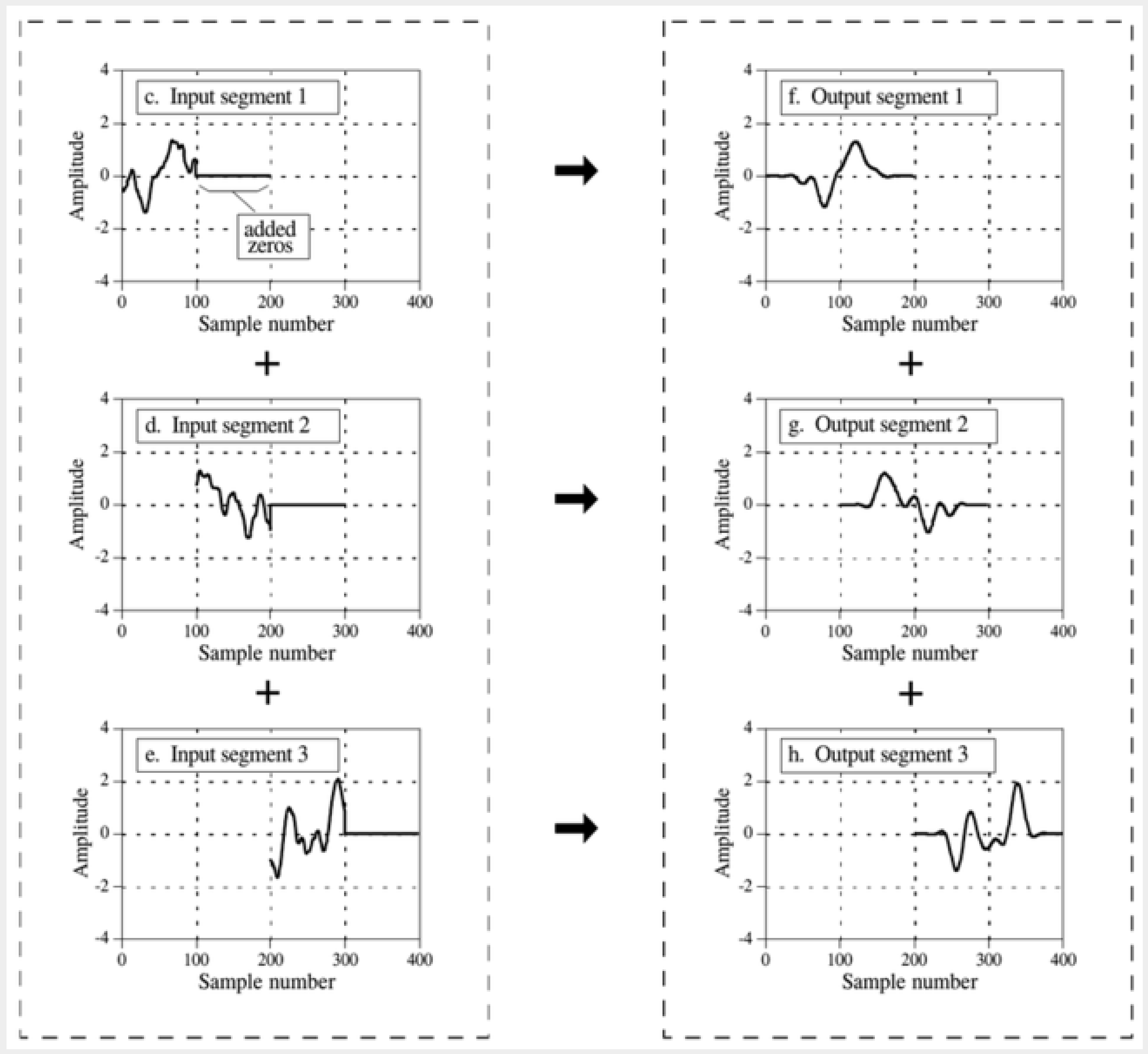
[S03]

OLA — Partitioning into segment 1 and 2



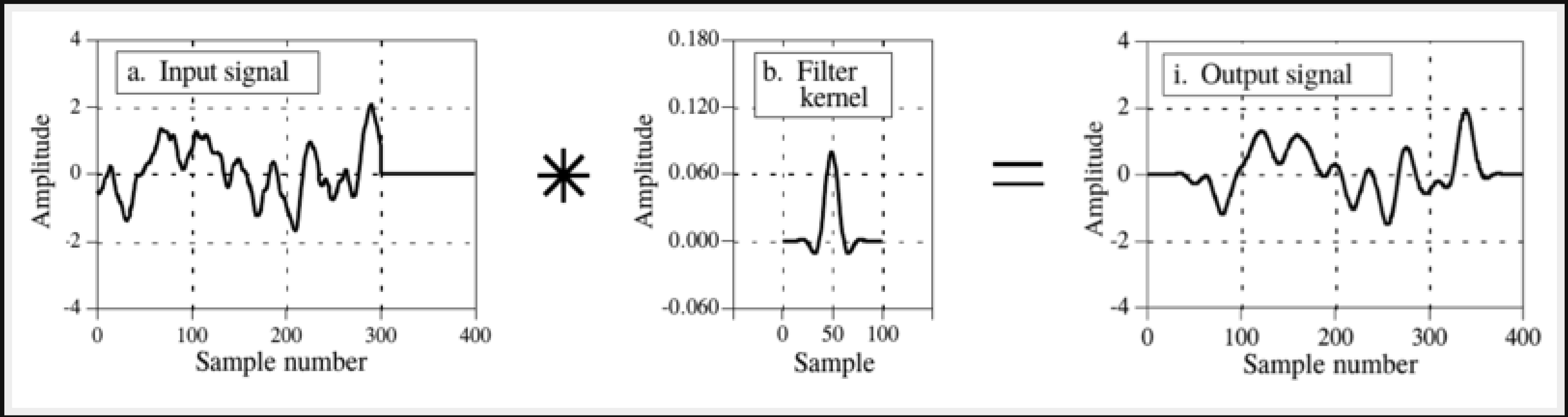
[S03]

OLA — Partitioning



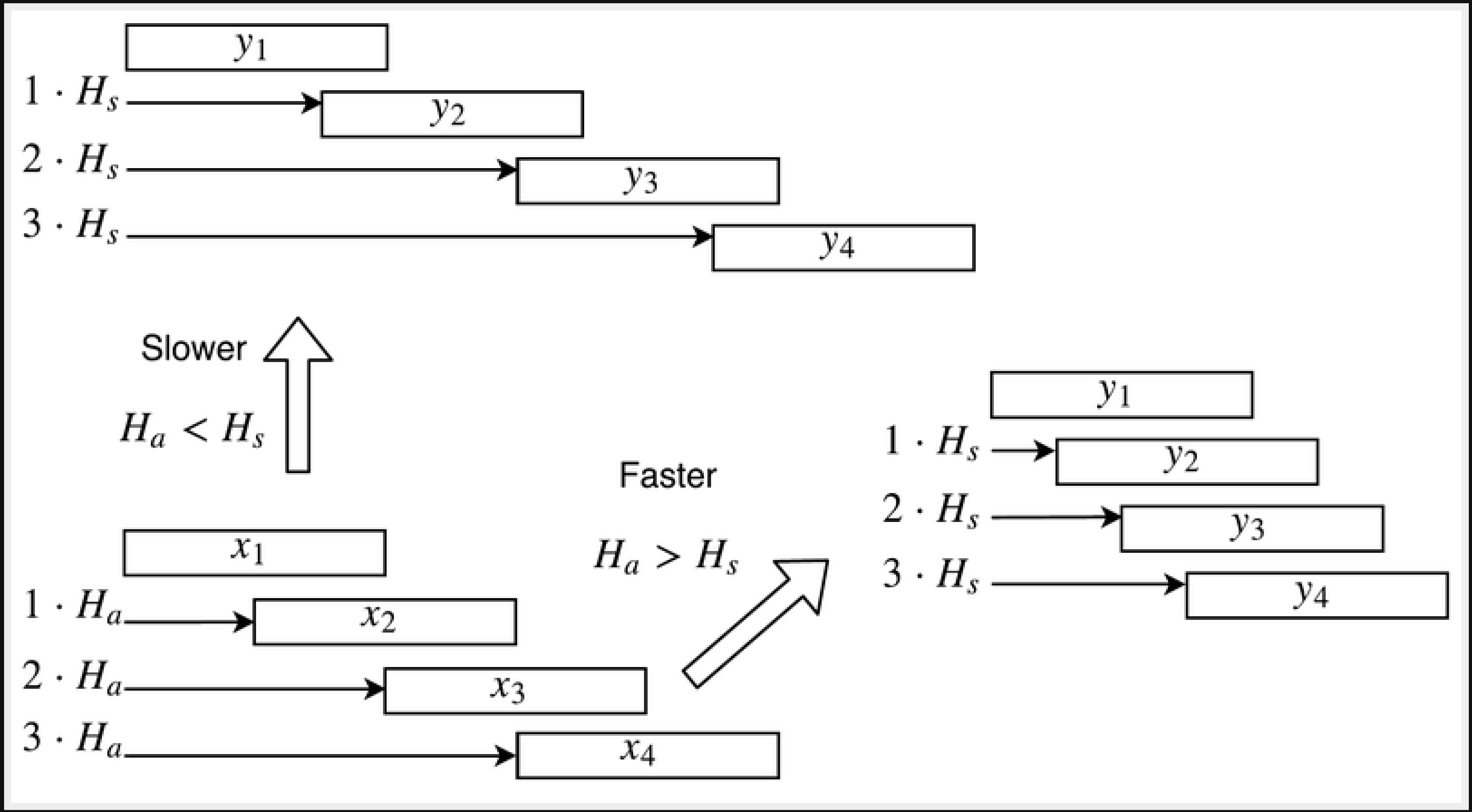
[S03]

OLA — Result



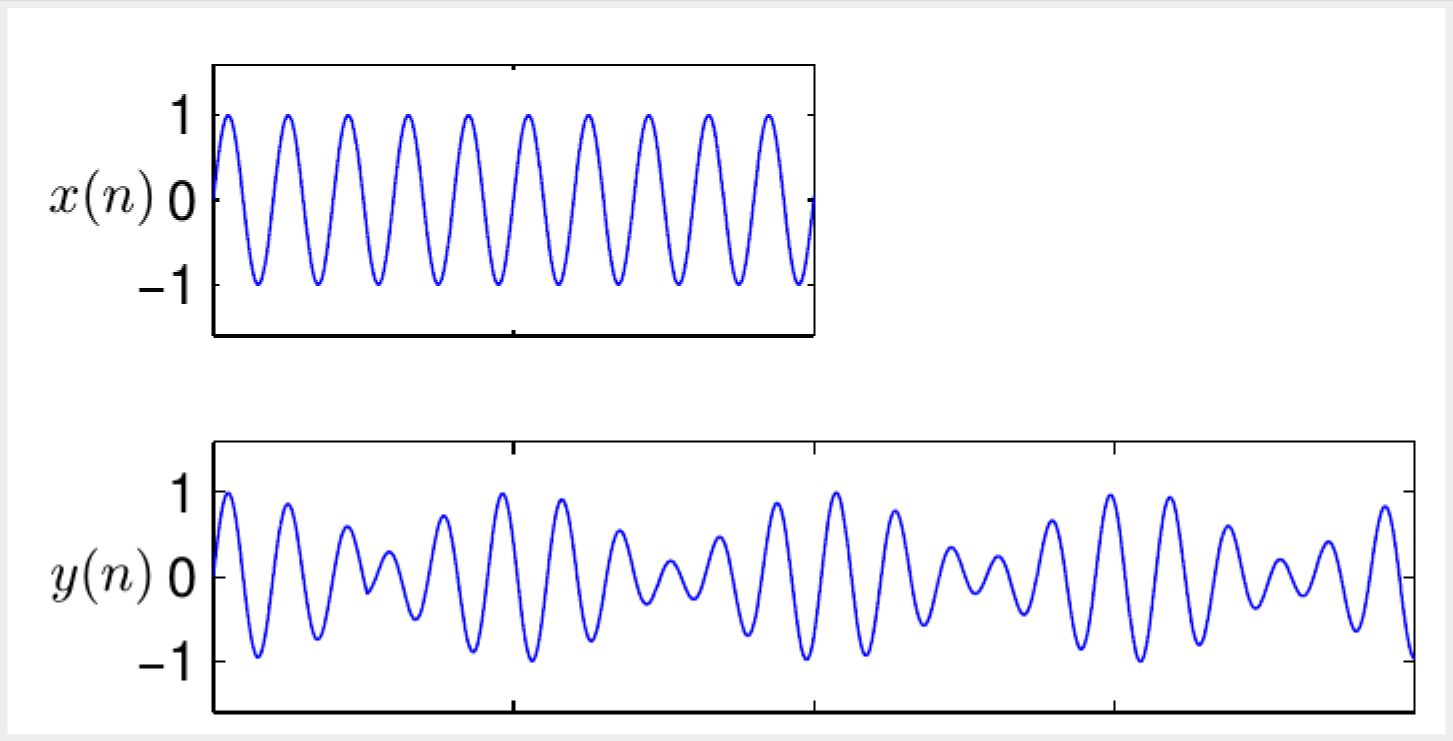
[S03]

OLA — Usage for tempo adjustment



[DMDP16]

OLA – Downsides



OLA does not preserve phase relations between consecutive frames. This means that in the worst case, heavy cancellation effects can occur.

[D11]

OLA – Example

OLA produces significant artifacts in the output signal, which is especially noticable in harmonic structures.

Play original music

Play music made 20% faster with OLA

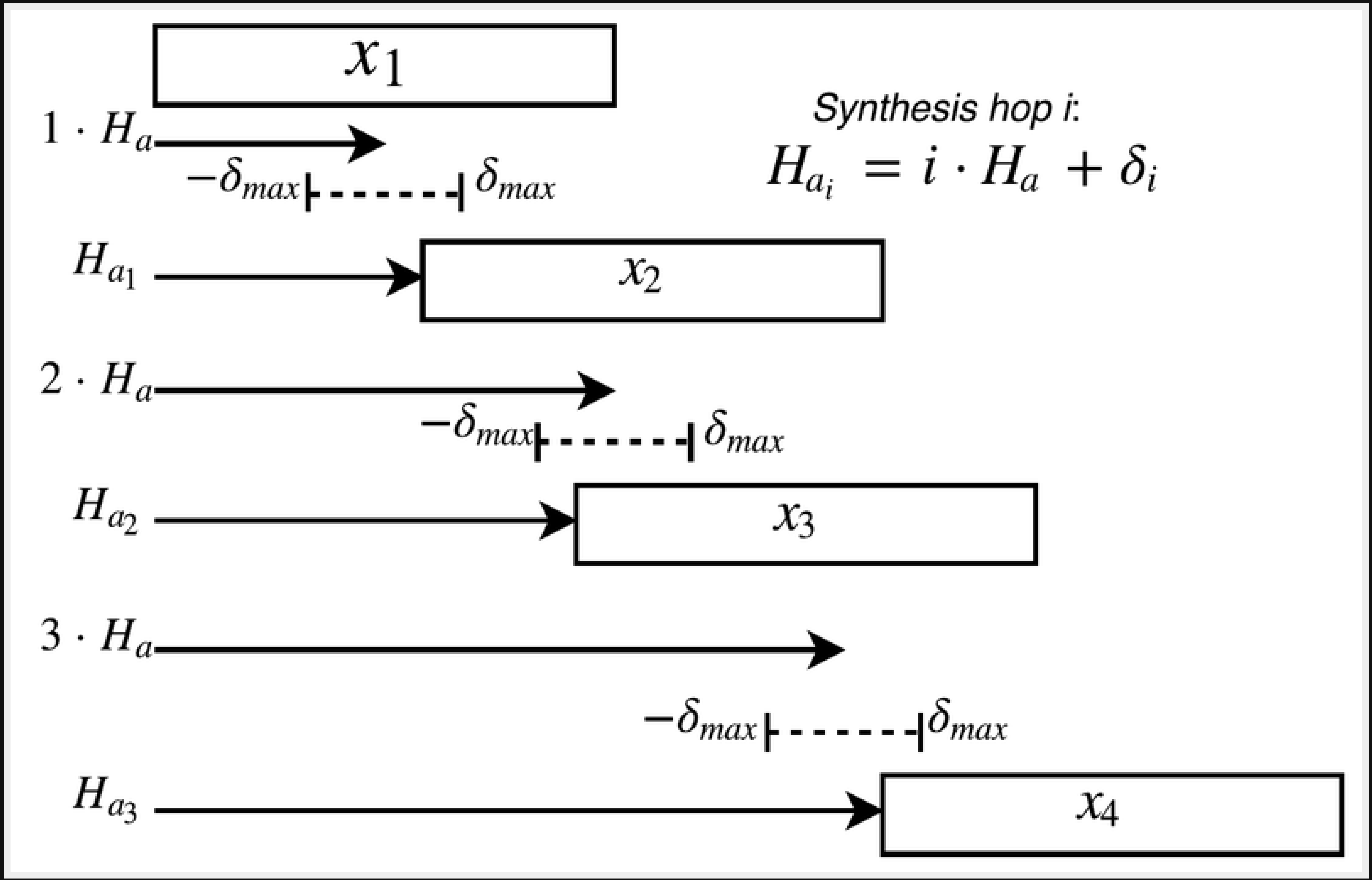
[S13]

Waveform-similarity based OLA (WSOLA)

Developed by W. Verhelst and M. Roelands in 1993 at Vrije Universiteit Brussel [VR93]

Still used today via various audio processing libraries that are used in programs such as Foobar2000, Audacity, Rhythmbox, Firefox and Chrome [DMDP16]

WSOLA — δ windows for similarity matching



Idea: Move each pair of overlapping frames around a bit before merging them, such that their waveforms are as similar as possible

[DMDP16]

WSOLA – Complexity

Space complexity $\mathcal{O}(n)$ (with n being the frame size)

Time complexity $\mathcal{O}(n \cdot \log_2 n)$ [DMDP16]

Therefore, with the right equipment, suited for real-time usage.

There exist many proposals for further reduction of WSOLA complexity (e.g. by estimating the optimal shift [KLK+10])

WSOLA – Audio example

Play original music

Play music made 20% faster with OLA

Play music made 20% faster with WSOLA

[S13]

Recent R&D

- Time Stretching algorithms are numerous, the implementations on different devices are the current problem
- As many applications move to the web, so do audio editing tools like DJ mixing software
- Currently, there exist only few JavaScript implementations that can be used by web audio applications

Name	Algorithm	Audio Artifacts
Vexwarp	Phase Vocoder	Metal Tunnel
tempo-sox.js	WSOLA	Unknown
PhaseVocoder.JS	Phase Vocoder	Smeared Transients
OLA-TS.JS	Modified OLA	Modulation in harm. struct.

[DMDP16]

References

[D11] J. Driedger, *Time-Scale Modification Algorithms for Music Audio Signals*, M.Sc. Thesis, Saarland University, 2011

[DMDP16] B. Dias, D. M. Matos, M. Davies and H. S. Pinto, *Time Stretching & Pitch Shifting with the Web Audio API: Where are we at?* in *Proceedings of Web Audio Conference (WAC)*, 2016

[KLK+10] D. S. Kim et al., *Complexity Reduction of WSOLA-Based Time-Scale Modification Using Signal Period Estimation* in *Future generation Communication and Networking (FGCN)*, 2010, pp. 155—557

References

[S03] S. W. Smith, *FFT Convolution* in *Digital Signal Processing*, Newnes, USA, 2003, pp. 311—318

[S13] R. Schmakeit, *Destiny Can Wait, Music, 2013 (https://youtu.be/J1FX7Klafng)*

[VR93] W. Verhelst, M. Roelands, *An overlap-add technique based on waveform similarity (WSOLA) for high quality time-scale modification of speech* in *Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, 1993, pp. 554—557

Thank You

Questions?