Recent Developments in Signal Processing for Audio and Music

Julius Orion Smith and Kurt James Werner
CCRMA, Stanford University

EURASIP/IEEE Workshop on
Recent Trends in Signal Processing (RTSP-2015)

Cluj-Napoca, Romania

July 6, 2015

Outline

New Enabling Technologies

1. Music Tech Snapshot (Recent CCRMA Dissertations)
   - Analysis, Spatial Audio, Speech, Audio EEG, Music Information Retrieval, Source Separation, Perception, Interaction, Physical Modeling / Virtual Analog
2. Smart-Phones and Tablets
   - High-quality audio in (mono) and out (stereo)
   - Fast multicore processors (exponentially growing speed)
   - Multitouch display screens (5 for iPhone, 11 for iPad)
3. Theory/Architecture Advances
   - 2D Bridge Modeling for Bowed Strings
   - Scattering Delay Networks (SDN)
   - Recent Virtual Analog Dissertations
   - Recent Wave Digital Filter (WDF) research

Recent CCRMA Research

Recent CCRMA Research

Analysis

name	year	dissertation title (or topic)
Jorge Herrera	current	"Computational Beat Tracking with Gradient Frequency Neural Networks"
Craig Sapp	2011	"Computational methods for the analysis of musical structure"
Gregory Sell	2011	"Diffusion-Based Music Analysis"
Woon Seung Yeo	2007	"Raster Scanning: A New Approach to Image Sonification, Sound Visualization, Sound Analysis, and Synthesis"
Randal Leistikow	2006	"Bayesian Modeling of Musical Expectations via Maximum Entropy Stochastic Grammars"
Yi-Wen Liu	2005	"Audio Watermarking Through Parametric Signal Representation"

Spatial Audio

name	year	dissertation title (or topic)
François Germain	current	(sound field reproduction)
Björn Erlach	current	(microphone array)
Tim O'Brien	current	(spherical microphone array)
Elliot Kermit-Canfield	current	(ambisonics)

Recent CCRMA Research

Speech

name	year	dissertation title (or topic)
Jieun Oh	2014	"Affective Analysis and Synthesis of Laughter"
Sook Young Won	2014	"Simulating how humans hear themselves vocalize: a two-parameter spectral model"
Pamornpol Jinachitra	2007	"Robust Structured Voice Extractino for Flexible Expressive Resynthesis"
Rodrigo Segnini	2006	"On the Vocalization of Non-Speech Sounds: Implicit Mechanisms and Musical Applications"

Audio EEG

name	year	dissertation title (or topic)
Irán Román	current	(neurological expectation / ERAN)
Madeline Huberth	current	(EEG / polyphony)
Alexander Chechile	current	(otoacoustic emissions)
Blair Kaneshiro	current	(harmony, meter, tempo; categorical representation)

Recent CCRMA Research

Music Information Retrieval

name	year	dissertation title (or topic)
Jeffrey Smith	2014	"Correlation Analysis of Encoded Music Performance"
Juhan Nam	2013	"Learning Feature Representations for Music Classification"
Kyogu Lee	2007	"A System for Chord Transcription, Key Extraction, and Cadence Recognition from Audio Using Hidden Markov Models Trained with Audio-from-Symbolic Data"
Parag Chordia	2006	"Automatic Transcription of Solo Tabla Music"
Unjung Nam	2004	"A method of automatic recognition of structural boundaries in recorded musical signals"

Source Separation

name	year	dissertation title (or topic)
Dennis Sun	current	(nonnegative matrix factorization)
Nicholas Bryan	2014	"Interactive Sound Source Separation"
Gautham Mysore	2010	"A Non-negative Framework for Joint Modeling of Spectral Structure and Temporal Dynamics in Sound Mixture"
Aaron Steven Master	2006	"Stereo Music Source Separation via Bayesian Modeling"

Recent CCRMA Research

Perception

name	year	dissertation title (or topic)
Hyung-Suk Kim	current	(musical texture)
Emily Graber	current	(timbre perception)
Kitty Shi	current	(computational models of auditory perception and stream segregation)
Miriam Kolar	2013	"Archaeological Psychoacoustics at Chavin de Huántar, Perú"
Hiroko Terasawa	2009	"A hybrid model for timbre perception"
Ryan James Cassidy	2008	"Auditory Signal Processing to Improve Impaired Listening Experiences via Efficient, Loudness-Based Algorithms"
Matthew Wright	2007	"The shape of an instant: measuring and modeling perceptual attack time with probability density functions (if a tree falls in the forest, when did 57 people hear it make a sound?)"
Harvey Thornburg	2005	"Detection and Modeling of Transient audio Signals with Prior Information"

Recent CCRMA Research

Interaction

name	year	dissertation title (or topic)
Romain Michon	current	(interfaces for real-time physical model control, 3D printing)
John Granzow	current	(3D printing for acoustics)
Spencer Salazar	current	“Handwritten Computer Music Composition and Design”
Luke Dahl	2014	“Timing of Discrete Music Air-Gestures”
Rob Hamilton	2014	“Perceptually Coherent Mapping Schemata for Virtual Space and Musical Method”
Hongchan Choi	2014	“The Design of a collaborative music system: rethinking music creation and production platform through web technology”
Ed Berdahl	2009	“Applications of Feedback Control to Musical Instrument Design”
Jonathan Norton	2008	“Motion Capture To Build A Foundation For A Computer Controlled Instrument By Study Of Classical Guitar Performance”
Ge Wang	2008	“The ChucK Audio Programming Language: A Strongly-timed and On-the-fly Environ/mentality”
Lonny Chu	2004	“Haptic Interfaces for Audio Navigation”

Recent CCRMA Research

Physical Modeling / Virtual Analog

name	year	dissertation title (or topic)
Kurt James Werner	current	“Like an 808: Circuit Models and Analog Musicology of the TR-808, including Advances in Wave Digital Filter Theory”
Jack Perng	2012	“Physical Modeling of the Harpsichord Plectrum-String Interaction”
David Yeh	2009	“Digital Implementation of Musical Distortion Circuits by Analysis and Simulation”
Michale Gurevich	2007	“Computational Acoustic Modeling of Cetacean Vocalizations”
Scott Van Duyne	2007	“Digital Filtering Applications to Modeling Wave Propagation in Springs, Strings, Membranes, and Acoustical Space”
Tim Stilson	2006	“Efficiently-Variable Non-Oversampled Algorithms in Virtual-Analog Music Synthesis: a Root-Locus Perspective”
Patricio de la Cuadra	2006	“The sound of oscillating air jets: Physics, modeling, and simulation in flute-like instruments”
Tamara Smyth	2004	“Applications of Bioacoustics to the Development of Musical Instrument Technology”
Stefania Serafin	2004	“The sound of friction: real-time models, playability and musical applications”

Smart-Phone/Tablet Example: moForte Guitar

moForte Guitar for iOS and (later) Android

Real-time on iPhone 4S and iPad 2 (and later)

Guitar and effects written in the FAUST language:

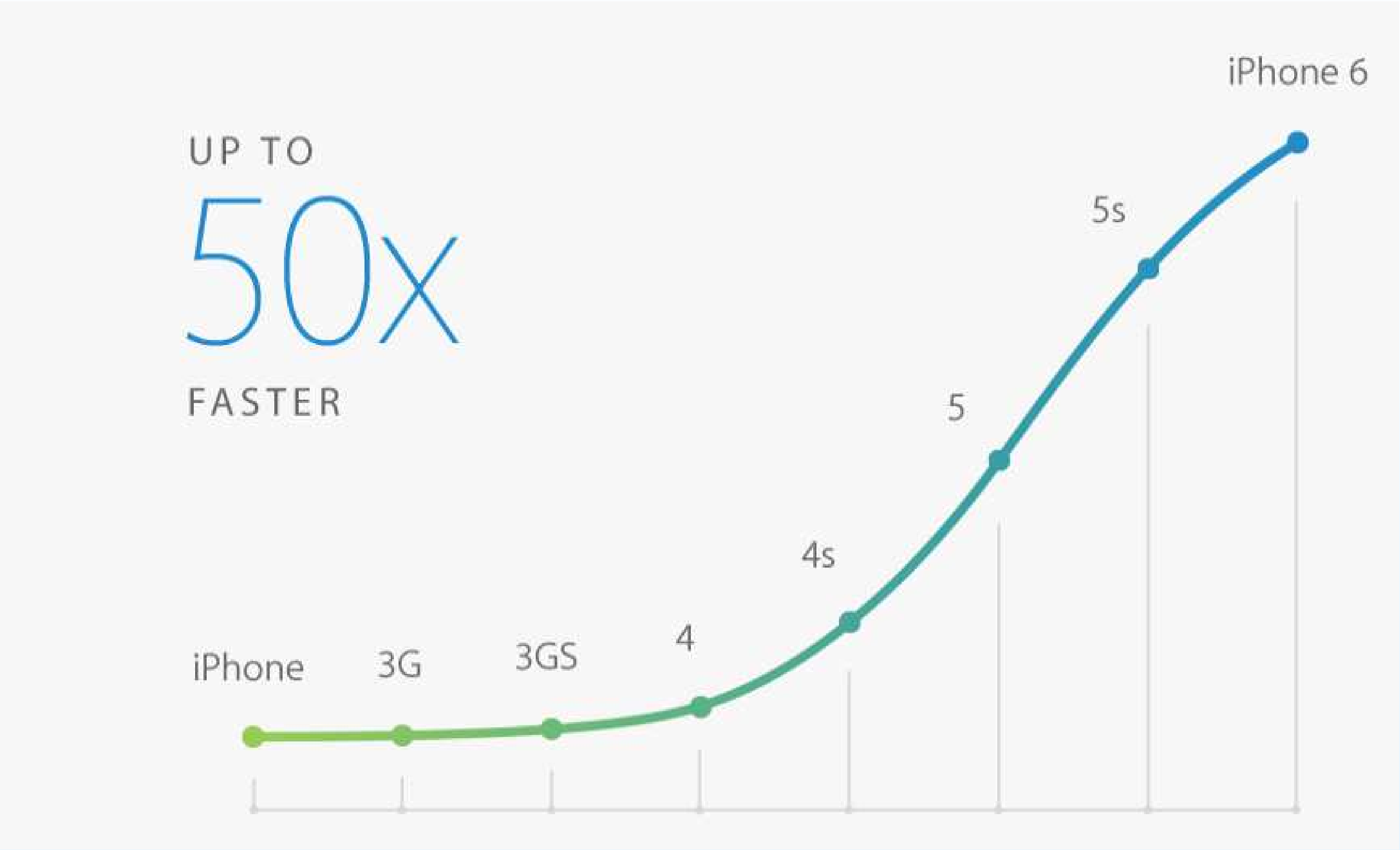
- Full physically modeled electric-guitar + effects:

Six vibrating strings —	general excitations
Distortion	Feedback
Compression	Wah pedal or Autowah
Phaser	Flanger
Five-band parametric equalizer	Reverb

- Responds to
  accelerometer, gyros, touches (plucks), swipes (strumming), ...
- It is challenging to fully utilize *five* points of multitouch on the iPhone and *eleven* on the iPad!
- Android audio latency is still a significant issue

All These Guitar Effects Run in Real Time on the iPhone 4S or iPad 2

CPU Performance in Phones



Distortion Guitar Sound Examples on a Tablet

GeoShred Demo

- YouTube (./mov/GeoShredPreviewNAMM2015.mov)

Recent Theory/Architecture Advances

Enabling Technologies

CCRMA Research

Tablet Instruments

Theory/Architecture
- Overview
- 2D Bridge Modeling
- Waveguide Reverb
- recent VA diss.
- WDF Overview
- WDF: single NL
- WDF: one-port NL
- WDF: cross-control
- WDF: linearized
- WDF: PWL
- WDF: iterative
- WDF: topologies

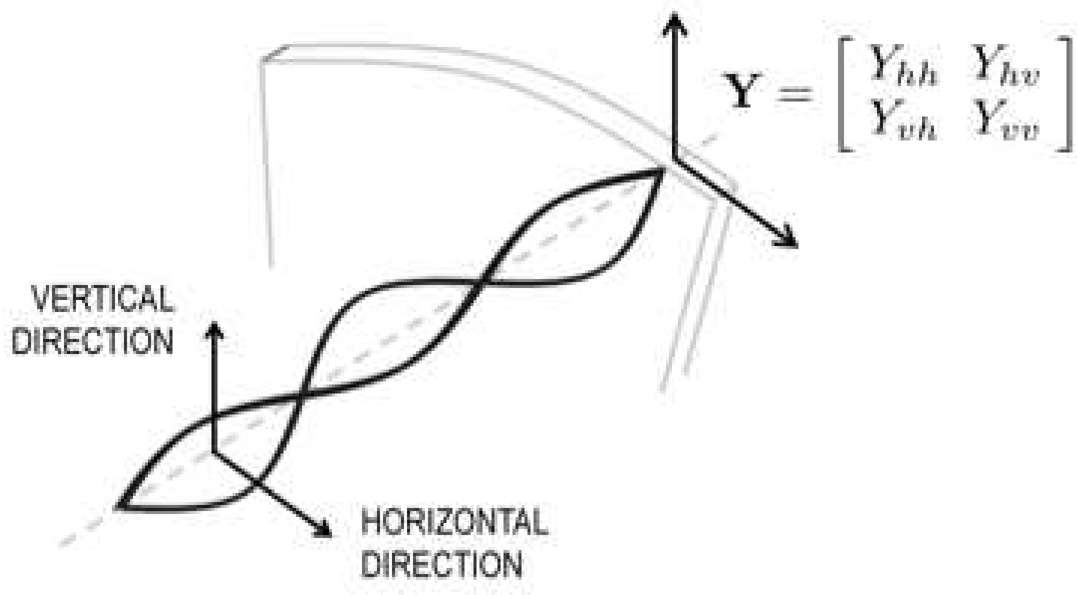
Summary

Overview

1. 2D Bridge Modeling for Bowed Strings
2. Scattering Delay Networks (SDN)
3. Recent Virtual Analog Dissertations
4. Recent Wave Digital Filter (WDF)research

2D Bridge Modeling for Bowed Strings

Reference: E. Maestre, G. P. Scavone, and J. O. Smith III, "Digital Modeling of String Instrument Bridge Reflectance and Body Radiativity for Sound Synthesis by Digital Waveguides," 2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, Oct. 18–21. See also SMAC-13 paper: `http://www.speech.kth.se/smac-smc-2013/`

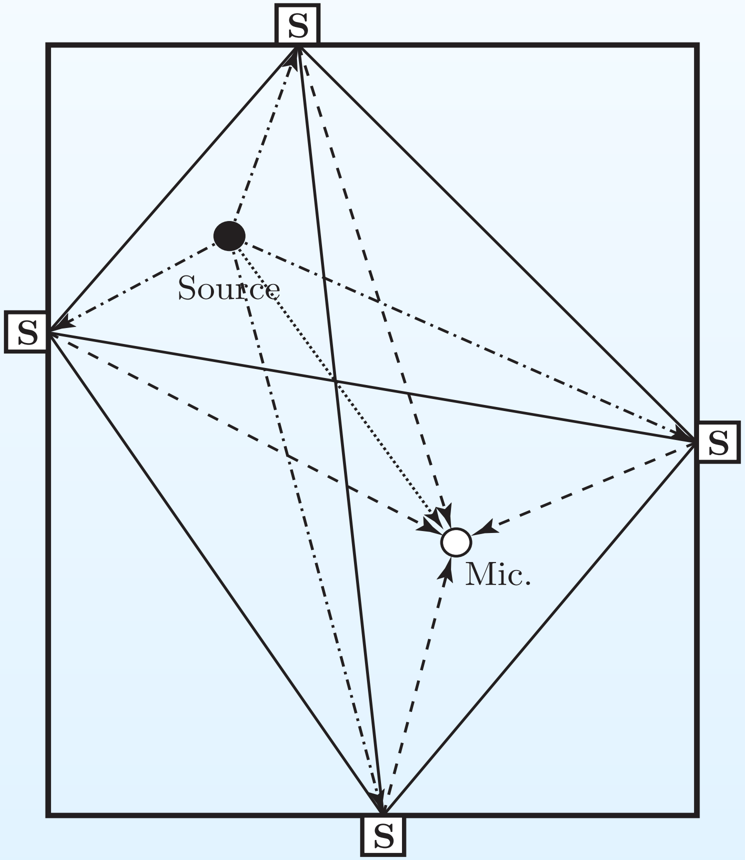


Two-dimensional Bridge Model

Scattering Delay Networks (SDN)

Reference: E. De Sena, H. Hacıhabiboğlu, Z. Cvetković, and J. O. Smith III, "Efficient Synthesis of Room Acoustics via Scattering Delay Networks," IEEE/ACM Trans. Audio, Speech, Language Process., vol. 23, no. 9, Sept. 2015.

`http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7113826`



Recent Virtual Analog Dissertations

name	year	dissertation title
Tim Stilson	2006	"Efficiently-Variable Non-Oversampled Algorithms in Virtual-Analog Music Synthesis: a Root-Locus Perspective"
David Yeh	2009	"Digital Implementation of Musical Distortion Circuits by Analysis and Simulation"
Kristjan Dempwolf	2012	"Modellierung analoger Gitarrenverstrker mit digitaler Signalverarbeitung"
Mačák, Jaromír	2012	"Real-Time Digital Simulation of Guitar Amplifiers as Audio Effects"
Rafael C. D. de Paiva	2013	"Circuit modeling studies related to guitars and audio processing"
Julian Parker	2013	"Dispersive Systems in Musical Audio Signal Processing"
Jussi Pekonen	2014	"Filter-Based Oscillator Algorithms for Virtual Analog Synthesis"
Stefan D'Angelo	2014	"Virtual Analog Modeling of Nonlinear Musical Circuits"
Kurt James Werner	2016(?)	"Like an 808: Circuit Models and Analog Musicology of the TR-808, including Advances in Wave Digital Filter Theory"

Wave Digital Filters

Wave Digital Filters (WDFs) work by:

- Port-wise consideration of Kirchhoff-domain circuit
- Discretize reactive elements w/ Bilinear Transformation (BLT):

$$s \leftarrow c\frac{1 - z^{-1}}{1 + z^{-1}}, \quad c = 2/T \text{ or } 1, \text{ typically} \tag{1}$$

- Move into the wave domain, with incident wave a_n and reflected wave b_n, incorporating **arbitrary** port resistance R_n, at each port n

$$a_n = v_n + i_n R_n \tag{2}$$

$$b_n = v_n - i_n R_n \tag{3}$$

- Scattering matrices at every impedance mismatch (typically 3-port series and parallel connections)
- Choose port resistances to cancel delay-free loops

Wave Digital Filters

We'll review developments in nonlinear WDFs:

1. Single Nonlinearity
2. Consolidated One-Port Combinations
3. Cross-Controlled Multiport
4. Linearized
5. Piecewise Linear Models
6. Iterative Schemes
7. Topological Aspects

WDFs: Single Nonlinearity I

Reference: K. Meerkötter and R. Scholz, “Digital simulation of nonlinear circuits by wave digital filter principles,” Proc. IEEE Int. Symp. Circuits Syst. June 1989.

- Accomodate ONE one-port NL element, e.g.:
  - Ideal rectifier (ideal diode)
  - Piecewise linear resistance
- Can view as lookup table with interpolation or piece-wise linear segments

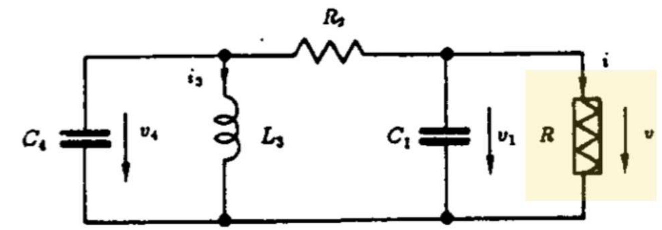


Fig. 2. Nonlinear circuit according to Ref. 2.

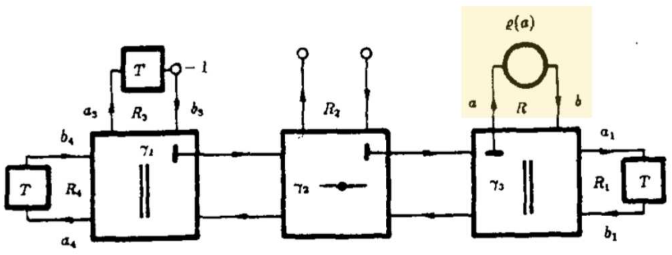


Fig. 4. Wave digital model of the circuit of Fig. 2.

WDFs: Single Nonlinearity II

Reference: K. Meerkötter and R. Scholz, "Digital simulation of nonlinear circuits by wave digital filter principles," Proc. IEEE Int. Symp. Circuits Syst. June 1989.

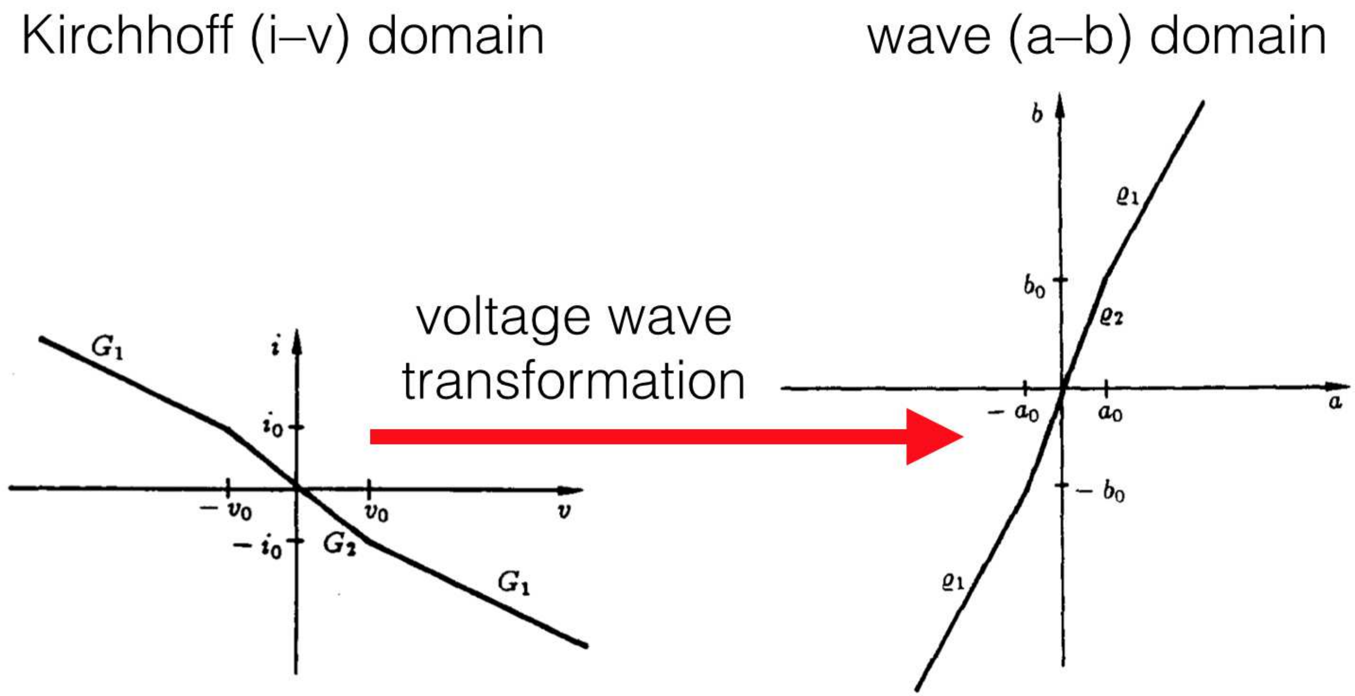


Fig. 3. Characteristic of the nonlinear resistance defined by (16).

$$i = G_1 v + \frac{1}{2}(G_2 - G_1)\left(|v + v_0| - |v - v_0|\right), \qquad (16)$$

Fig. 5. Plot of the characteristic defined by (17).

$$b = \varrho(a) = \varrho_1 a + \frac{1}{2}(\varrho_2 - \varrho_1)\left(|a + a_0| - |a - a_0|\right), \quad (17)$$

WDFs: Single Nonlinearity III

Reference: A. Sarti and G. De Sanctis, “Systematic methods for the implementation of nonlinear wave-digital structures,” IEEE Trans. Circuits Syst. I: Reg. Papers, vol. 56, no. 2, pp. 460–472, 2009.

- Binary connection tree (BCT) systematizes WDF with only series and parallel connections
- Up to one nonlinearity
- N-port series (parallel) implemented with $N-2$ 3-port adaptors
- see: A. Fettweis and K. Meerkötter, “On adaptors for wave digital filters,” 1975.

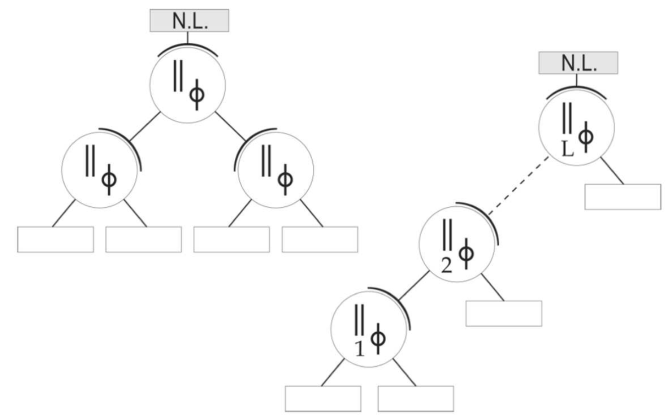


Fig. 5. Two examples of BCTs. (left) Generic one and (right) chainlike circuit. The circular box represents an instantaneous adaptor, in which the adapted port is clearly specified. This particular notational choice simplifies the drawing of connection trees with a great amount of branching.

WDFs: Single Nonlinearity IV

Reference: A. Sarti and G. De Poli, "Toward nonlinear wave digital filters," IEEE Trans. Signal Process., vol. 47, no. 6, pp. 1654–1668, 1999.

- Use "mutators" from classical network theory to enable, e.g., nonlinear q–v relationships
- These are needed for nonlinear elements "with memory"
- For example, nonlinear capacitors and inductors where flux or charge can saturate

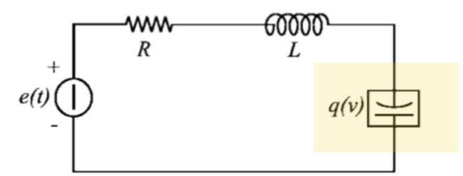


Fig. 6. Electrical circuit of the anharmonic oscillator.

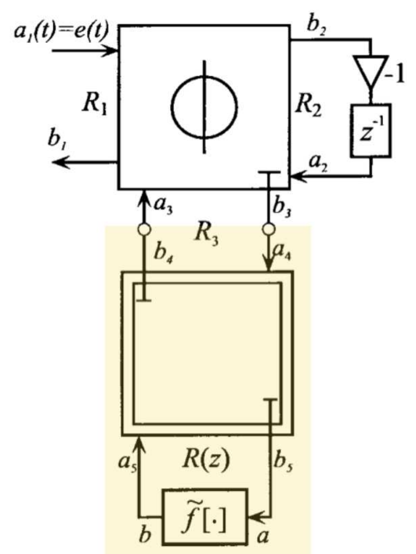


Fig. 9. Wave implementation of the anharmonic oscillator based on instantaneous adaptation. The double-bordered box represents an R-C mutator, and the presence of two "stubs" in its outputs denotes the absence of local instantaneous reflections.

WDFs: One-Port Combinations I

Reference: D. T. Yeh and J. O. Smith III, "Simulating guitar distortion circuits using wave digital and nonlinear state-space formulations," Proc. Int. Conf. Digital Audio Effects (DAFx-08), pp. 19–26, 2008.

$$\frac{b-a}{2R} = I_s\left(e^{\frac{a+b}{2V_T}} - 1\right) - I_s\left(e^{-\frac{a+b}{2V_T}} - 1\right) \tag{4}$$

- Multiple nonlinearities handled by combining into a single one-port
- Implicit nonlinear function solved as $b = f(a)$ with numerical methods

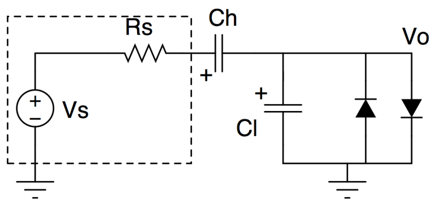


Figure 6: Schematic of the diode clipper with high-pass and low-pass capacitors.

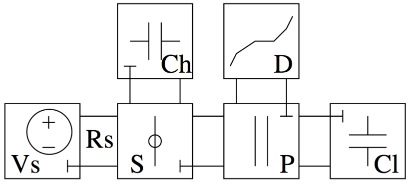


Figure 7: WDF tree of the two-capacitor diode clipper. Diode D is root.

WDFs: One-Port Combinations II

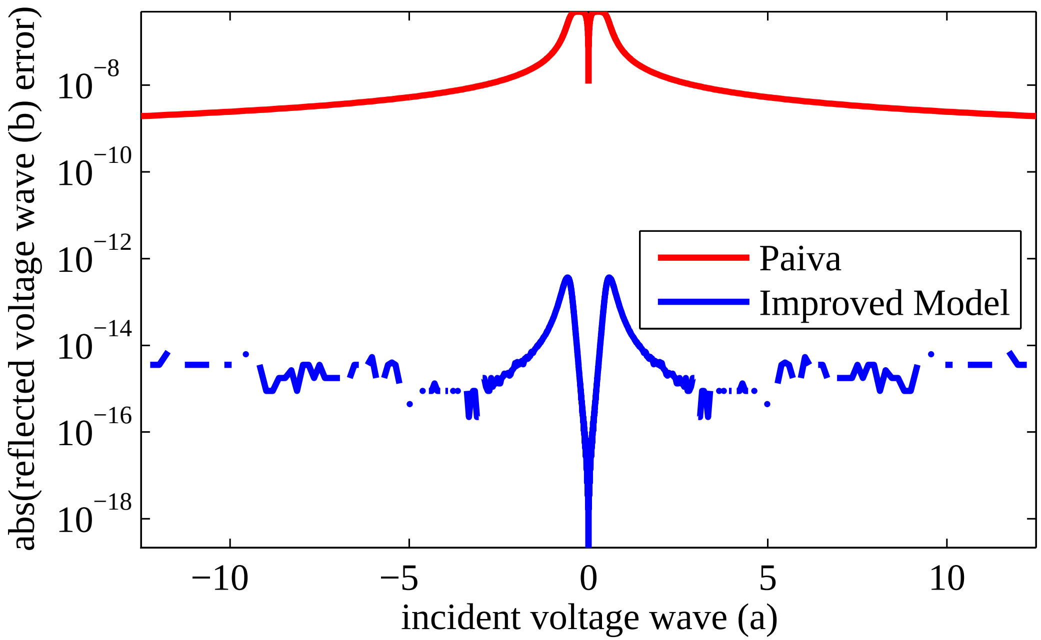
Reference: R. C. D. Paiva, S. D'Angelo, J. Pakarinen, and V. Välimäki, "Emulation of operational amplifiers and diodes in audio distortion circuits," IEEE Tran. Circuits Syst. II: Expr. Briefs, vol. 59, no. 10, pp. 688-692, 2012.

$b = f(a)$ for diode pair solved using *Lambert* $\mathcal{W}$ *function*, assuming one diode dominates:

- One diode (Kirchhoff): $i = I_s \left(e^{\frac{v}{V_T}} - 1 \right)$
- One diode (wave, implicit): $e^{\frac{a+b}{2V_T}} = \frac{a-b}{2RI_s} + 1$
- One diode (wave, explicit):
$$b = f(a) = a + 2RI_s - 2V_T \mathcal{W} \left(\frac{RI_s}{V_T} e^{\frac{RI_s + a}{V_T}} \right)$$
- Diode pair (approx., assuming one diode dominates):
$$b = \text{sgn}(a) \cdot f(|a|)$$

WDFs: One-Port Combinations III

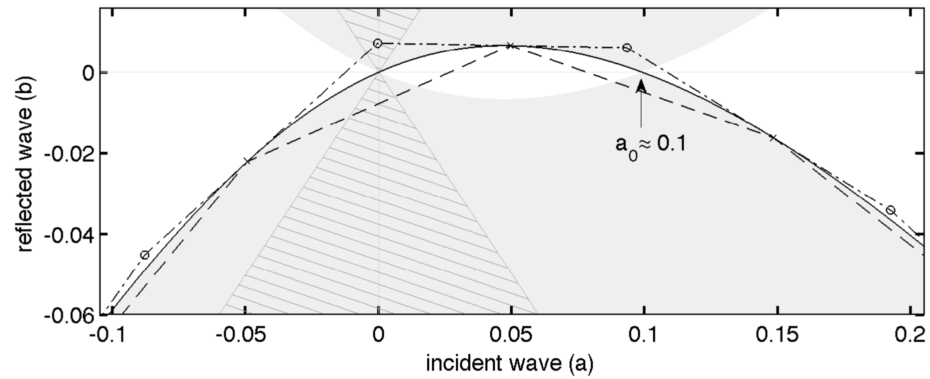
Reference: K. J. Werner, V. Nangia, A. Bernardini, J. O. Smith III, and A. Sarti, "An Improved and Generalized Diode Clipper Model for Wave Digital Filters," Proc. Audio Eng. Soc. (AES), New York, NY, Oct.–Nov. 2015.



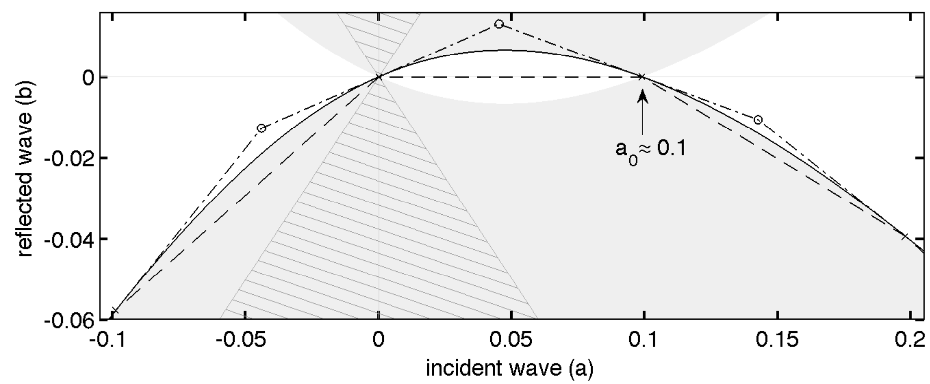
Explicit $b = f(a)$ model improved by cancelling some approximation error of Piava *et al.* model with additional Lambert $\mathcal{W}$ term.

WDFs: One-Port Combinations IV

Reference: K. J. Werner and J. O. Smith III, “An Energetic Interpretation of Nonlinear Wave Digital Filter Lookup Table Error,” Proc. Int. Symp. Signals, Circuits, Syst. (ISSCS), Iasi, Romani, July 2015.



Linear secant interpolation and tangent extrapolation can be incrementally (gray) or globally (thatched) non-passive.



Choosing table points and secant/tangent properly (considering $\text{sgn}(a)$ and a'') yields interpolation methods that respect passivity.

WDFs: Cross-Control Multiport I

Reference: M. Karjalainen and J. Pakarinen, “Wave digital simulation of a vacuum-tube amplifier,” Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Toulouse, France, May 14–19 2006.

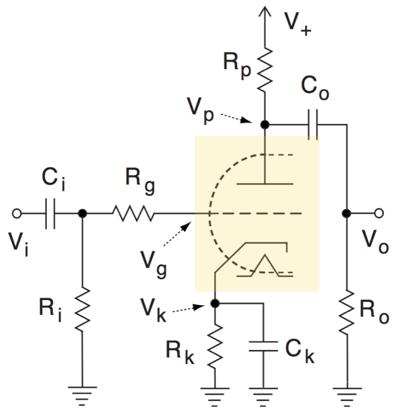


Fig. 1. A typical triode amplifier stage.

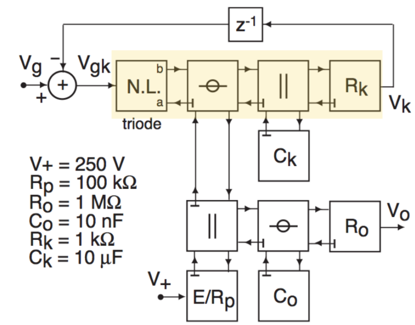
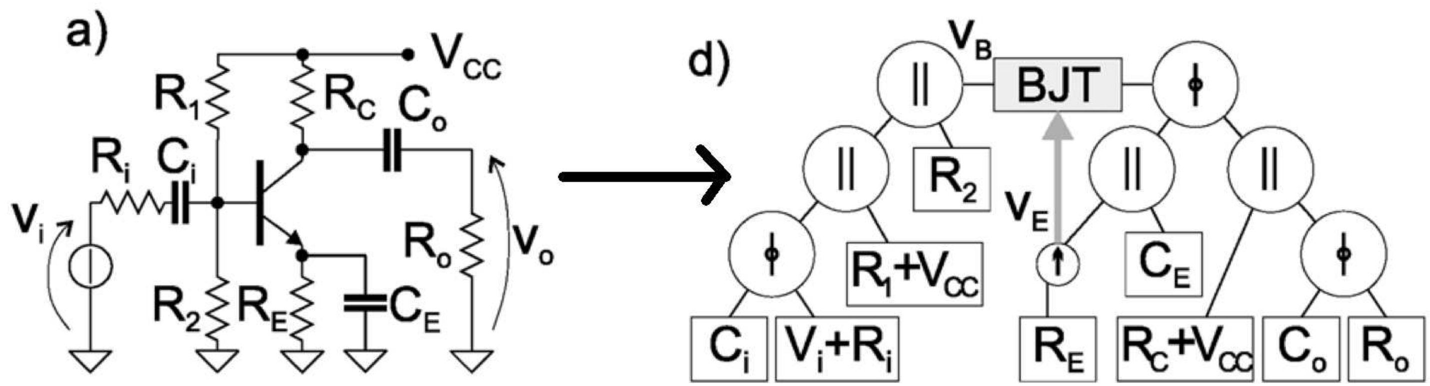


Fig. 4. WDF binary tree for simulation of the triode stage in Fig. 1. The input circuit (C_i, R_i, R_g) is omitted and the cathode voltage V_k is used throught a unit delay to get the grid-to-cathode voltage V_{gk}.

- Grid–Cathode voltage “cross-control” w/ ad-hoc delay to aid realizability.
- Refined in J. Pakarinen and M. Karjalainen, “Enhanced wave digital triode model for real-time tube amplifier emulation,” IEEE Trans. Audio, Speech, Language Process., vol. 18, no. 4, pp. 738–746, May 2010.

WDFs: Cross-Control Multiport II

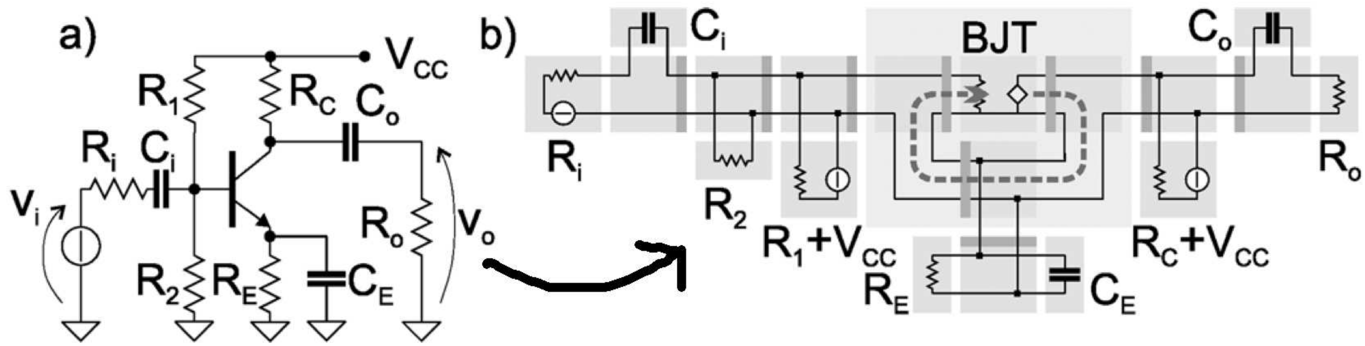
Reference: G. De Sanctis and A. Sarti, "Virtual analog modeling in the wave-digital domain," IEEE Trans. Audio, Speech, Language Process., vol. 18, no. 4, pp. 715–727, 2010.



- Frame BJT as two-port nonlinear element
- Two linear WDF subtrees
- Emitter voltage V_E proposed as "cross-control" on BJT

WDFs: Linearized

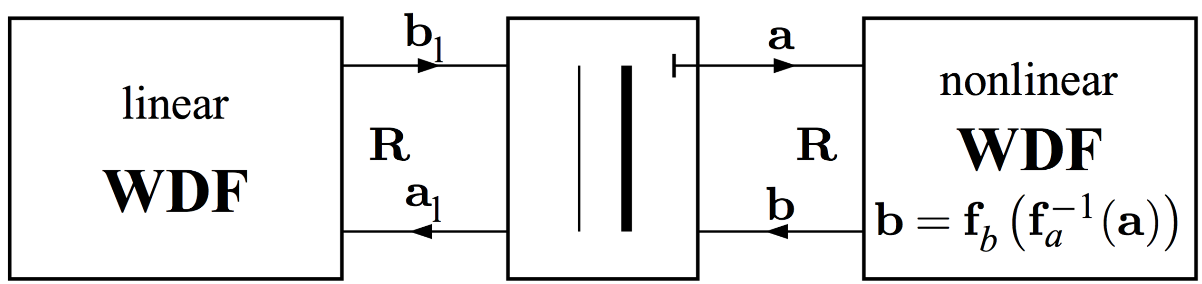
Reference: G. De Sanctis and A. Sarti, "Virtual analog modeling in the wave-digital domain," IEEE Trans. Audio, Speech, Language Process., vol. 18, no. 4, pp. 715–727, 2010.



- Proposes linearized Hybrid-π model of BJT
- Three linear WDF subtrees

WDFs: Multiple NLE, Piecewise Linear Model I

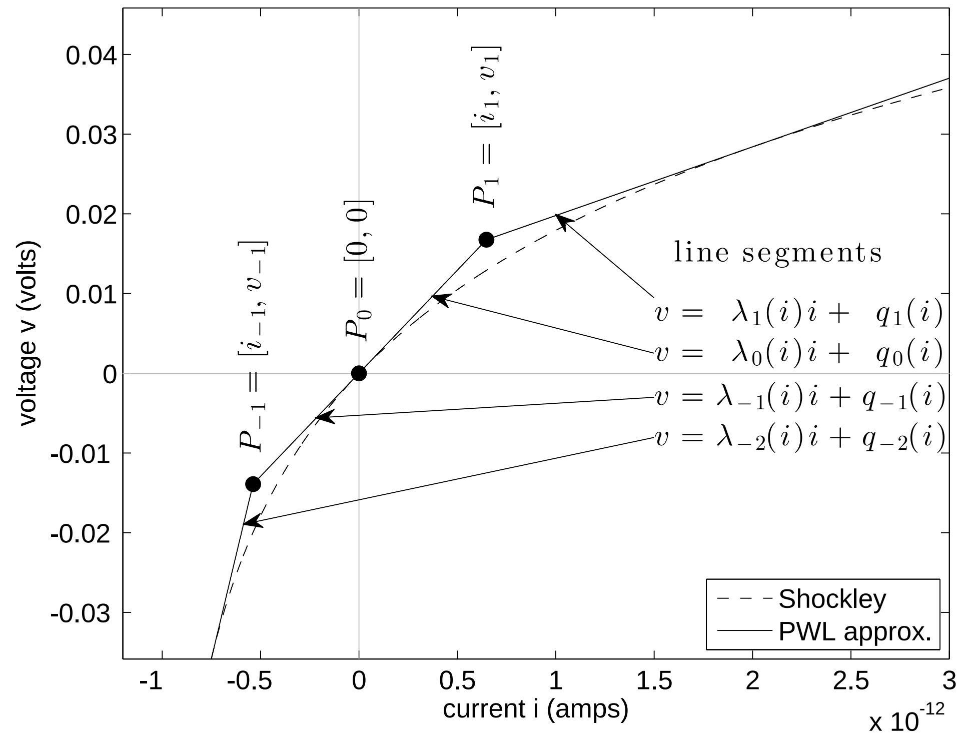
Reference: S. Petrausch, R. Rabenstein, “Wave digital filters with multiple nonlinearities,” Proc. European Signal Process. Conf. (EUSIPCO), vol. 12. Vienna, Austria, Sept. 2004.



- Represent vector of nonlinear root elements with piecewise linear approximations
- Limited to vector parallel relationship between “internal” ($\mathbf{a}$ and $\mathbf{b}$) and “external” ($\mathbf{a}_l$ and $\mathbf{b}_l$) root ports

WDFs: Piecewise Linear Model II

Reference: A. Bernardini, K. J. Werner, A. Sarti, and J. O. Smith III, "Multi-Port NonLinearities in Wave Digital Structures," IEEE Int. Symp. Signals, Circuits, Syst. (ISSCS), Iasi, Romania, July 9–10 2015.



Addresses Petrausch & Rabenstein limit to vector parallel case for multiple one-port nonlinearities.

WDFs: Iterative Schemes I

Reference: S. D'Angelo, J. Pakarinen, and V. Välimäki, "New Family of Wave-Digital Triode Models," IEEE Trans. Audio, Speech, Language Process., vol. 21, no. 2, pp. 313–321, 2013.

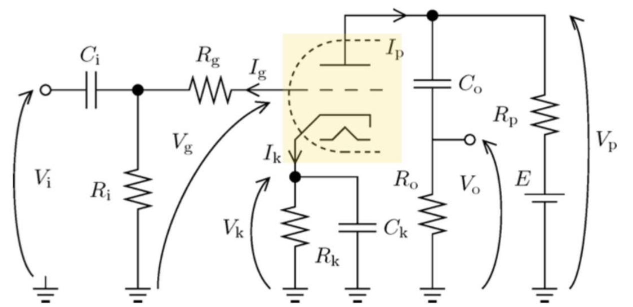


Fig. 2. The common-cathode triode gain stage, typically found in tube amplifiers.

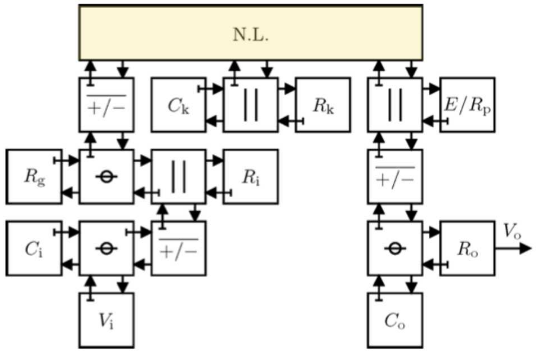


Fig. 4. Implementation of new WDF simulators. The same structure is used in both cases (w/o and w/ grid current).

- Entire triode nonlinearity contained in root element
- Three linear WDF subtrees
- Root solved with customized secant method (specific to triode model)

WDFs: Iterative Schemes II

Reference: T. Schwerdtfeger and A. Kummert, "A Multidimensional Approach to Wave Digital Filters with Multiple Nonlinearities," Proc. European Signal Process. Conf. (EUSIPCO), Lisbon, Portugal, Sept. 1–5 2014.

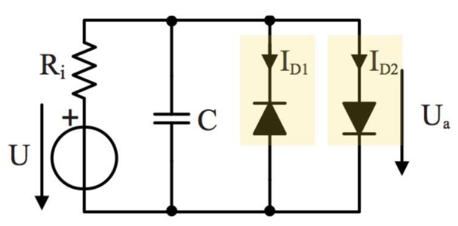
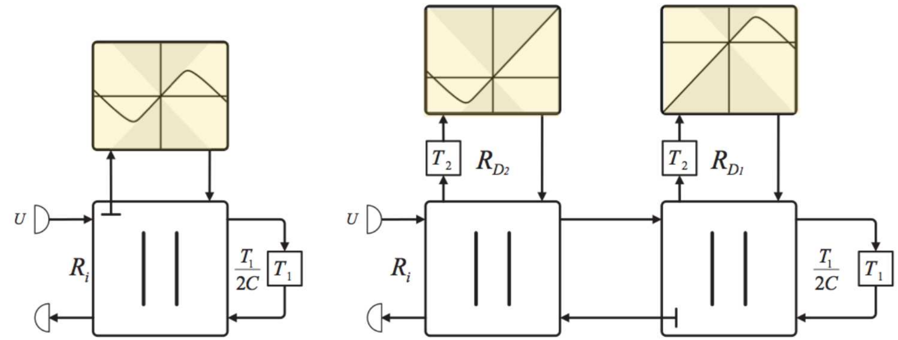


Fig. 1. Prototype circuit



(a) Reference WDF (b) Lumped multidimensional WDF

- Multiple nonlinearities create delay-free loops
- Resolved by inserting extra delay elements as second time dimension (T_2)
- Framed as extension to multidimensional case
- T_2 solved by iteration

WDFs: Iterative Schemes III

Reference: T. Schwerdtfeger and A. Kummert, "A multidimensional signal processing approach to wave digital filters with topology-related delay-free loops," Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Florence, Italy, May 4–9 2014, pp. 389–393.

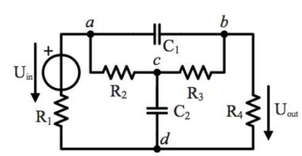
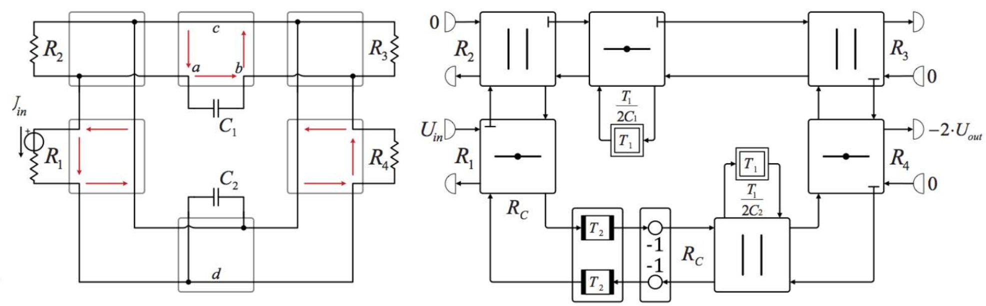


Fig. 1. Prototype circuit with bridged-T topology



(a) Rewritten prototype circuit

(b) WDF realization

Same technique applied to topological problems in linear circuits (bridged-T).

WDFs: Topologically Problematic circuits

- Any circuit that can't be decomposed completely into series and parallel connections
- This includes some linear circuits
- Especially problematic for circuits with multiple or multiport nonlinearities
- Especially problematic for circuits with controlled sources (VCVS, VCCS, CCCS, CCVS, op-amps, etc.) in feedback configuration
- Issues arise in deceptively simple circuits (simple amplifiers, bridged-T, tone stack)

WDFs: Topological Insights

Reference: D. Fränken, J. Ochs, K. Ochs, “Generation of wave digital structures for connection networks containing mutiport elements,” IEEE Trans. Circuits Syst. I: Reg. Papers, vol. 52, no. 3, pp. 586–596, 2005.

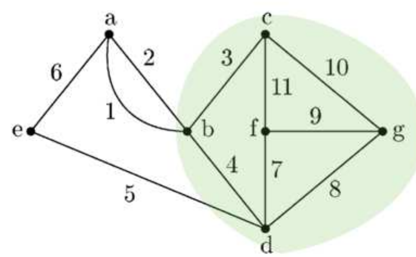


Fig. 7. Biconnected graph.

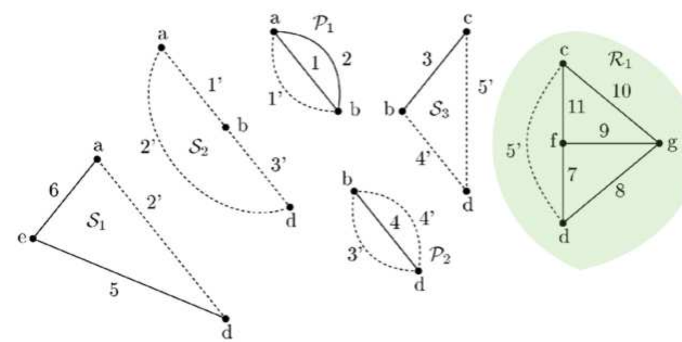
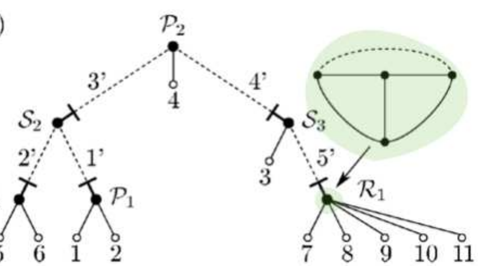


Fig. 8. Split components of the graph in Fig. 7.

(a)



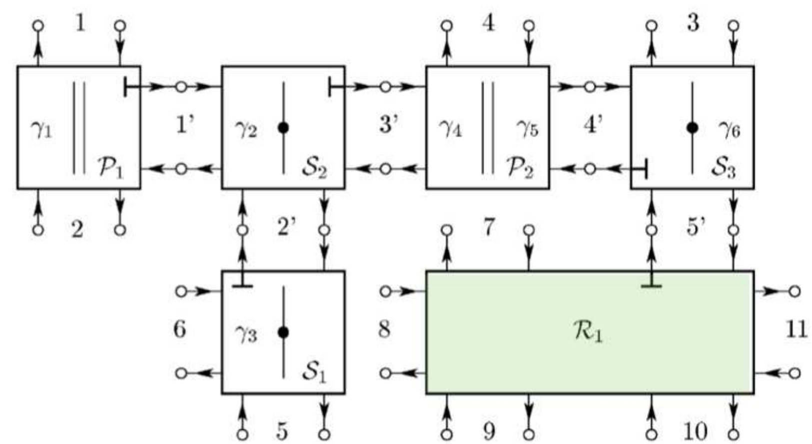
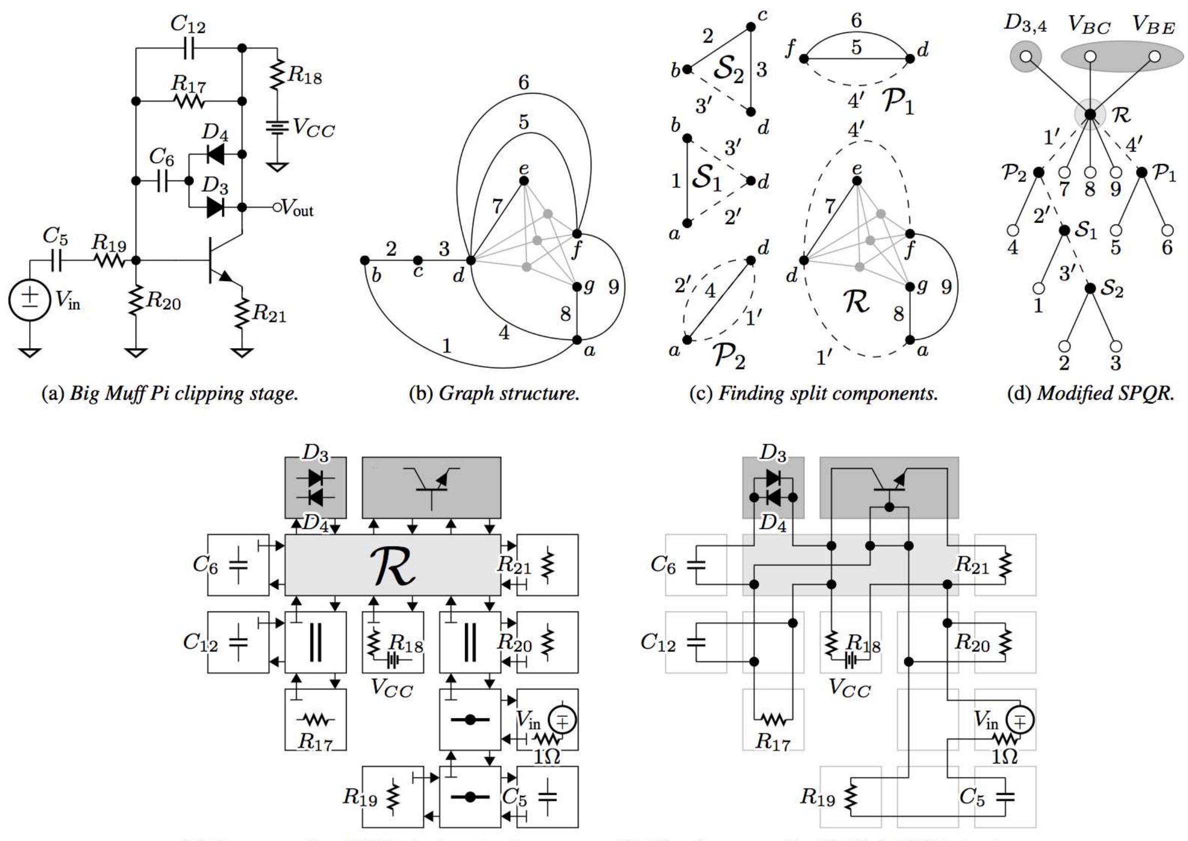


Fig. 10. Adaptor structure for the graph in Fig. 7 as determined by the generalized SPQR-tree in Fig. 9(a).

WDFs: New Use of SPQR Tree

Reference: K. J. Werner, V. Nangia, J. O. Smith III, and J. S. Abel, “A General and Explicit Formulation for Wave Digital Filters with Multiple/Multiport Nonlinearities and Complicated Topologies,” Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, NY, Oct. 18–21.



(a) *Big Muff Pi clipping stage.* (b) *Graph structure.* (c) *Finding split components.* (d) *Modified SPQR.*

(e) *Corresponding WDF adaptor structure.* (f) *Circuit arranged to highlight WDF structure.*

Summary

Summary

1. Hardware advances enable efficient and complex implementations of older theory
2. Software advances ease implementation effort and portability
3. Theory advances continue to improve models and render elegant formulations (WDF) applicable to a wider class of reference systems

Thank you for listening! :)

http://ccrma.stanford.edu/~jos/pdf/RTSP-2015.pdf