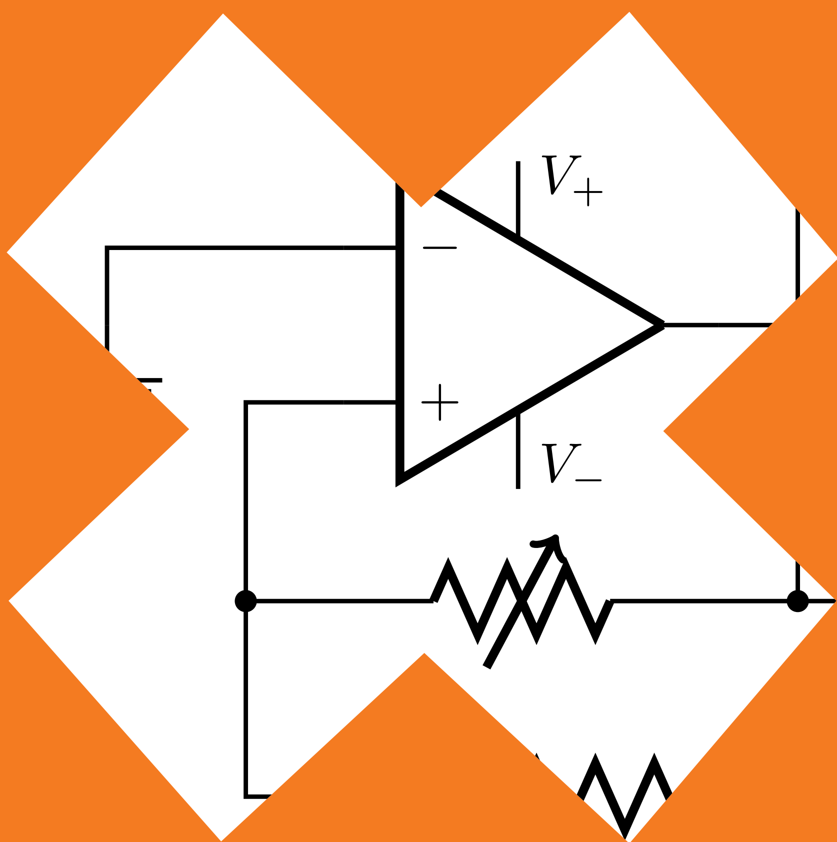
Department of Signal Processing and Acoustics

Filter-Based Oscillator Algorithms for Virtual Analog Synthesis

Jussi Pekonen



Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 26/2014

Filter-Based Oscillator Algorithms for Virtual Analog Synthesis

Jussi Pekonen

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Electrical Engineering, at a public examination held at the lecture hall S1 of the school on 4 April 2014 at 12.

Aalto University
School of Electrical Engineering
Department of Signal Processing and Acoustics

Supervising professor
Prof. Vesa Välimäki

Preliminary examiners
Dr. Stéphan Tassart, STMicroelectronics, France
Dr. Tuomas Virtanen, Tampere University of Technology, Finland

Opponent
Prof. Sylvain Marchant, Université de Bretagne Occidentale, Brest, France

Aalto University publication series
DOCTORAL DISSERTATIONS 26/2014

© Jussi Pekonen

ISBN 978-952-60-5588-6
ISBN 978-952-60-5586-2 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-5586-2

Unigrafia Oy
Helsinki 2014

Finland

NORDIC ECOLABEL
441 697
Printed matter

Aalto University

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Jussi Pekonen

Name of the doctoral dissertation
Filter-Based Oscillator Algorithms for Virtual Analog Synthesis

Publisher School of Electrical Engineering

Unit Department of Signal Processing and Acoustics

Series Aalto University publication series DOCTORAL DISSERTATIONS 26/2014

Field of research audio signal processing

Manuscript submitted 9 September 2013 **Date of the defence** 4 April 2014

Permission to publish granted (date) 20 December 2013 **Language** English

☐ **Monograph** ☒ **Article dissertation (summary + original articles)**

Abstract

This thesis deals with virtual analog synthesis, i.e., the digital modeling of subtractive synthesis principle used in analog synthesizers. In subtractive synthesis, a spectrally rich oscillator signal is modified with a time-varying filter. However, the trivial implementation of the oscillator waveforms typically used in this synthesis method suffers from disturbing aliasing distortion. Filter-based algorithms that produce these waveforms with reduced aliasing are studied in this thesis.

An efficient antialiasing oscillator technique expresses the waveform as a bandlimited impulse train or a sum of time-shifted bandlimited step functions. This thesis proposes new polynomial bandlimited function generators and introduces optimized look-up table and polynomial-based functions for these algorithms. A new technique for generating nonlinear-phase bandlimited functions is also presented.

In addition to the aforementioned technique, the research focus in oscillator algorithms is on ad-hoc approaches that either post-process the output of the trivial oscillator algorithm or produce signals that look similar to the classical waveforms. Linear post-processing algorithms that suppress aliasing of the waveform generated, in principle, by any oscillator algorithm are introduced in this thesis.

Perceptual aspects of the audibility of aliasing are also addressed in this thesis. The results of a listening test that studied the audibility of aliasing distortion in a trivially sampled sawtooth signal are shown. Based on the test results, design criteria for digital oscillator algorithms are obtained and the usability of previously used computational measures for the evaluation of aliasing audibility is analyzed.

In addition, modeling of analog synthesizer oscillator outputs is addressed in this thesis. Two separate models for the sawtooth signal generated by the oscillator circuitry of the MiniMoog Voyager analog synthesizer are developed. The first model uses phase distortion to generate sawtooth waveforms that resemble that of the MiniMoog. The second model filters the output of a digital oscillator algorithm with a fundamental frequency dependent post-processing filter.

The techniques described in this thesis can be used in the development of alias-free oscillator algorithms for virtual analog synthesis. Also, the output of this oscillator can be processed to sound and look like the respective waveform of any analog synthesizer using the methods proposed here.

Keywords audio systems, interpolation, sound synthesis, oscillators, music

ISBN (printed) 978-952-60-5588-6 **ISBN (pdf)** 978-952-60-5586-2

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2014

Pages 165 **urn** http://urn.fi/URN:ISBN:978-952-60-5586-2

A! Aalto-yliopisto

Tiivistelmä

Aalto-yliopisto, PL 11000, 00076 Aalto www.aalto.fi

Tekijä
Jussi Pekonen

Väitöskirjan nimi
Suodattimiin perustuvat oskillaattorialgoritmit virtuaalianalogisynteesissä

Julkaisija Sähkötekniikan korkeakoulu

Yksikkö Signaalinkäsittelyn ja akustiikan laitos

Sarja Aalto University publication series DOCTORAL DISSERTATIONS 26/2014

Tutkimusala äänenkäsittelytekniikka

Käsikirjoituksen pvm 09.09.2013 **Väitöspäivä** 04.04.2014

Julkaisuluvan myöntämispäivä 20.12.2013 **Kieli** Englanti

☐ **Monografia** ☒ **Yhdistelmäväitöskirja (yhteenveto-osa + erillisartikkelit)**

Tiivistelmä

Tämä väitöskirja käsittelee virtuaalianalogisynteesiä eli analogisyntetisaattoreissa käytetyn vähentävän synteesin toimintamallin digitaalista mallintamista. Vähentävässä synteesissä spektriltään rikasta oskillaattorisignaalia muokataan aikamuuttuvalla suodattimella. Tässä menetelmässä tyypillisesti käytetyttyjen lähdesignaalien triviaali digitaalinen toteutus tuottaa häiritsevää laskostumista signaaliin. Tässä työssä tutkitaan suodattimiin perustuvia algoritmeja, joilla voidaan generoida lähdesignaaleja, joissa laskostumista on vaimennettu.

Eräs tehokas menetelmä ilmaisee oskillaattorin tuottaman aaltomuodon kaistarajoitettujen impulssien jonona tai ajassa siirrettyjen kaistarajoitettujen askelfunktioiden summana. Tässä työssä esitetään uusia polynomeihin pohjautuvia kaistarajoitettuja funktiogeneraattoreita ja optimoidaan sekä taulukko- että polynomipohjaisia funktioita kyseiselle menetelmälle. Lisäksi työssä esitellään uusi tapa luoda epälineaarivaiheisia kaistarajoitettuja funktioita.

Edellä mainitun menetelmän lisäksi viime aikoina oskillaattorialgoritmien tutkimuskohteena ovat olleet niin sanotut ad-hoc-algoritmit, jotka joko jatkokäsittelevät triviaalin oskillaattorin ulostuloa tai tuottavat signaaleja, jotka muistuttavat klassisia aaltomuotoja. Väitöskirjassa esitellään lineaarisia jälkikäsittelyalgoritmeja, joilla voi vähentää laskostumista periaatteessa minkä tahansa oskillaattorialgoritmin tuottamasta signaalista.

Työssä käsitellään myös laskostumisen kuulumista psykoakustiikan näkökulmasta esittelemällä tuloksia kuuntelukokeesta, joka tutki laskostumisen havaitsemista triviaalisti generoidun saha-aallon tapauksessa. Kokeen tuloksista saadaan suunnittelukriteerit digitaalisille oskillaattorialgoritmille, ja tuloksia verrataan aiemmin käytettyihin laskennallisiin mittoihin, joilla on arvioitu laskostumisen kuulemisesta.

Lisäksi tässä väitöskirjassa käsitellään MiniMoog Voyager -analogisyntetisaattorin oskillaattorin tuottaman saha-aallon digitaalista mallintamista. Malleista ensimmäinen käyttää vaihesärömenetelmää luomaan saha-aaltoa, joka näyttää samalta kuin MiniMoogin saha-aalto. Toinen malli suodattaa digitaalisen oskillaattorin tuottamaa saha-aaltoa perustaajuudesta riippuvalla suodattimella.

Työssä esitettyjä menetelmiä voidaan hyödyntää laskostumisvapaan oskillaattorialgoritmin kehittämisessä virtuaalianalogisynteesiä varten. Tämän oskillaattorin lähtösignaalia voidaan myös muokata vastaamaan analogisyntetisaattorin tuottamaa aaltomuotoa väitöskirjassa ehdotettujen menetelmien avulla.

Avainsanat audiojärjestelmät, interpolaatio, äänisynteesi, oskillaattorit, musiikki

ISBN (painettu) 978-952-60-5588-6 **ISBN (pdf)** 978-952-60-5586-2

ISSN-L 1799-4934 **ISSN (painettu)** 1799-4934 **ISSN (pdf)** 1799-4942

Julkaisupaikka Helsinki **Painopaikka** Helsinki **Vuosi** 2014

Sivumäärä 165 **urn** http://urn.fi/URN:ISBN:978-952-60-5586-2

To M.A.K. and H.H.P.

Preface

Monday May 17, 2004. That was the day I joined the TKK Acoustics lab. I had just completed my first year of studies at the university, but still I was considered to be a qualified research assistant. There were two persons who made that decision, and both of them have been, and still are, inspiring guides. One of them was Prof. Vesa Välimäki, my supervisor, to whom I want to express my gratitude. He has supported and instructed me to find the thing I would love to do. The other one was Prof. Matti Karjalainen (in memoriam 1946–2010) whose enthusiasm towards making funny things a hardcore science has been the driving motivation in my work.

In addition to these two, I wish to thank my co-authors (in the order of appearence in the list of publications) Mr. Juhan Nam, Prof. Julius O. Smith III, Prof. Jonathan S. Abel, Dr-Ing. Martin Holters, Dr. Heidi-Maria Lehtonen, Dr. Joseph Timoney, Dr. Victor Lazzarini, and Dr. Jari Kleimola without whom this thesis work would have been only half-ready. I would like to thank my pre-examiners, Dr. Stephan Tassart and Dr. Tuomas Virtanen for their valuable comments. I am grateful to Lic.Sc. (Tech.) Luis Costa for the careful proofreads of my manuscripts, including this thesis.

Obviously, this work wouldn't have been possible to do without someone supporting it financially. This thesis was supported by the Academy of Finland (project number 122815), Graduate School of Electrical and Communications Engineering, and Tekniikan Edistämissäätiö.

In addition, I would like to thank Mark Smart for the permission to use his photo of the Moog Modular synthesizer in this thesis. I am grateful to Arne Barlindhaug Ellingsen and Toman Johansson from Clavia DMI AB for the photo of the original Clavia Nordlead.

During the seven and half years I worked at the Acoustics lab in its various incarnations, I met incredibly fantastic people. There have been a large number of people working in the lab during these years, and it is obvious that I cannot remember to thank all of you. However, special thanks go to Paavo, Hynde, Unski, Kalle, Vile, and Jouni, the ones who have been there through all these years. I would like to thank also “the old gang”, Henkka, Jykke, Heikku, Hannu, Cumi, Mairassi, Balázs, Tomppa, Laura E., Tontsa, Patty, Hanna, Pete, Laura S. (née L.), Carlo (in memoriam 1980–2008), Juha, Tuomas, Mara, Jukkis, Lea, Miikka, Toomas, and David, for introducing me the lab spirit. Of the younger generation(s) of labsters (current or former), I would like to thank Mauno, Seppo, Javier, Rafael, Magge, Antti, Mikkis, Digiänkyrä, Julian, Tapsa, Tuomo, Okko, Olli, Symeon, Akis, Ville (the other one), Jussi, Henna, Marko, Juha, and Teemu for being fast in learning the standard procedures in making fun of each other. I would also like to thank Lauri and Tapio, who I consider to be members of the lab even though they work elsewhere.

I would be rude if I wouldn’t thank the DAFx community, which has organized one of the nicest and welcoming conferences I have ever been able to take part in. I have met so many great people through DAFx, and I wish I could one day meet you all again.

After I left the Acoustics lab, I joined the wonderful MRF community at Ericsson. I would like to thank my colleagues for all the fun times we have had.

I am grateful for the love and support I have gotten from my parents, Kaarina and Heikki. I would also want to thank my sister Virpi and my brother Kari and their families for their support.

Last, but definitely not the least, I would like to thank my wonderful wife, Susanna, for all the love. And our children, Kaisa, Maija, Arri, and Ilkka, there are not enough words in this world to describe how much I love you.

Espoo, February 26, 2014,

Jussi Pekonen

Contents

List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I V. Välimäki, J. Pekonen, and J. Nam. Perceptually informed synthesis of bandlimited classical waveforms using integrated polynomial interpolation. *Journal of the Acoustical Society of America*, Special issue on Musical Acoustics, vol. 131, no. 1, part 2, pp. 974–986, January 2012.

II J. Pekonen, J. Nam, J. O. Smith, J. S. Abel, and V. Välimäki. On minimizing the look-up table size in quasi bandlimited classical waveform oscillators. In *Proceedings of the 13th International Conference on Digital Audio Effects (DAFx-10)*, pp. 57–64, Graz, Austria, September 2010.

III J. Pekonen, J. Nam, J. O. Smith, and V. Välimäki. Optimized polynomial spline basis function design for quasi-bandlimited classical waveform synthesis. *IEEE Signal Processing Letters*, vol. 19, no. 3, pp. 159–163, March 2012.

IV J. Pekonen and M. Holters. Nonlinear-phase basis function generators for quasi-bandlimited waveform synthesis. Accepted for publication in *Journal of the Audio Engineering Society*, 2014.

V J. Pekonen and V. Välimäki. Filter-based alias reduction for digital classical waveform synthesis. In *Proceedings of the 2008 IEEE International Conference on Audio, Speech, and Language Processing (ICASSP'08)*, pp. 133–136, Las Vegas, NV, USA, April 2008.

VI H.-M. Lehtonen, J. Pekonen, and V. Välimäki. Audibility of aliasing distortion in sawtooth signals and its implications to oscillator algorithm design. *Journal of the Acoustical Society of America*, vol. 132, no. 4, pp. 2721–2733, October 2012.

VII J. Pekonen, V. Lazzarini, J. Timoney, J. Kleimola, and V. Välimäki. Discrete-time modelling of the Moog sawtooth oscillator waveform. *EURASIP Journal on Advances in Signal Processing*, Special issue on Musical Applications of Real-Time Signal Processing, vol. 2011, Article ID 785103, 15 pages, 2011.

Author's Contribution

Publication I: "Perceptually informed synthesis of bandlimited classical waveforms using integrated polynomial interpolation"

The author of this thesis developed the oscillator algorithms in collaboration with the first and third author. The author was responsible for making the aliasing analysis of the presented algorithms. The author was the sole author of Sections 2 and 4 apart from Figures 1 and 2, which were designed in collaboration with the first author, and contributed in writing Sections 1, 3, and 5 of the article.

Publication II: "On minimizing the look-up table size in quasi bandlimited classical waveform oscillators"

The present author conducted the research presented in this article together with the second author. The author coordinated the writing of the article, wrote Sections 2, 3, and 4.1, and contributed in writing Sections 1, 4.2, and 5.

Publication III: "Optimized polynomial spline basis function design for quasi-bandlimited classical waveform synthesis"

The author of this thesis conducted the research together with the second author. The author wrote the article almost completely and edited the article based on the comments of the other authors.

Publication IV: "Nonlinear-phase basis function generators for quasi-bandlimited waveform synthesis"

The presented algorithm was developed in collaboration by the present author and the second author. The author wrote Sections 0, 1, 2, and 4, and contributed in writing Section 3 and the Appendix.

Publication V: "Filter-based alias reduction for digital classical waveform synthesis"

The author of this thesis developed the algorithms and conducted the aliasing analysis together with the second author. The author wrote the article and edited it based on the comments of the second author.

Publication VI: "Audibility of aliasing distortion in sawtooth signals and its implications to oscillator algorithm design"

The author of this thesis designed and ran the listening experiment together with the first author. The author wrote Sections III.B, IV.B, and V.E and contributed in writing Sections I, II, and VI of the article.

Publication VII: "Discrete-time modelling of the Moog sawtooth oscillator waveform"

The present author conducted the parameter estimation for both presented approaches using sound examples recorded by the second author. The author coordinated the writing of the article and contributed in writing all parts of the paper.

List of Abbreviations

BLEP	Bandlimited step function (sequence)
BLIT	Bandlimited impulse train
DC	Direct current
DPW	Differentiated polynomial waveform
DSF	Discrete summation formulae
FIR	Finite impulse response
FM	Frequency modulation
IIR	Infinite impulse response
NMR	Noise-to-mask ratio
PD	Phase distortion
PEAQ	Perceptual evaluation of audio quality

List of Symbols

d		Fractional delay (value)
f	[Hz]	Frequency variable
f_0	[Hz]	Fundamental frequency, $f_0 = 1/T_0$
f_c	[Hz]	Cut-off frequency
f_s	[Hz]	Sampling frequency, sample rate, $f_\mathrm{s} = 1/T$
k		Indexing variable, $\in \mathbb{Z}$
n		Discrete-time variable, sample index, $\in \mathbb{Z}$
$p(n)$		Phasor signal (discrete-time), $\in [0,1[$
P		Duty cycle, pulse width, $\in [0,1]$
$r(\cdot;\cdot)$		Rectangular pulse wave
$s(\cdot)$		Sawtooth wave
$s_\mathrm{t}(\cdot;\cdot)$		(Asymmetric) Triangular wave
t	[s]	Continuous-time variable
T	[s]	Sampling interval, $T = 1/f_\mathrm{s}$
T_0	[s]	Oscillation period, $T_0 = 1/f_0$
$u(\cdot)$		Heaviside unit step function
$\mathbb{Z}$		Set of integers, $0, \pm 1, \pm 2, \ldots$
τ		Dummy variable
$\phi(t)$		Phasor signal (continuous-time), $\in [0, 2\pi[$
ϕ_0		Initial phase
$\varphi(t)$		Normalized continuous-time phasor signal, $\in [0,1[$

1. Introduction

The history of electronic music dates back to the late 19th century. The first well-known electronic synthesizer was the Telharmonium [Cahill, 1897]. However, the Telharmonium was a huge system that required a roomful of electronics. The first electronic synthesizers that were portable by a man were constructed in the early 20th century. The development of electronic synthesizers truly began between the World Wars when sound-generation units like the Theremin [Théremin, 1925], the Ondes Martenot [Martenot, 1931], and the Hammond organ [Hammond, 1934] were introduced.

These early electronic synthesizers build up the complex tone using primitive waveforms, like sinusoids and triangular waves, without any modification. In the 1950s, and especially in the 1960s and 1970s, many electronic synthesizers added a modifying filter to the sound production chain. This sound generation technique, called subtractive synthesis, starts with a spectrally rich source signal that is shaped with a time-varying, typically a lowpass-type, filter.

Especially the subtractive synthesizers built by the Moog Music company, one of which is shown in Fig. 1.1, were popular during the 1960s and 1970s. For instance, one of the most sold *classical music* records of that era, "Switched-on Bach" by Wendy Carlos[1], contained compositions of J. S. Bach played on the Moog Modular synthesizer. In addition to music productions, analog synthesizers started to gain popularity also in movie studios in the 1960s and 1970s, and the first popular movie whose film score contained parts played on a Moog synthesizer was the James Bond film "On Her Majesty's Secret Service" from 1969.[2]

In the 1980s and 1990s subtractive synthesis appeared to become a rare sound

[1]This record received three Grammy awards in 1969, including the best classical record award. Source: `http://www.grammy.com/nominees/search?artist=&field_nominee_work_value=Switched-On+Bach&year=1969&genre=All` (last viewed on February 26, 2014).

[2]Sources: `http://en.wikipedia.org/wiki/Analog_synthesizer` and `http://en.wikipedia.org/wiki/John_Barry_(composer)` (last viewed on February 26, 2014)

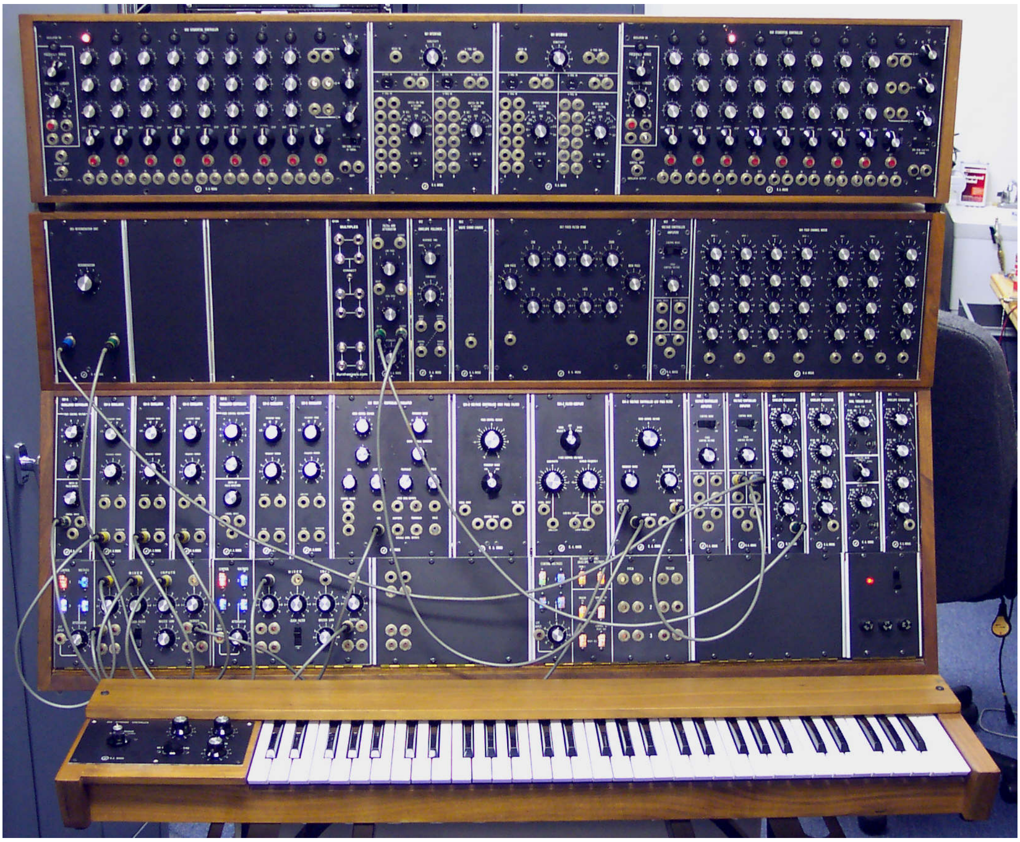


Figure 1.1. Moog Modular synthesizer. Photo copyright by Mark Smart. Used with permission.

generation principle as frequency modulation (FM) synthesis and sampling synthesis gained popularity. At the same time, digital signal processing started to overtake the analog electronics as the technological paradigm also in music synthesizers. However, in the mid-1990s musicians rediscovered the “warm” timbre of the analog synthesizers.

To meet the increased interest in subtractive synthesis, a Swedish music technology company Clavia introduced the NordLead synthesizer (shown in Fig. 1.2) in 1995. That synthesizer was the very first digital synthesizer that emulated the complete sound generation chain of analog synthesizers using digital signal processing tools [Smith, 1996; Välimäki et al., 2006; Erkut et al., 2008], though some of the features of analog synthesizers were modeled earlier in Roland’s JD-series of synthesizers.

Together with the NordLead synthesizer, Clavia coined the term “virtual analog”. It represents the digital simulation of analog audio devices [Smith, 1996; Välimäki et al., 2006; Erkut et al., 2008]. Since 1995, research on virtual analog synthesis, i.e., digital emulation of subtractive synthesis, has increased

Figure 1.2. Original Clavia NordLead synthesizer from 1995. Photo copyright by Clavia DMI AB. Used with permission.

both in academia and in music technology companies such as Yamaha, Korg, Roland, Native Instruments, Access, and Arturia. Nowadays, both hardware and software implementations are available from the aforementioned and many other companies.

Furthermore, more and more interest has been shown on the topic in the past few years [Pekonen and Välimäki, 2011]. A special focus in the research has been on source signal generation, i.e., oscillator algorithms, a subtopic that has justly been studied due to the aliasing issue in the generation of the traditionally used source signals.

Almost all oscillator algorithm studies have focused on finding a method that does not produce audible aliasing. However, the ultimate objective of virtual analog synthesis is to produce a faithful digital representation of the signal generated by an analog synthesizer. It should be noted that these two objectives are not mutually exclusive. The antialiasing oscillator algorithms can be used in modeling of an analog synthesizer waveforms, as will be illustrated in this thesis.

1.1 Scope and content of this thesis

This thesis presents the recent development of the oscillator algorithms used in virtual analog synthesis with a special focus on the advances in time-varying filter-based approaches. The thesis consists of a summary and seven articles that have been published in or accepted for publication in international, peer-reviewed journals or scientific conferences.

The summary part of the thesis first presents the traditionally used source waveforms and the aliasing problem in trivial oscillator algorithms in Section 2. Section 3 provides an overview of antialiasing oscillator algorithms, the topic of Publications I–V. Section 4 discusses the audibility of aliasing distortion, a topic that was investigated in more detail in Publication VI. Modeling of the waveforms of analog synthesizers, a topic that was addressed for the first time in Publication VII, is presented in Section 5. Finally, Section 6 summarizes the main results of the thesis, and Section 7 concludes the thesis and discusses directions for future research on the topic.

2. Trivial Oscillator Algorithms

The waveforms that are typically used in subtractive synthesis [Olson et al., 1955; Moog, 1964] are depicted in Fig. 2.1. These waveforms, which originate from the function generators used for the analysis of analog circuits, are composed of piece-wise linear or constant segments. Due to the well-defined shape of these waveforms, they can be referred to as geometric waveforms. In addition, they are often called classical waveforms. This name is justified by the fact that they are often used to exemplify the classical analysis tools of signals and systems theory (see for example [Carlson et al., 2002, pp. 25–29]).

2.1 Continuous-time classical geometric waveforms

Because the classical geometric waveforms are periodic, the phase of a waveform can be understood to wrap around whenever a new period begins. The phase signal, the so-called *phasor* signal, can be expressed mathematically as

$$\phi(t) = 2\pi f_0 t - 2\pi \lfloor f_0 t \rfloor = 2\pi f_0 t \mod 2\pi, \tag{2.1}$$

where t is the time (continuous variable) in seconds, f_0 is the (time-varying) fundamental frequency of oscillation in Hertz, and $\lfloor\cdot\rfloor$ denotes the floor function, i.e., rounding to the closest integer *smaller than or equal to* the function argument.

The phasor values given by (2.1) range from zero to 2π. However, it is practical to express the phasor value as a fraction of the period, in which $\phi(t)$ can be normalized to be between zero and one:

$$\varphi(t) = \frac{\phi(t)}{2\pi} = f_0 t - \lfloor f_0 t \rfloor = f_0 t \mod 1. \tag{2.2}$$

The normalized phasor signal $\varphi(t)$ can be efficiently used in the oscillator algorithms, as will be seen shortly.

The sawtooth waveform, plotted in Fig. 2.1(a), is given by

$$s(t) = 2\varphi(t) - 1 = 2f_0 t + 1 - 2\sum_{k=-\infty}^{\infty} u(t - kT_0), \tag{2.3}$$

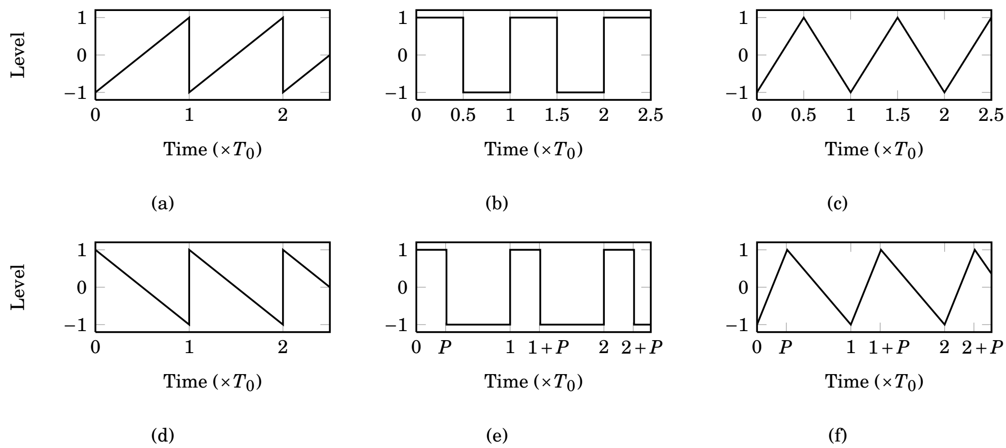


Figure 2.1. Amplitude-normalized classical waveforms typically used in subtractive synthesis: (a) sawtooth, (b) square, and (c) triangular wave. (d) Inverted sawtooth, (e) rectangular pulse, and (f) asymmetric triangular waveforms are also used. P denotes the pulse width, or the duty cycle, of the asymmetric waveforms.

where $T_0 = 1/f_0$ is the oscillation period in seconds, $k \in \mathbb{Z}$ is an indexing variable, and $u(\tau)$ is the Heaviside unit step function [Kreyszig, 1999, pp. 265–266],

$$u(\tau) = \begin{cases} 1, & \text{when } \tau > 0, \text{ and} \\ 0, & \text{when } \tau < 0. \end{cases} \tag{2.4}$$

Note that $u(0)$ is not defined, but it is typically set to 0.5 [Abramowitz and Stegun, 1972, p. 1020]. The inverted sawtooth (see Fig. 2.1(d)) can be obtained by multiplying the expressions of (2.3) by -1. Note that the term $2f_0t + 1$ in the latter form of (2.3) represents the rising ramp of the sawooth waveform, and the sum of the time-shifted Heaviside unit step functions, $u(t - kT_0)$, are responsible for resetting the sample value back to -1 when the phasor wraps around.

The closed-form expression of the square wave, shown in Fig. 2.1(b), is

$$\begin{aligned} r(t) &= \begin{cases} 1, & \text{when } \varphi(t) < 0.5, \\ 0, & \text{when } \varphi(t) = 0.5, \text{ and} \\ -1, & \text{when } \varphi(t) > 0.5, \end{cases} \\ &= \operatorname{sgn}(0.5 - \varphi(t)) = s(t - 0.5T_0) - s(t) \\ &= 2\sum_{k=-\infty}^{\infty} [u(t - kT_0) - u(t - (k + 0.5)T_0)] - 1, \end{aligned} \tag{2.5}$$

where $\operatorname{sgn}(\tau)$ is the signum (sign) function [Carlson et al., 2002, p. 64],

$$\operatorname{sgn}(\tau) = 2u(\tau) - 1 = \begin{cases} 1, & \text{when } \tau > 0, \\ 0, & \text{when } \tau = 0, \text{ and} \\ -1, & \text{when } \tau < 0. \end{cases} \tag{2.6}$$

The first Heaviside unit step functions inside the sum of the last form of (2.5) steps up the amplitude of the waveform and the latter step function steps it down.

The triangular wave (see Fig. 2.1(c)) is given by

$$\begin{aligned} s_{\mathrm{t}}(t) &= \begin{cases} 4\varphi(t)-1, & \text{when } \varphi(t) \leq 0.5, \text{ and} \\ -4\varphi(t)+3, & \text{when } \varphi(t) \geq 0.5, \end{cases} \\ &= 1-2|s(t)| \\ &= 4f_0 \int_{-\infty}^{t} r(\tau)\,\mathrm{d}\tau \\ &= 8f_0 \sum_{k=-\infty}^{\infty} [(t-kT_0)u(t-kT_0) \\ &\qquad - (t-(k+0.5)T_0)u(t-(k+0.5)T_0)] - 4f_0t-1. \end{aligned} \tag{2.7}$$

The first form of (2.7) can be obtained from Fig. 2.1(c) by writing linear functions for the ascending and descending slopes of a single triangle pulse. The second form is obtained by noting that by taking the absolute value of the sawtooth waveform one gets an inverted triangle wave that is between 0 and 1. The third form is obtained by noting that the time derivative of the triangle waveform is a scaled square wave. In the last form of (2.7), the first term inside the sum represents the ascending ramp and the second term the descending ramp of a triangle pulse, and the term $4f_0t-1$ removes the drifting DC offset created by the summation term.

The square and triangular waves given above are symmetric, i.e., they have a *duty cycle*, or *pulse width*, of 50 %. In principle, the triangular and rectangular pulse waves could also be asymmetric, as depicted in Figs. 2.1(e) and (f). The asymmetric rectangular pulse wave, plotted in Fig. 2.1(e), having a duty cycle of $P \in [0,1]$, which can be time-varying, is given by

$$\begin{aligned} r(t;P) &= \begin{cases} 1, & \text{when } \varphi(t) < P, \\ 0, & \text{when } \varphi(t) = P, \text{ and} \\ -1, & \text{when } \varphi(t) > P, \end{cases} \\ &= \operatorname{sgn}(P-\varphi(t)) \\ &= s(t-PT_0)-s(t)+2P-1 \\ &= 2\sum_{k=-\infty}^{\infty} [u(t-kT_0)-u(t-(k+P)T_0)]-1. \end{aligned} \tag{2.8}$$

The asymmetric triangular wave (see Fig. 2.1(f)) is

$$
\begin{aligned}
s_{\mathrm{t}}(t;P) &= \begin{cases} \frac{2\varphi(t)-P}{P}, & \text{when } \varphi(t) \le P, \\ \frac{1+P-2\varphi(t)}{1-P}, & \text{when } \varphi(t) \ge P, \end{cases} \\
&= \frac{1-r(t;P)-2r(t;P)P+4r(t;P)\varphi(t)}{1-r(t;P)+2r(t;P)P} \\
&= \frac{f_0}{P(1-P)} \int_{-\infty}^{t} [r(\tau;P)+1-2P]\,\mathrm{d}\tau \\
&= \frac{2f_0}{P(1-P)} \sum_{k=-\infty}^{\infty} [(t-kT_0)u(t-kT_0) \\
&\qquad - (t-(k+P)T_0)u(t-(k+P)T_0)] - \frac{2f_0 t}{1-P} - 1.
\end{aligned}
\tag{2.9}
$$

The first form of (2.9) is obtained by writing linear functions for the ascending and descending slopes of a single pulse, and second formula combines the two piece-wise definitions into a single formula. The third form of (2.9) is obtained by noting that the time derivative of the asymmetric waveform is a scaled asymmetric rectangular pulse wave with the DC component removed. In the last form of (2.9), the terms inside the sum again represent the ramps of the triangular wave and remaining term removed the drifting DC offset caused by the summation term.

Note that, when $P = 0.5$, (2.8) and (2.9) are equal to (2.5) and (2.7), respectively. When $P = 0$, the asymmetric triangular wave becomes the inverted sawtooth wave and the rectangular pulse wave is equal to -1 for all t. When $P = 1$, the rectangular pulse is a constant $+1$ and the asymmetric triangular wave becomes the sawtooth wave.

2.2 Trivial digital implementations

As the closed-form expressions of the classical waveforms show, they can be constructed from the normalized phasor signal. Therefore, using the phasor signal in the digital generation of the waveforms is efficient. The discrete-time phasor is trivially obtained by sampling the normalized phasor signal

$$p(n) \equiv \varphi(nT) = f_0 nT \mod 1, \tag{2.10}$$

where $n \in \mathbb{Z}$ is the discrete-time variable, i.e., the sample index, and T is the sampling interval in seconds. By examining the difference between the phasor signal values at consecutive sample indices, (2.10) can be rewritten as

$$p(n) = (p(n-1) + f_0 T) \mod 1. \tag{2.11}$$

The trivial digital oscillator algorithms can be constructed by replacing the continuous-time phasor signal with the discrete-time phasor signal in the respective closed-form expressions given above. Block diagrams of the trivial algorithms are shown in Fig. 2.2. Note that Fig. 2.2(a) shows the block diagram for the general asymmetric triangular oscillator. The sawtooth and symmetric triangular oscillators can be implemented more efficiently than with the general algorithm (see (2.3) and (2.7)), and the block diagrams of these efficient implementations are given Figs. 2.2(c) and 2.2(d), respectively.

2.3 Aliasing problem in trivial oscillators

Unfortunately, the trivial digital oscillators suffer from aliasing distortion because the continuous-time waveforms are not bandlimited. This can be verified by deriving the Fourier series representation (see for example [Kreyszig, 1999, pp. 240–242] and [Carlson et al., 2002, pp. 25–26] for the theory) of the continuous-time waveforms.

The Fourier series representation of the rectangular pulse waveform is given by

$$\begin{aligned} r(t;P) &= 1-2P+4\sum_{k=1}^{\infty}\frac{\sin(kP\pi)}{k\pi}\cos(2\pi k f_0 t - kP\pi) \\ &= 1-2P+4P\sum_{k=1}^{\infty}\operatorname{sinc}(kP)\cos(2\pi k f_0 t - kP\pi), \end{aligned} \tag{2.12}$$

where $\operatorname{sinc}(\tau)=\sin(\pi\tau)/(\pi\tau)$ [Carlson et al., 2002, p. 26] is the sinc, i.e., sine cardinal, function. The Fourier series representation of the asymmetric triangular wave is

$$\begin{aligned} s_{\mathrm{t}}(t;P) &= -2\sum_{k=1}^{\infty}\frac{\sin(k(1-P)\pi)}{k^2P(1-P)\pi^2}\sin(2\pi k f_0 t - k(P-1)\pi) \\ &= -\frac{2}{P\pi}\sum_{k=1}^{\infty}\frac{\operatorname{sinc}(k(1-P))}{k}\sin(2\pi k f_0 t - k(P-1)\pi) \\ &= \frac{2}{(1-P)\pi}\sum_{k=1}^{\infty}\frac{\operatorname{sinc}(kP)}{k}\sin(2\pi k f_0 t - kP\pi). \end{aligned} \tag{2.13}$$

Again, when P is one of the special cases given above (0, 0.5, or 1), (2.12) and (2.13) can be simplified. When $P=0.5$, (2.12) becomes

$$r(t;0.5) = 2\sum_{k=1}^{\infty}|\operatorname{sinc}(k/2)|\cos(2\pi k f_0 t) \tag{2.14}$$

and (2.13) simplifies to

$$s_{\mathrm{t}}(t;0.5) = -\frac{4}{\pi}\sum_{k=1}^{\infty}\frac{|\operatorname{sinc}(k/2)|}{k}\cos(2\pi k f_0 t). \tag{2.15}$$

Note that $\operatorname{sinc}(\tau)=0\ \forall\tau\in\{\mathbb{Z}\setminus 0\}$ [Carlson et al., 2002, p. 26], which means that the symmetric rectangular pulse and triangular waves have only odd harmonics.

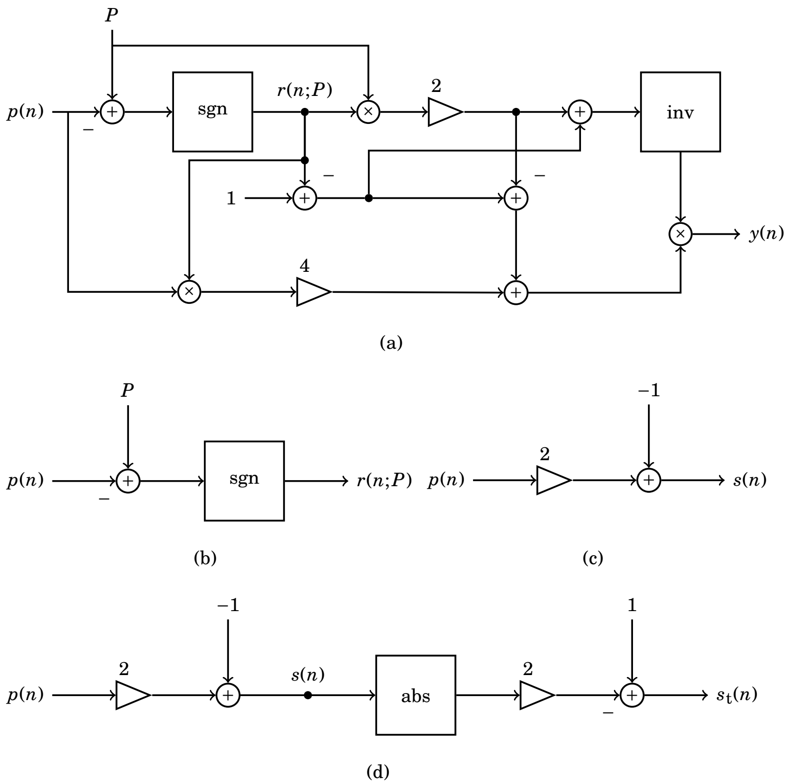


Figure 2.2. Trivial digital algorithms for the classical waveforms: (a) asymmetric triangular, (b) rectangular pulse, (c) sawtooth, and (d) symmetric triangular oscillators. Box "sgn" applies the signum function (see (2.6)) to its input, box "inv" inverts its input, and box "abs" outputs the absolute value of its input.

When $P = 1$, the second expression in (2.13) becomes

$$s_t(t;0) = s(t) = -\frac{2}{\pi}\sum_{k=1}^{\infty}\frac{1}{k}\sin(2\pi k f_0 t) \tag{2.16}$$

because sinc(0) = 1 [Carlson et al., 2002, p. 26]. With $P = 0$, the asymmetric triangular wave becomes the inverted sawtooth, whose Fourier series representation is the same as (2.16) but without the minus sign. This can be verified by substituting $P = 0$ into the last expression of (2.13). Similarly, when $P = 0$ or $P = 1$, the scaling factor $\sin(k\pi P)$ in (2.12) is equal to 0 for all $k \in \mathbb{Z}$, meaning that (2.12) reduces to -1 and 1, respectively.

As the Fourier series representations of the classical geometric waveforms show, the waveforms have infinitely many harmonics. However, in digital systems the frequencies that will be represented faithfully by the samples is limited to the Nyquist limit, which is the half of the sampling frequency $f_s = 1/T$. This means that a component whose frequency is $f_s/2 + f$ will alias (fold back) to

$f_s/2 - f$ [Bateman, 1980, pp. 105–106; Roads, 1998, pp. 27–31] because

$$\cos(2\pi(f_s/2 + f)nT + \phi_0) = \cos(2\pi(f_s/2 - f)nT + \phi_0) \tag{2.17}$$

for all $n \in \mathbb{Z}$ and for any values of the initial phase ϕ_0. This equality can be proved by rewriting the left side of the equation using the angle sum, symmetry, shift, and periodicity properties of the sine and cosine functions.

The spectra of the trivial sawtooth wave, rectangular pulse waves that have a duty cycle of 0.5 and 0.25, and the (symmetric) triangular wave are given in Fig. 2.3 for the fundamental frequency 2.093 kHz and sampling rate 44.1 kHz. The non-aliased components are indicated with crosses, the rest is aliasing distortion.

The effect of the duty cycle on the harmonic structure is visible in Fig. 2.3. The symmetric rectangular pulse and triangular waves are missing the even harmonics (see Figs. 2.3(b) and (d), respectively), and the rectangular pulse wave with duty cycle of 0.25 lacks every fourth harmonic (Fig. 2.3(c)) and it has a direct current (DC) component.

The sawtooth and rectangular pulse waves have a spectral tilt of about 6 dB per octave (component amplitudes $\sim 1/k$), as indicated with the dashed line in Figs. 2.3(a), (b), and (c). The spectral envelope of the (asymmetric) triangular wave decays approximately 12 dB per octave ($\sim 1/k^2$), as shown with the dashed line in Fig. 2.3(d). This means that especially with high fundamental frequencies the trivially generated waveforms will contain relatively strong aliased components. This aliasing distortion can be heard as an annoying noise-like hiss and hum, and/or as beating [Alles, 1980; Moore, 1990; Burk, 2004; Puckette, 2007].

However, human hearing renders some of aliasing distortion inaudible, which effectively means that some aliasing can be allowed. Further information about algorithms that take this into account is given in Section 3, and details about the processes that affect the audibility of aliasing are discussed in Section 4.

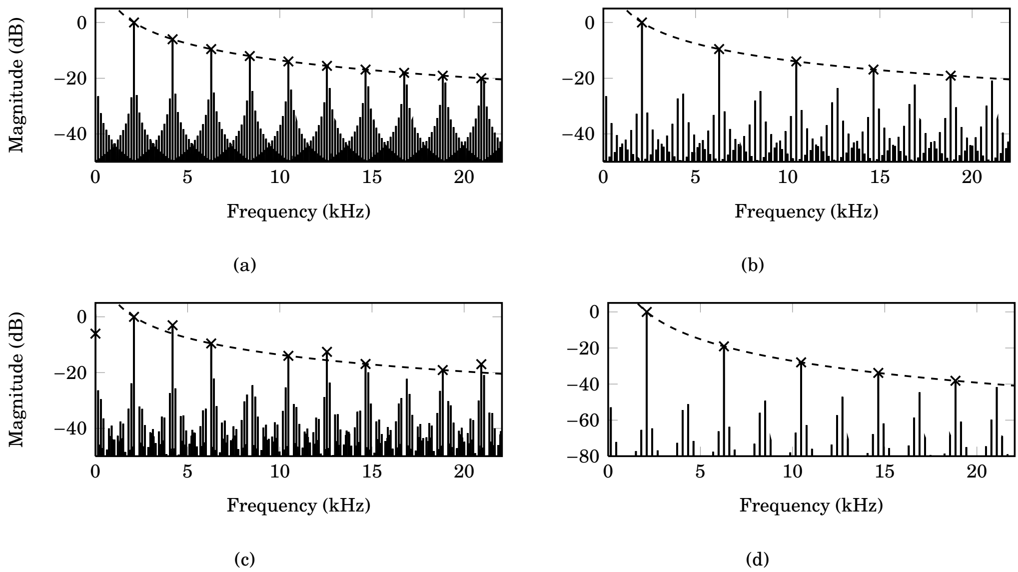


Figure 2.3. Spectra of trivially generated (a) sawtooth wave, (b) rectangular pulse wave with a duty cycle of 0.5, (c) rectangular pulse wave that has duty cycle of 0.25, and (d) symmetric triangular wave for the fundamental frequency 2.093 kHz and sampling rate 44.1 kHz. The non-aliased components are indicated with crosses. The dashed line in (a), (b), and (c) illustrates the approximately 6-dB-per-octave spectral envelope. In (d), the dashed line is the approximately 12-dB-per-octave spectral tilt.

3. Antialiasing Oscillator Algorithms

This section reviews the existing antialiasing oscillator algorithms, which can be categorized into four groups: ideally bandlimited, quasi-bandlimited, alias-suppressing, and ad-hoc oscillator algorithms. Each group has distinct properties that differentiate it from the others, and in the following sections these different features are discussed.

3.1 Ideally bandlimited oscillator algorithms

Some of the existing antialiasing oscillators try to avoid aliasing completely. These algorithms can be understood to generate waveforms that have a finite number of harmonics so that the highest component does not exceed the Nyquist limit. An obvious choice for a such algorithm is the additive synthesis approach [Chaudhary, 1998], where the waveforms are generated by summing individual harmonics below a given cut-off frequency $f_c \in [0, f_s/2]$. In other words, the additive synthesis oscillator computes the first $\lfloor f_c/f_0 \rfloor$ terms of the Fourier series representation of the waveform (see Sec. 2.3).

However, the number of components the additive synthesis oscillator is required to add up is inversely proportional to the fundamental frequency. This means that the computational load of the additive synthesis oscillator heavily depends on the fundamental frequency, and especially at low fundamental frequencies the load may become an issue in applications that require minimal load. This issue can be partly relaxed by ignoring the non-aliased components that are, for example, more than 60 dB softer than the fundamental frequency. This relaxation sets an upper limit for the components to be generated[1].

The issue of the f_0-dependent computational load can be overcome in an alternative way. Because the samples of the waveform require a summation

[1]The actual number of harmonics to be generated is the minimum of the number of non-aliased components and this limit.

of well-defined functions, single periods of the finite sums can be precomputed and stored as tables that are read in a loop to generate the waveform. This algorithm, the wavetable synthesis oscillator [Chamberlin, 1985; Burk, 2004], has a constant computational load at all fundamental frequencies. By reading the look-up table at different rates, the fundamental frequency can be varied from its nominal value [Chamberlin, 1985; Massie, 1998; Burk, 2004].

Although the wavetable synthesis oscillator provides a great computational saving with respect to the additive synthesis oscillator, the memory space required for the look-up tables of the wavetable synthesis oscillator may be an issue in memory-sensitive applications. When all harmonics below a given cut-off frequency are desired to be generated at all fundamental frequencies, the number of look-up tables becomes huge. As with the additive synthesis oscillator, the number of tables can be reduced by ignoring weak components.

Alternatively, a finite sum of harmonically related sinusoids can be expressed as a fraction of two sinusoids both of which depend on the current sample index and the fundamental frequency [Winham and Steiglitz, 1970; Moorer, 1976; Lazzaro and Wawrzynek, 2004]. Note, however, that this "discrete summation formulae" (DSF) approach typically requires a modifying filter to produce the classical waveforms. Furthermore, it can have numerical issues when the term in the denominator of the expression is close to zero.

The bandlimited waveforms can also be generated by synthesizing their spectral content in the frequency domain and applying the inverse (fast) Fourier transform [Rodet and Depalle, 1992] to the synthetic spectrum [Stilson, 2006, p. 212; Deslauriers and Leider, 2009]. However, since the harmonic components of a waveform may not exactly correspond to the frequency bins, the spectrum synthesis requires interpolation. Ideal interpolation is a theoretical idea, and in order to have an implementable algorithm the ideal interpolation needs to be approximated [Deslauriers and Leider, 2009]. However, the interpolation approximation induces noise in the spectrum, which can be understood to be a special kind of aliasing distortion.

From the discussion given above one can conclude that all of the ideally bandlimited oscillator algorithms are a trade-off between sound quality and the computational issues. Therefore, they may not be the ideal solution for every application, but they do offer good reference implementations for the other techniques reviewed next.

3.2 Quasi-bandlimited oscillator algorithms

While the ideally bandlimited oscillator algorithms do not allow any aliasing, quasi-bandlimited oscillator algorithms take some advantage of the psychophysiology of human hearing. Human hearing is known to be insensitive at high frequencies (see Section 4 for more details), which means that the oscillator algorithm can be allowed to produce some aliasing at high frequencies. In other words, a quasi-bandlimited oscillator algorithm can be interpreted to produce a signal that is a lowpass-filtered version of the continuous-time waveform so that the lowpass filter is a realizable filter with a non-infinitesimal transition band.

3.2.1 Fundamentals of the quasi-bandlimited oscillator algorithms

The first quasi-bandlimited oscillator algorithm was suggested by Stilson and Smith [1996a]. They noted that by differentiating a continuous-time classical geometric waveform with respect to time, once in the case of the sawtooth and rectangular pulse wave and twice in the case of the triangular wave, one obtains a sequence of impulse functions. When the time derivative of the waveform is lowpass filtered, each impulse is replaced with the impulse response of the filter. Then, by integrating[2] the obtained bandlimited impulse train (BLIT), an approximately bandlimited waveform is obtained [Stilson and Smith, 1996a; Stilson, 2006].

In the ideal BLIT synthesizer, the impulses are replaced with the impulse response of the ideal lowpass filter, the well-known sinc function. The sinc function can also be called the ideal basis function of the BLIT algorithm. With this ideal basis function the BLIT algorithm becomes in fact an alternative formulation to the DSF expression [Stilson and Smith, 1996a]. However, the sinc function is infinitely long, which means that it needs to be windowed in order to have an implementable realization of the BLIT algorithm [Stilson and Smith, 1996a]. The windowed sinc function is typically stored in a look-up table [Stilson, 2006; Välimäki and Huovilainen, 2007].

Yet, because the fundamental frequency and hence the oscillation period can be arbitrary, the discontinuities are not located at sampling instants. Therefore, the mid-point of the look-up table needs to be shifted in time, which means that the table needs to be oversampled in order to have proper positioning of the bandlimited impulse function for each discontinuity [Stilson and Smith, 1996a; Stilson, 2006]. Furthermore, the look-up table can be interpolated to improve the accuracy of the non-tabulated basis function values.

[2] Again, twice in the case of the triangular wave.

As mentioned above, the BLIT oscillator needs an integrator[3] in order to produce the desired waveform. However, if the BLIT values are slightly off from their infinitely accurate values, the integration leads to a drifting DC offset. A second-order leaky integrator that has a transmission zero at DC can be used to avoid this issue [Brandt, 2001].

Brandt also suggested that the integration can be performed in advance. By integrating the BLIT basis function, or the minimum-phase representation of the look-up table as Brandt originally proposed, an approximation of the bandlimited step function is obtained [Brandt, 2001]. Because the classical waveforms can be constructed by summing time-shifted step functions (see Sec. 2.1), replacing the unit step functions in the respective formulations with the bandlimited step function approximation effectively yields a bandlimited waveform.

At every discontinuity the integrated function is read from the table, and when the end of the table is reached, the table-read process outputs a constant one. This technique, called the bandlimited step function (BLEP) algorithm, leads to an efficient realization that avoids the numerical issues the BLIT algorithm has.

However, Brandt did not provide the ideal basis function for the BLEP algorithm. In Publication I, the closed-form expression for the ideal BLEP basis function, which is the integral of the sinc function, is derived. Like the sinc function, it is also infinitely long and impractical to use as is. Therefore, it also needs to be approximated in order to have an implementable realization. One approach is to apply a window function to the BLEP *residual*, that is the difference between the ideal BLEP basis function and the unit step function. This residual function is then *added* onto the waveform at the discontinuities [Välimäki and Huovilainen, 2007; Leary and Bright, 2009].

The table-based BLIT and BLEP oscillators can be implemented in two ways. The look-up process can be centralized so that the algorithm first reads or interpolates the table values that are used as the coefficients of a finite impulse response (FIR) filter. The FIR filter is then triggered with an impulse that has the same polarity as the discontinuity[4]. Alternatively, the look-up table entries can be decomposed into a polyphase FIR filter structure. In this approach all polyphase branches are triggered with the impulse, and the samples of the bandlimited impulse function are obtained by mixing the outputs of the branches depending on the needed shift [Välimäki and Huovilainen, 2007].

[3]In digital systems, the integrator is in practice implemented as an accumulator.

[4]Here, the polarity of a discontinuity means the direction of the waveform transition: a transition upwards (for instance from -1 to 1 in the rectangular pulse wave) has positive polarity while the transition in the other direction has negative polarity.

3.2.2 Basis function approximations

In order to have a good alias-reduction performance, both the oversampling factor and the effective length of the look-up table need to be quite large (see for example Nam et al. [2010, Sec. IV], and Publication I, Sec. IV.A, for analyses of table-based BLIT and BLEP algorithms, respectively). This can be become an issue as the effective length of the table also affects the highest obtainable fundamental frequency when the implementation does not allow overlapping impulse functions [Pekonen, 2007]. When the algorithm is implemented using only one filter structure, the oscillation period cannot be shorter than the effective length of the table.

Alternative basis function approximations have been developed to overcome the aforementioned issues. These approximations typically have short basis functions but their alias-reduction performance is clearly better than that of the tabulated basis functions of same (effective) length. One of these approximations generates the bandlimited impulses using modified FM [Timoney et al., 2008; Lazzarini and Timoney, 2010b]. Another approach, a feedback delay loop [Nam et al., 2009], generates the BLIT using an infinite loop that delays the previously outputted impulse by the given period. This algorithm, in fact, yields a truly bandlimited impulse train, but the generated signal is slightly inharmonic [Nam et al., 2009]. Recently Tassart proposed an approach where the bandlimited impulse train is obtained by modeling an analog lowpass filter digitally using its state-space representation [Tassart, 2013a]. Generation of nonlinear-phase bandlimited impulse functions was also studied by Rodet [1984] but in the context of granular synthesis.

An efficient way to approximate the ideal basis function in the BLIT oscillator is to use fractional delay filters [Nam et al., 2010]. This approximation approach is motivated by the fact that the ideal BLIT basis function, the sinc function, is also the approximation target in fractional-delay filter design [Nam et al., 2010; Pekonen et al., 2010]. The fractional-delay filter designs used for the approximation include allpass filters [Nam et al., 2010] as well as interpolation polynomials implemented as FIR filters [Frei, 2002; Välimäki and Huovilainen, 2007; Pekonen, 2007; Nam et al., 2010]. Especially the polynomial approach provides a great improvement in alias-reduction performance compared to the table-based approach [Välimäki and Huovilainen, 2007; Nam et al., 2010]. Furthermore, the polynomial-based basis functions remove the need for a look-up table completely because the basis function can be computed while generating the impulses.

In addition to the derivation of the ideal BLEP basis function, Publication I extends the polynomial-based basis function design to BLEP residuals. The BLEP basis function approximations are constructed by integrating interpolation polynomials in closed form, yielding basis functions that can be evaluated during the waveform synthesis. The proposed polynomial basis functions provide intuitive control of the alias-reduction performance in terms of polynomial length and order, and the best basis functions presented in Publication I are shown to provide excellent alias-reduction performance.

In Publication I, the fourth-order B-spline polynomial BLEP oscillator is found to provide the best alias-reduction performance based on computational measures (see Section 4.3 for more details). It is shown to produce a sawtooth waveform that is perceptually alias-free in the range of fundamental frequencies typically used in music. Moreover, it is computationally efficient to implement, as one can observe from its Farrow-like [Farrow, 1988; Välimäki, 1995; Franck, 2008] block diagram given in Fig. 3.1.

In Publication II, the performance of table-based quasi-bandlimited oscillators are shown to be improved by tabulating a function other than the sinc function. When the effective length of the look-up table is short, the windowed sinc function is not the best option to be tabulated in the BLIT oscillator. Instead, one can use a controllable window function as proposed in Publication II, Sec. 4.1. Alternatively, one can also optimize the look-up table entries using a perceptually informed optimization problem to reduce aliasing distortion (see Publication II, Sec. 4.2).

Both the controllable window functions and the optimization procedure presented in Publication II yielded bell-shaped look-up tables[5]. An example of a such look-up table is shown in Fig. 3.2 where the weighted least-squares-optimized basis-function table and the sawtooth signal generated with the BLIT algorithm using the optimized table and the second-order leaky integrator proposed by Brandt [2001] are plotted.

Publication III extends the optimization procedure introduced in Publication II to polynomial basis-function designs. By optimizing the polynomial coefficients so that a perceptually informed measure of aliasing distortion is minimized, the alias-free range of fundamental frequencies of the BLIT generator and the BLEP sawtooth oscillator are shown to be extended to be well above the range of fundamental frequencies used in music.

---

[5] Plots of the basis function tables exemplified in Publication II can be found at its companion page `http://www.acoustics.hut.fi/publications/papers/dafx10-optosctables/`.

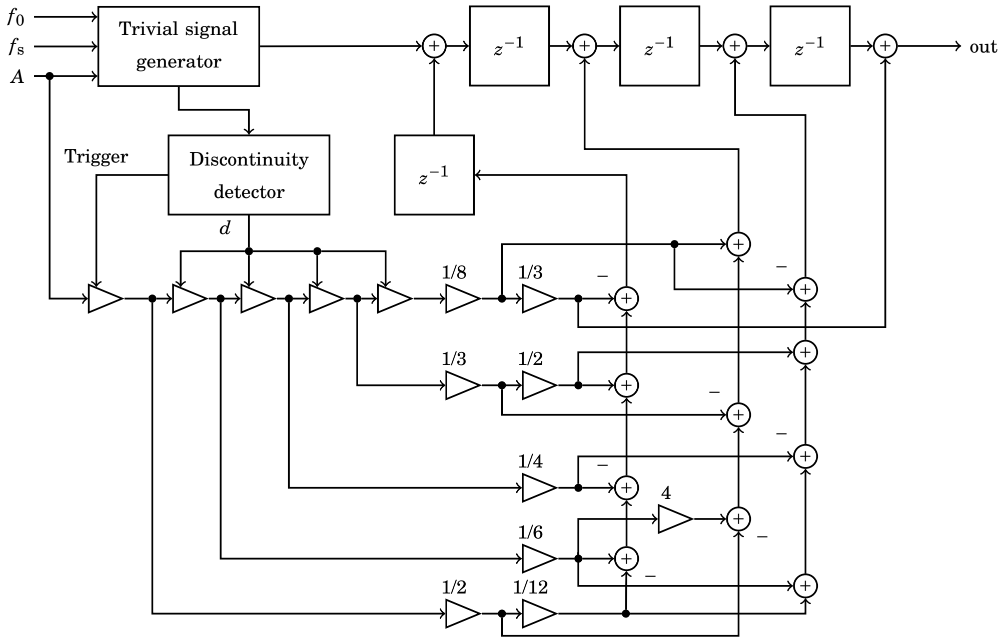


Figure 3.1. Block diagram of the fourth-order B-spline BLEP sawtooth or rectangular pulse oscillator algorithm. The trigger signal is zero when the trivial signal does not have a discontinuity in between samples. When a discontinuity is detected, the trigger signal is either 1 or −1, depending on the polarity of the discontinuity. d is the fractional delay of the discontinuity, as explained in Publication I, pp. 3–4.

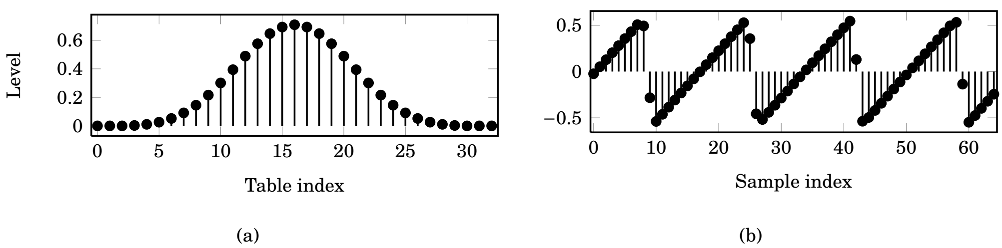


Figure 3.2. (a) Weighted least-squares-optimized basis function table and (b) the sawtooth signal generated by filtering a sequence of the basis functions. The sawtooth signal is obtained by integrating the impulse train with the leaky second-order integrator proposed by Brandt [2001].

Figure 3.3 shows the optimized polynomial basis functions for the BLIT and BLEP oscillators exemplified in Publication III. These polynomial basis functions can be synthesized with a third-order FIR filter as discussed in Publication III, Sec. II. The coefficients for the optimized BLIT filter as a function of the fractional delay d of the discontinuity (see Publication I, pp. 977–978, for details)

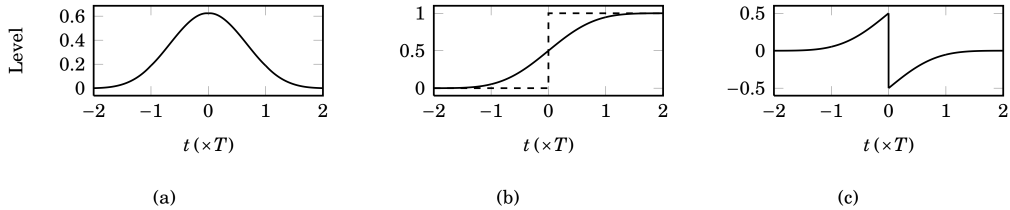


Figure 3.3. Optimized polynomial basis functions for the BLIT and BLEP oscillators: (a) BLIT basis function, (b) BLEP function (integrated BLIT function), and (c) BLEP residual function. The unit step function is plotted with a dashed line in (b).

are

$$\begin{aligned}
b_0(d) &= 0.00029 + 0.01474d + 0.01851d^2 + 0.15485d^3, \\
b_1(d) &= 0.18783 + 0.45327d + 0.44866d^2 - 0.46625d^3, \\
b_2(d) &= 0.62351 + 0.04817d - 0.95010d^2 + 0.46625d^3, \text{ and} \\
b_3(d) &= 0.18839 - 0.51631d + 0.48306d^2 - 0.15485d^3,
\end{aligned} \tag{3.1}$$

and for the BLEP residual filter they are

$$\begin{aligned}
b_0(d) &= 0.00029d + 0.00737d^2 + 0.00617d^3 + 0.03871d^4, \\
b_1(d) &= 0.05254 + 0.18783d + 0.22663d^2 + 0.14955d^3 - 0.11656d^4, \\
b_2(d) &= -0.5 + 0.62351d + 0.02409d^2 - 0.31670d^3 + 0.11656d^4, \text{ and} \\
b_3(d) &= -0.05254 + 0.18839d - 0.25816d^2 + 0.16102d^3 - 0.03871d^4.
\end{aligned} \tag{3.2}$$

Note that the polynomial coefficients given in Publication III are given as a function of the general time variable. The polynomial coefficients given here are obtained by translating the polynomial definition range to $[0,1]$.

In Publications I, II, and III, as well as in the most of the other papers dealing with quasi-bandlimited oscillator algorithms, the basis-function designs resulted in linear-phase functions. The only nonlinear-phase basis-function generators have been proposed by Brandt [2001], Nam et al. [2009, 2010], and Tassart [2013a]. However, the minimum-phase BLEP proposed by Brandt and allpass filter based algorithms discussed by Nam et al. utilize discrete-time functions that do not have well-defined continuous-time representations. The algorithm developed by Tassart samples the transition matrix of the state-space representation, yielding an approximation of the continuous-time impulse response of the prototype filter.

In Publication IV, a general technique that transforms a nonlinear-phase continuous-time basis function to a discrete-time basis-function generator is derived. In this technique, the continuous-time function is designed using analog lowpass filter design, and it is transformed to obtain a set of parallel infinite-impulse response (IIR) filters that generate the basis function with given fractional offsets.

Figure 3.4 shows an impulse train and the rectangular pulse wave generated with a transformed fifth-order elliptic filter, excitation signals of which are approximated with third-order polynomials. The impulse train (see Fig. 3.4(a)) shows that the generated basis function is clearly not linear-phase and that most of the impulse energy is concentrated at the beginning of the impulse signal. This means that the rectangular pulse wave has most of its amplitude variations right after the discontinuities, as can be seen in Fig. 3.4(b). With linear-phase basis functions, the variations are always symmetric with respect to the discontinuity (see, for example, Publication I, Fig. 2(c)).

The quasi-bandlimited oscillator algorithms typically produce waveforms in which the higher harmonics are attenuated from the their nominal levels. This problem can be overcome by applying a post-equalizing filter that boosts those components close to their nominal levels while keeping the level of the lower harmonics almost intact. Using a low-order post-equalizing filter suffices to perform the boosting, and the filter can be designed using any design process. Moreover, the post-equalizing filter can be a FIR filter, as exemplified in Publication I, or an IIR filter, as is illustrated in Publications II and III.

3.3 Alias-suppressing oscillator algorithms

The alias-suppressing oscillator algorithms differ greatly from the ideally bandlimited and quasi-bandlimited algorithms. Whereas in the two other categories of oscillator algorithms the objective is to *remove* aliasing as much as possible especially at low frequencies, the alias-suppressing algorithms allow aliasing in the entire audio band. However, the objective is to *suppress* aliasing distortion so that, in the best case, it becomes inaudible.

The alias-suppressing oscillator algorithms can be interpreted to sample a signal that has the same harmonic structure as the target waveform but with a faster decaying spectrum than the target. The sampled signal contains aliasing in the entire audio band, like in the trivial approach, but clearly suppressed compared to the trivial algorithm. After the sampling, the spectral tilt is restored to the target tilt by filtering the sampled signal with a digital post-processing highpass filter.

The earliest of alias-suppressing algorithms was based on oversampling of the trivial algorithm [Chamberlin, 1985]. By doubling the sample rate the aliased components of the sawtooth and rectangular pulse wave will have an amplitude

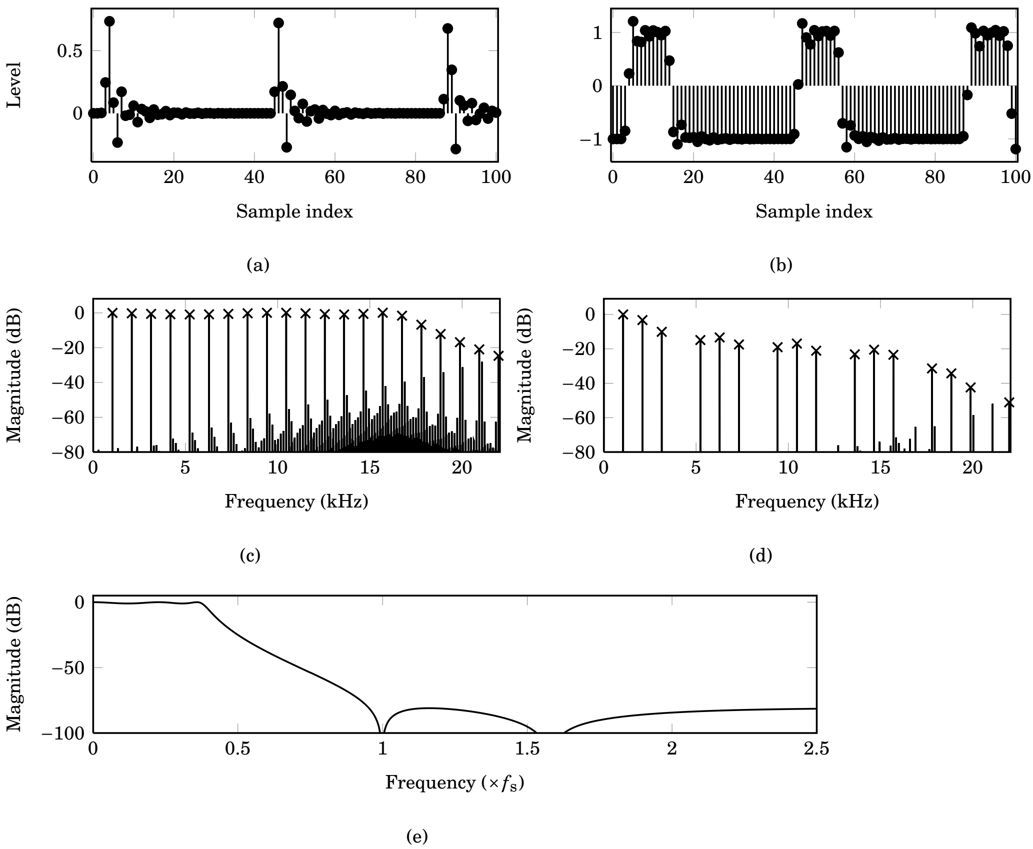


Figure 3.4. (a) Bandlimited impulse train and (b) the rectangular pulse wave generated by the nonlinear-phase oscillator algorithm when the fundamental frequency of the waveform is 1.047 kHz and the sample rate is 44.1 kHz. Their spectra are shown in (c) and (d), respectively. A fifth-order elliptic (Cauer) filter that has the cutoff frequency at $\frac{3}{4}\pi$, passband ripple of 1 dB, and stopband attenuation of 81 dB is used as the prototype filter. The magnitude response of the prototype filter is given in (e). The excitation signals are approximated with third-order polynomials.

that is at least 6 dB lower (see Sec. 2.3) than in the case of trivial sampling. In the case of triangular wave, the aliased components will be attenuated by at least 12 dB. With a high enough sampling rate aliasing distortion can be reduced to become inaudible [Chamberlin, 1985; Puckette, 2007]. However, in order to have a good alias-suppression performance, the oversampling factor needs to be quite high. Furthermore, the higher sampling rate results in increased computational load of the algorithm. On the other hand, the sampled signal has exactly the spectral tilt of the target, which means that this approach does not need a post-processing filter.

Another alias-suppressing sawtooth algorithm starts by full-wave rectifying a sine wave whose frequency is half of that of the target waveform [Lane et al., 1997]. The spectral shape of the resulting signal decays faster than that of the sawtooth waveform, and by applying a f_0-tracking highpass filter and a fixed lowpass filter, an approximation of the sawtooth waveform is obtained. By applying additional operations to the sawtooth algorithm, the rectangular pulse and triangular waveforms can also be generated from the sinusoid [Lane et al., 1997].

An algorithm related to the one proposed by Lane et al. [1997] generates the sawtooth signal by differentiating a squared, trivial sawtooth [Välimäki, 2005]. This approach is motivated by noting that the square of the trivial sawtooth waveform is also the integral of the waveform. The integration benefits the suppression of aliasing because the spectrum of an integrated signal decays by approximately 6 dB per octave faster than the spectrum of the integrand. This can be proved by analyzing the properties of the Laplace transform [Kreyszig, 1999, pp. 258–263]. Hence, the sampled squared sawtooth waveform contains less aliasing than the trivial waveform. Then, by applying a difference filter, the first-order FIR filter that has a zero at DC, the spectral tilt is adjusted to the original decay rate while keeping aliasing at low and middle frequencies practically unaltered [Välimäki, 2005].

Because integration increases the spectral tilt of a signal by about 6 dB per octave while differentiation reduces it by the same amount, it is obvious that the operation described above can be extended to higher integration and differentiation orders [Välimäki et al., 2010]. By increasing the order of the approach, i.e., the number of integrations applied, aliasing can be suppressed more than with a lower-order approach. Moreover, by running the differentiated polynomial waveform (DPW) algorithm at a higher sampling rate, aliasing distortion can be suppressed even more [Välimäki, 2005; Välimäki et al., 2010]. Furthermore, this approach is not limited to only oscillator algorithms: it has been extended to wavetable synthesis too [Franck and Välimäki, 2012, 2013].

However, this DPW algorithm requires a post-scaling operation that is inversely proportional to the fundamental frequency raised to a power smaller than the order of the algorithm. This means that the required post-scaling gain can exceed the numerical limits of the implementation platform at low fundamental frequencies even with moderately small algorithm orders [Välimäki et al., 2010].

The rectangular pulse and triangular waveforms can be generated by introducing additional operations to the sawtooth DPW algorithm [Huovilainen and Välimäki, 2005; Välimäki and Huovilainen, 2006; Välimäki et al., 2010; Huovilainen, 2010]. Alternatively, the rectangular pulse waveform can be obtained from the sawtooth signal by applying a FIR comb filter [Lowenfels, 2003][6]. The duty cycle of the resulting waveform is controlled via the delay line length that sets the locations of the notches of the comb filter. Similarly, the FIR comb filter can be applied to the squared sawtooth signal to produce a triangular wave [Puckette, 2007].

3.4 Ad-hoc oscillator algorithms

This category of oscillator algorithms contains a set of techniques that synthesize signals that resemble the classical geometric waveforms. The objective of these algorithms is not necessarily to be antialiasing. Instead, they typically utilize readily available simple signal processing methods to produce similar-looking waveforms that sound approximately the same as the target waveforms.

A set of ad-hoc algorithms try to reduce aliasing by post-processing the output of a trivial oscillator. By full-wave rectifying a slightly DC-shifted sawtooth signal and then by removing the offset, the spectral envelope of the aliased portion of the sawtooth waveform gets modified [Lisle and McDonald, 1993]. Alternatively, by taking the absolute value, the logarithm, and exponentiating the trivial sawtooth, the sharp discontinuities of the waveform will be replaced with smooth transitions [Chidlaw and Muha, 2004].

The algorithm proposed by Chidlaw and Muha [2004] performs its operations on all samples of the waveform. However, it suffices to replace a few samples around the discontinuities with values computed from a sinusoidal [Kleimola et al., 2011c] or polynomial [Kleimola and Välimäki, 2012; Ambrits and Bank, 2013] transition function. These approaches can be applied to any discontinuities, and the polynomial function has been shown to be an efficient technique to remove the transient issues occurring in the DPW algorithm[7] when the fundamental frequency is varied rapidly [Kleimola and Välimäki, 2012].

Publication V proposes two linear post-processing algorithms for the task of aliasing distortion reduction. Because aliasing is more easily heard at low frequencies below the fundamental (see Section 4), it can be suppressed by highpass filtering the alias-corrupted waveform. Aliasing between the harmonic components can be suppressed by filtering the oscillator output with a comb

[6]Actually, the rectangular pulse wave can be obtained from *any* sawtooth signal by applying a FIR comb filter.

[7]The reader may note that the DPW algorithm too can be understood to perform post-processing on the trivial sawtooth waveform. However, it has a close connection to continuous-time waveforms while these other post-processing algorithms do not. Therefore, it is not categorized as an ad-hoc oscillator algorithm.

filter. In Publication V, the applicability of both FIR and IIR comb filters to the task is analyzed.

However, the filters proposed in Publication V suffer from a couple of issues, as one can see in Fig. 3.5. The simplified highpass filter, whose f_0-dependent pole is computed from the linear function in Eq. (3) of Publication V, has a lower cutoff frequency than the exact filter as shown in Fig. 3.5(a). Moreover, the omitted scaling factor boosts high frequencies slightly. This can be seen in Fig. 3.5(b), where the spectrum of the trivially generated sawtooth waveform that has been filtered with the FIR comb filter and the simplified highpass filter. However, in general, the comb filters suppress the higher harmonics of the original signal because the filter response is inharmonic. The inharmonicity is due to the fractional delay filter that applies frequency-dependent, i.e., dispersive, delay to the signal. The dispersion effect is visible in the spectrum of the IIR comb and highpass filtered response plotted in 3.5(c). The crosses and circles indicate the levels of the non-aliased harmonics and their target levels, respectively.

The attenuation caused by the dispersion depends heavily on the fundamental frequency that is used to define the location of the first 0-dB peak of the comb filter. However, Figs. 3.5(b) and (c) also show that the suppression caused by the dispersion effect is much milder with the FIR than with the IIR comb filter. Yet, the IIR comb filter provides better alias suppression than the FIR comb filter between the waveform harmonics.

The filtered waveforms are plotted in Figs. 3.5(d), (e), and (f) for the simplified highpass filter, the cascade of the FIR comb and highpass filter, and the cascade of the IIR comb and highpass filter, respectively. The output of the cascade of the FIR comb filter and the simplified highpass filter (see Fig. 3.5(e)) is almost an exact replica of the only highpass filtered waveform (Fig. 3.5(d)), indicating that the effect of the FIR comb filtering is fairly subtle.

The rest of the ad-hoc algorithms utilize nonlinear tricks to generate classical-looking, pseudo-geometric waveforms. One of these algorithms uses waveshaping, or amplitude distortion, that is applied to the output of a sinusoidal oscillator. The rectangular pulse wave can be easily obtained from a (DC-shifted) sinusoid by applying a sign-function-like sigmoid function [Timoney et al., 2009a; Lazzarini and Timoney, 2010a]. The sawtooth and triangular waves can also be generated with waveshaping functions. Alternatively, the sinusoid can be distorted with a mix of Chebychev polynomials [Pekonen, 2007].

In addition to the amplitude distortion, the sinusoidal oscillator can be distorted by shaping the phasor signal. By passing the phasor signal through a two-piece linear phase distortion (PD) function, the shape of the sinusoid starts

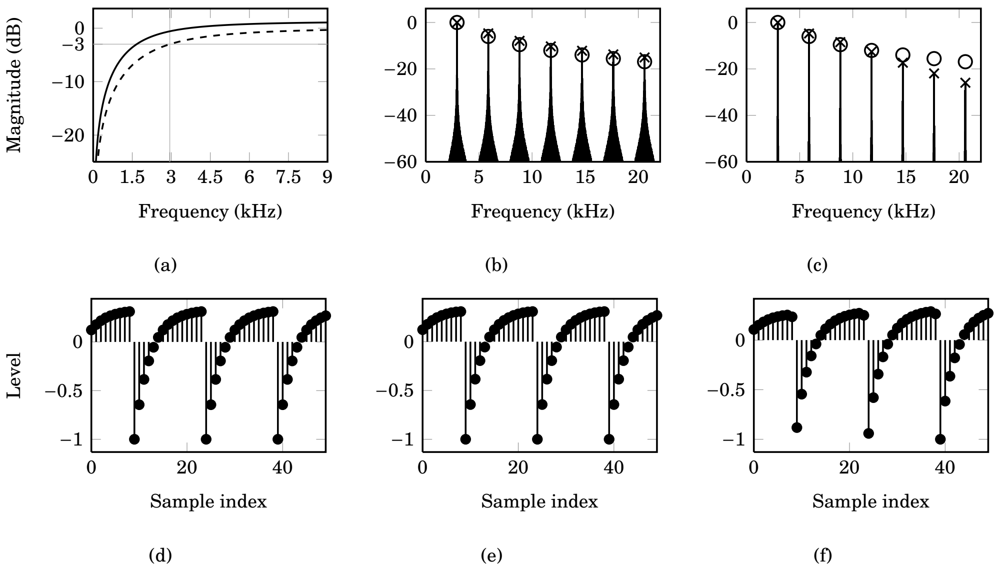


Figure 3.5. (a) Frequency response of the highpass filter with the approximated filter pole (solid line) and with the exact filter pole (dashed line), and the spectra of the (b) FIR-comb filtered and (c) IIR-comb filtered trivial sawtooth waveform that is also filtered with the simplified highpass filter. The fundamental frequency, which is also the cutoff frequency of the highpass filter, is 2.941 kHz and the sample rate is 44.1 kHz. The fractional delay filter in the FIR and IIR comb filters is a first-order allpass filter. The crosses and circles in (b) and (c) indicate the actual and ideal levels of the waveform harmonics, respectively. The output waveforms of the simplified highpass filter, the cascade of the FIR comb and the highpass filters, and the cascade of the IIR comb and the highpass filter are shown in (d), (e), and (f), respectively.

to resemble a sawtooth waveform [Ishibashi, 1987]. A pseudo-rectangular pulse wave is obtained with a four-piece PD function. The sharpness of the transition of the pseudo-classical waveforms and the pulse width of the pseudo-rectangular pulse wave depend on the control points used in the PD function.

A PD function can be decomposed into a linear part and a time-varying part, implying that phase distortion is effectively a subset of a more general phase modulation synthesis approach. Therefore, the PD function can be implemented by adding a modulation function to the unmodified (linear) phasor signal. This approach can be used in the traditional (wavetable) oscillator implementation that generates the phase-modified sinusoid.

The PD oscillator algorithm can also be implemented in couple of alternative ways. Lazzarini et al. [2007] proposed a structure that enables FM to be applied to arbitrary signals. By modulating the delay-line length (in practice, the point at which the output is read), the phase of the signal is modified. This approach can use any sinusoidal oscillator as the input signal, relaxing possible issues raised by the modification of the phase of the actual sinusoidal oscillator. However, the modulated delay-line length is not necessarily an integer. Therefore, this method requires a fractional delay filter to take this issue into account [Laakso et al., 1996]. In order to avoid amplitude modifications, the fractional delay filter used is an allpass type [Lazzarini et al., 2007].

The phase modulation introduced by the modulated delay line can also be generated with a single time-varying first-order allpass filter [Timoney et al., 2009b]. By modulating the coefficients of the allpass filter, the filter introduces a time-varying phase delay that effectively results in the same output as phase modulation. By applying different coefficient modulation signals, the different pseudo-classical waveforms can be obtained [Timoney et al., 2009a,b; Lazzarini et al., 2009b; Lazzarini and Timoney, 2010a].

Because the allpass filter is a recursive filter, the coefficient modulation may cause the filter to become unstable even when the coefficients vary in a range where the time-invariant filter is stable [Laroche, 2007]. The properties of time-varying first-order allpass filter were analyzed by Pekonen [2008], who showed that the time-varying first-order allpass filter responds to a bounded input with a bounded output, i.e., it is stable, when the modulation signal is in the range $[-1, 1]$, end points included, independent of the input signal.

It should be noted that the implementation of the time-varying allpass filter affects the output dramatically. The filter structure discussed by Pekonen [2008] is the transposed direct form II, which has transient issues when the modulation signal changes rapidly [Timoney et al., 2009a,b]. With alternative filter structures the transient can be removed or suppressed, as shown in Fig. 3.6. The stability condition for these other implementations can be shown to be exactly the same as with the transposed direct form II.

However, the modulation signal of the single time-varying allpass filter is not simple when pseudo-classical waveforms are synthesized [Timoney et al., 2009a,b; Lazzarini and Timoney, 2010a]. This is shown in Fig. 3.6(c). The modulation signal can be simplified by cascading a set of time-varying first-order allpass filters. This approach was proposed by Kleimola et al. [2009] who noted that the input sinusoid as the modulation signal is sufficient when the length of the allpass filter chain is large enough. The stability conditions of the time-varying allpass filter chain were analyzed by Pekonen et al. [2009].

Kleimola et al. [2011c] generalized the phaseshaping principle and defined the PD function as a control interface for the resulting timbre. By choosing an alternative PD function and possibly cascading it with another function, one can obtain pseudo-classical waveforms that are different from the original waveforms generated by the two- or four-piece PD functions. However, the design of the

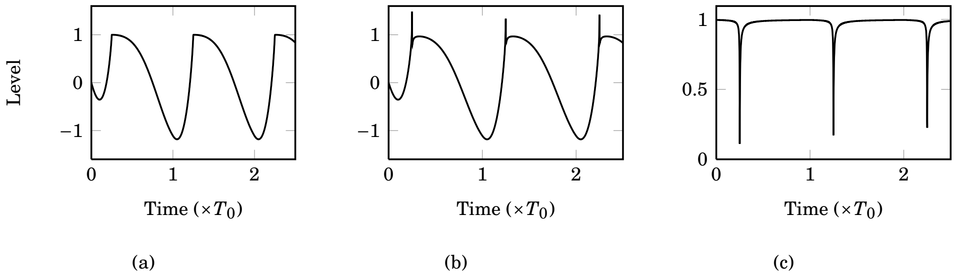


Figure 3.6. PD-like sawtooth waveform generated by a coefficient-modulated allpass filter when the filter structure is (a) direct form I and (b) transposed direct form II. The coefficient modulation signal with which the waveform is obtained is plotted in (c).

multi-part PD function requires careful tuning, and it is prone to introducing rapid amplitude changes that result in harsh aliasing distortion [Kleimola et al., 2011c; Kleimola, 2013]. This issue can be softened by using the alias-reduction techniques mentioned earlier in this section.

Pseudo-classical waveforms can also be generated using the feedback FM algorithm [Tomisawa, 1981], where the output of a sinusoidal oscillator is fed back to its frequency control [Schoffhauzer, 2007]. By controlling the amount of feedback, the spectral content of the oscillator output and hence the amount of aliasing can be controlled. The same idea can be applied the amplitude control of the sinusoidal oscillator [Lazzarini et al., 2009a; Kleimola et al., 2011b,a]. The classical-looking waveforms can be generated by modifying the feedback signal with a waveshaping function.

Another amplitude-related approach is to apply bit-wise logical operations to a sinusoid [Kleimola, 2008]. In this approach, the individual bits of the digital representation of the samples are modified via Boolean arithmetic rules that are applied to the bits of the sinusoid and a modulation signal. This purely digital amplitude modification can result in harsh amplitude changes, which effectively increases aliasing distortion [Kleimola, 2008].

4. Audibility of Aliasing in Classical Waveforms

As pointed out in the previous section, many antialiasing oscillator algorithms take advantage of the knowledge of human hearing. While the ideally band-limited waveform could be the ultimate objective of an antialiasing algorithm, the physiological constraints of human hearing relaxes the requirement for a such an extreme objective. In fact, these constraints allow the waveform eventually to have quite a lot of aliasing distortion before it becomes audible.

4.1 Psychoacoustic phenomena affecting the audibility of aliasing

There are two psychoacoustic phenomena that contribute to the audibility of aliasing. The first one is the frequency masking phenomenon. When an aliased component is close to a non-aliased component in frequency, it is rendered inaudible if its amplitude is small enough [Zwicker and Fastl, 1990; Moore, 1997]. The closer the aliased component is to the harmonic peak, the larger the amplitude of the aliased component can be before it becomes audible.

The second factor that affects the audibility of aliasing is the hearing threshold. If the amplitude of the aliased component is, in general, very small, it may not have enough power to excite the sensory system [Zwicker and Fastl, 1990; Moore, 1997]. The hearing threshold has a great impact especially at high frequencies as the threshold increases steeply above 15 kHz. When the fundamental frequency of the waveform is high, the hearing threshold defines the audibility of aliasing at low frequencies too. In such cases, the frequency masking threshold of the fundamental component becomes a smaller contributor than the hearing threshold [Zwicker and Fastl, 1990; Moore, 1997].

The actual threshold of aliasing audibility results from the combination of both of these phenomena. While the hearing threshold affects the audibility in all conditions, the overall frequency masking threshold depends on the non-aliased components and the pattern they form. Therefore, the audibility of aliasing depends on both the waveform and on its fundamental frequency.

Both of the abovementioned psychoacoustic phenomena have been studied extensively [Wegel and Lane, 1924; Egan and Hake, 1950; Greenwood, 1961; Zwicker and Fastl, 1990; Moore, 1997]. From these studies, statistical threshold patterns have been drawn. While the statistically obtained threshold of hearing can be understood as a general threshold for any kind of signal, the setups used for the frequency masking threshold studies differ greatly from the aliasing issue occurring in the digital classical waveforms.

The frequency masking threshold studies use either a pure tone masking another [Wegel and Lane, 1924] or a noise signal masking a pure tone [Egan and Hake, 1950; Greenwood, 1961; Zwicker and Fastl, 1990]. However, the aliasing issue occurring in the digital classical waveform differ from those setups. Although aliasing distortion of digital classical waveforms do have a wideband noise-like spectrum, it is also periodic. This means that the aliased components, if audible, can sound like a noise signal, but some of its components can also be heard as tonal components. To date, only Schimmel [2012] has studied the audibility of aliasing distortion in digital audio synthesis.

4.2 Audibility of aliasing in trivially generated sawtooth signals

In Publication VI, the audibility of aliasing in the trivially generated digital sawtooth wave is for the first time investigated using a listening test. The test subjects were exposed to ideally alias-free and (trivially) alias-corrupted signals in an adaptive three-alternative forced-choice test where the threshold of audibility of aliasing was sought.

The test hypothesis in Publication VI assumes that the frequency masking phenomenon contributes more to the audibility of aliasing between the harmonic components, apart from very high frequencies. In practice, this means that aliasing distortion above the fundamental frequency is mostly masked by the non-aliased components. Below the fundamental frequency, on the other hand, the frequency masking is expected not to play such a big role because the frequency masking threshold rolls off faster below the masker than above it [Wegel and Lane, 1924; Egan and Hake, 1950; Greenwood, 1961; Zwicker and Fastl, 1990; Moore, 1997]. Instead, the hearing threshold is assumed to have a more significant effect at very low frequencies than the frequency masking.

The results of the listening test conducted in Publication VI support this hypothesis. The thresholds obtained with and without the masker, i.e., the ideally bandlimited sawtooth signal, are statistically different when aliasing distortion was above the fundamental frequency. Below the fundamental frequency no statistical difference was found.

Based on the verified hypothesis and the numerical results of the test, design rules for the aliasing distortion pattern generated by an antialiasing oscillator algorithm are also described in Publication VI. By complying with these rules, the waveform produced by the algorithm will be perceptually alias-free.

The design rules are obtained by computing the maximum allowed levels for the aliased components below and above the fundamental frequency. Below the fundamental frequency, the most prominent component is the first component from the first-order generation of aliasing, i.e., from the frequencies that would be between the Nyquist limit and the sampling frequency, that folds back to this frequency range. Based on informal listening tests, the most likely audible component above the fundamental frequency is the strongest component between the first and the second harmonic component. The aliased components at higher frequencies are not heard as easily.

After these components are found, their levels can be computed from their harmonic index using the Fourier series expression of the sawtooth waveform. These levels must be modified by the required attenuation obtained from the test. By attenuating this modified level also by the magnitude of the variation in the confidence interval, the design rules for that fundamental frequency are obtained. The general design rules are obtained by finding the minimum allowed levels at different frequencies.

At frequencies that fold back close to DC, this requirement can be considered to be too strict. It can be relaxed by finding the minimum allowed level up to the fundamental frequency that corresponds to the aliased frequency. In addition, the requirement can be ignored at frequencies that fold above 15 kHz because human hearing is insensitive to the components above that frequency (see the discussion in Sec. 4.1).

The general design rules derived in Publication VI are shown in Fig. 4.1 with the solid line. The dotted line indicates relaxed design levels at frequencies that fold back close to DC. Also shown in Fig. 4.1(a) are the spectral envelopes of the sawtooth signals that have a fundamental frequency of 3.951 kHz when they are generated using the second-order and fourth-order Lagrange polynomial BLEP (see Publication I) method. The spectral envelopes of the sawtooth obtained with the fourth-order B-spline and optimized polynomial (see Publication III) are plotted in Fig. 4.1(b).

Figure 4.1(a) shows that the sawtooth signal generated with the second-order polynomial BLEP (dashed line) is not alias-free at this fundamental frequency,

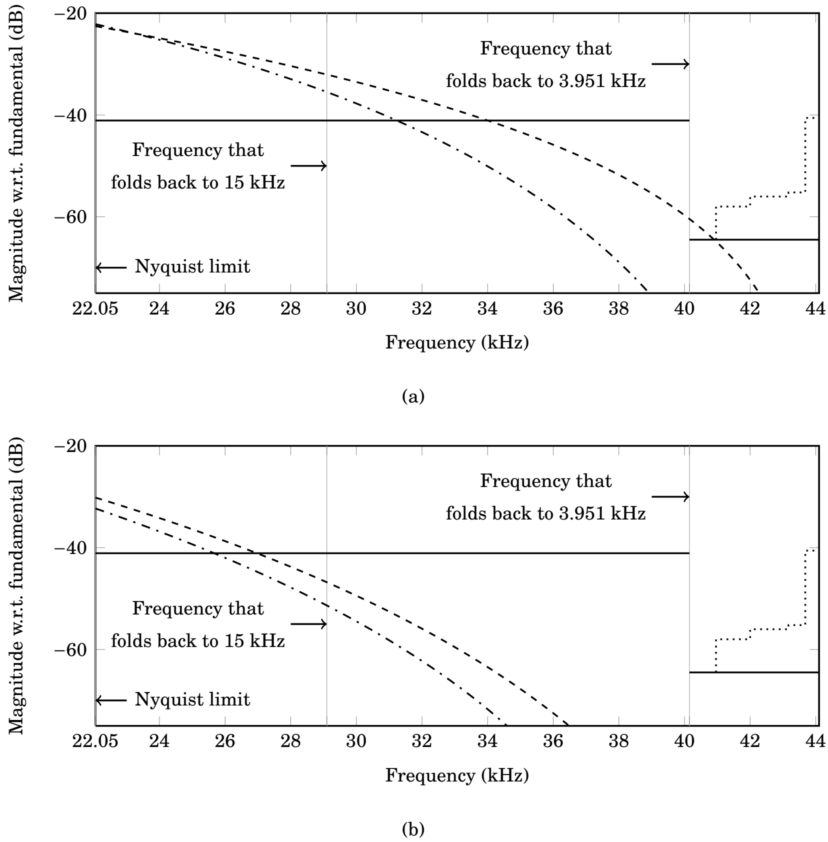


Figure 4.1. Maximum allowed levels of aliased components (solid line) and the spectral envelopes of the sawtooth waveform that is generated with the polynomial BLEP approach using (a) the second-order (dashed line) and fourth-order Lagrange (dash-dotted line), and (b) the fourth-order B-spline (dashed line) and optimized (dash-dotted line) polynomial basis function approximations. The envelopes of the sawtooth signals are for the fundamental frequency of 3.951 kHz, and the sample rate is 44.1 kHz. The dotted line represents the relaxed level requirement for the frequencies that fold back close to DC.

confirming the analysis made with the computational measures of aliasing distortion (discussed below in Sec. 4.3) and the observations made with informal listening tests. The fourth-order Lagrange approximation (dash-dotted line) appears to contain some aliasing above the fundamental frequency but at frequencies that fold back between 13 and 15 kHz. However, it should be noted that with the fundamental frequency of 3.951 kHz there are no first-order generation aliased components in that frequency range.

The fourth-order B-spline approximation and the optimized polynomial approach that comply with the design rules, spectral envelopes of which are shown in Fig. 4.1(b) with dashed and dash-dotted lines, respectively, are completely alias-free. This verifies the analysis made with the computational measures of the audibility of aliasing.

Note that the possibly applied signal manipulations, such as filtering, modulation, and distortion, can affect audibility of aliasing of the signal. Especially variations in the fundamental frequency of the waveform have a prominent effect. However, in Publication VI the focus was on audibility of aliasing in the pure sawtooth signals with fixed fundamental frequencies.

4.3 Computational measures of audibility of aliasing

Unfortunately, a listening test that examines the audibility of aliasing of a set of oscillator algorithms with all possible parameter combinations would take a huge amount of time to perform. In fact, a test that would analyze all possible combinations of a single algorithm would take a lot of time to complete. To avoid such tedious work, the algorithms can be analyzed using computational measures that simulate human hearing. Measures that have been used for the analysis of oscillator algorithms include a model of the hearing threshold and the frequency masking phenomenon and the noise-to-mask ratio (NMR).

The first measure utilizes the statistically obtained thresholds for both the hearing threshold and the frequency masking threshold obtained from the tone–masking-tone experiments. The use of this model for the evaluation of aliasing distortion was first proposed by Nam et al. [2010]. The overall threshold for the audibility of aliasing is obtained by computing the maximum of the hearing threshold and the individual frequency masking thresholds of the non-aliased components. If there are aliased components above the threshold, the waveform contains audible aliasing distortion according to this measure.

The NMR measure, which was originally proposed for the evaluation of audio codecs [Brandenburg, 1987; Brandenburg and Sporer, 1992], can be interpreted as a simplified version of the model proposed by Nam et al. While the Nam model uses as accurate expressions as possible for the thresholds, the NMR measure uses less complex expressions. The actual NMR algorithm compares a corrupted signal to the alias-free, or, more generally, the error-free signal. The algorithm gives a numeric value that tells what the ratio of the unwanted part of the corrupted signal to the threshold is. If the ratio is below 0 dB[1], the waveform is assumed to be alias-free.

In addition to the listening test, Publication VI analyzes the applicability of these computational measures to the automatic analysis of the audibility of

[1]However, in audio coding, −10 dB is considered to be the threshold of audibility for coding artefacts.

aliasing by comparing the results of the listening test with the results of the measures. The Nam model was found to yield conservative results. Moreover, the NMR measure was also observed to have limitations. Analysis of test signals that do not contain aliasing according to the listening test were found by these two measures to have clearly audible aliasing.

The contradiction with the Nam model is present especially when the signal contains aliased components whose frequencies are in the range where the frequency masking thresholds of two harmonic components are about to cross. The operation that combines the individual frequency masking thresholds is suspected to be the cause of this difference in the results, as discussed in Publication VI. The Nam model uses the maximum value of the components assuming that the higher threshold solely determines the overall threshold. However, there are studies that imply that also the weaker masking threshold contributes to the overall threshold and that the overall threshold can be obtained as a nonlinear combination of the individual thresholds [Green, 1967; Lufti, 1983, 1985; Humes and Jesteadt, 1989].

In the NMR analysis, signals that were alias-free according to the listening test yielded NMR results that suggested that the signals contained clearly audible aliasing distortion. A closer inspection of the NMR results indicated that aliasing distortion below the fundamental frequency of the signal was ranked disturbing by the NMR algorithm. The deviation was found to be caused by the hearing threshold model used by the NMR algorithm as it does not take into account the rise of the hearing threshold at low frequencies. To take the contribution of the threshold into account, A-weighted signals were analyzed to study the NMR in Publication VI. These NMR results were found to be more consistent with the listening test than the non-weighted results.

In addition, there is the "perceptual evaluation of audio quality" (PEAQ) measure. The PEAQ measure is, in principle, a combination of multiple computational measures developed for audio quality evaluations [International Telecommunication Union, 1998] mainly in audio codec development and analysis. The PEAQ algorithm runs a selection of evaluation algorithms and combines their results into a single number on the mean opinion score scale. However, it was noted to yield inconsistent results in the algorithm testing performed in Publication I. The inconsistency may be caused by the similar minor limitations of the algorithms used in the PEAQ analysis as the Nam model and the NMR measure have. In order to yield consistent results also with the PEAQ algorithm, the contradicting results of the underlying analysis algorithms should be analyzed more closely in a separate study.

5. Modeling of Analog Waveforms

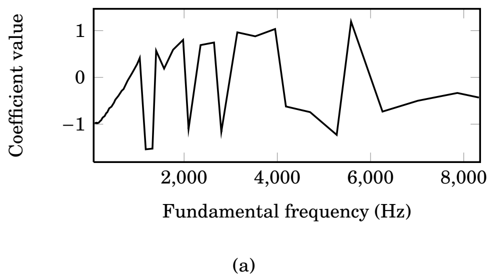
The goal of the oscillator algorithms presented in Section 3 was to approximate the ideal continuous-time waveforms in order to reduce aliasing. However, the actual source waveforms generated by the analog synthesizers differ from the ideal textbook waveforms [Kleimola, 2005; Kleimola et al., 2010; Kleimola, 2013]. Moreover, the analog waveforms typically sound less harsh than the textbook waveforms. In order to have a true analog oscillator model, these differences need to be taken into account.

An obvious approach to modeling of analog waveforms is the sampling/wavetable synthesis. However, when several different synthesizers are desired to be modeled in a single unit, the amount of memory needed for the different waveforms becomes easily huge. On the other hand, parametric models that, by a change of model parameters, can reproduce close approximations of the waveforms of the analog synthesizers result in dramatically reduced memory requirements.

Parametric models for an analog oscillator is developed for the first time in Publication VII. Before Publication VII, only ad-hoc modeling approaches were proposed. De Sanctis and Sarti [2010] developed a wave-digital filter model for an astable multivibrator circuit that can used to generate analog classical waveforms. Kleimola et al. [2010], in turn, proposed an ad-hoc PD model for the Moog sawtooth oscillator.

The authors of Publication VII introduced a general post-processing approach that filters the output of an antialiasing oscillator algorithm. In this approach, the objective is to find a low-order filter that tries to match the spectrum of the signal generated by an antialiasing oscillator algorithm to the spectrum of the analog waveform. The filter applied for the task in Publication VII is a first-order IIR filter.

One may wonder whether the accuracy of the post-processing approach can be improved by estimating higher-order post-processing filters. However, modeling of the fundamental-frequency-dependent filter parameters can become harder



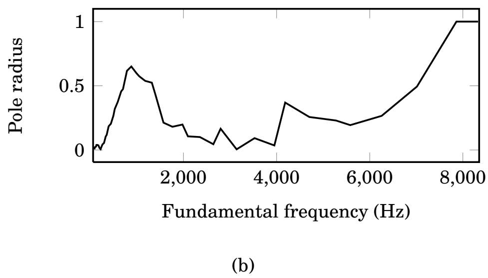


Figure 5.1. (a) Coefficient of the first-order feedback term and (b) a pole radius of the second-order post-processing filter as a function of the fundamental frequency estimated for the ideally bandlimited oscillator.

as the filter order is increased. This issue is illustrated in Fig. 5.1 with the feedback term of the second-order post-processing filter that was fitted to the ideally bandlimited oscillator in the least-squares sense in the case of the Moog sawtooth oscillator. Figure 5.1(a) shows that the estimated coefficient of the first-order feedback term varies quite a lot for neighbouring fundamental frequencies. Even the pole radius, plotted in Fig. 5.1(b), does not show an easily modelable behaviour as a function of the fundamental frequency. Therefore, using a higher-order filter may not ultimately provide a better model of the analog waveform than the first-order filter.

As mentioned in Publication VII, use of fundamental-frequency-dependent recursive post-processing filters requires careful design. Alternatively to general filter design techniques, one can use the parametric recursive filters, such as those introduced by, e.g., Regalia and Mitra [1987] and Tassart [2013b].

The PD model of the Moog sawtooth waveform, introduced by Kleimola et al. [2010], is generalized in Publication VII. The PD model is valid because the Moog sawtooth resembles the PD sawtooth waveform. While the model by Kleimola et al. uses only a quarter of a sinusoid with a small tweak at the transition regions, the PD model presented in Publication VII uses the whole period and sets the speed of the transition with the control point of the PD sawtooth model.

The post-processing approach of the Moog sawtooth waveform yields a more general modeling technique. While the PD model was well matched to the Moog sawtooth, the approach may not work with the waveforms generated by other analog synthesizers. On the other hand, as pointed out above, the post-processing approach can be applied to any analog synthesizer waveforms by estimating the filter parameters for the waveforms that are to be modeled. With this technique, a virtual analog synthesizer can generate the source waveforms of multiple analog synthesizers by filtering the output of an antialiasing oscillator algorithm with different post-processing filters.

6. Main Results of the Thesis

Publication I: "Perceptually informed synthesis of bandlimited classical waveforms using integrated polynomial interpolation"

In Publication I, the polynomial BLEP approach, originally proposed by Välimäki and Huovilainen [2007], is extended to higher-order integrated interpolation polynomials. In addition, the closed-form basis function of the ideal BLEP is derived in Publication I. The proposed polynomials are shown to yield excellent alias reduction when analyzed using computational measures. Moreover, the computational cost of these polynomials is indicated to be low.

Publication II: "On minimizing the look-up table size in quasi bandlimited classical waveform oscillators"

While Publication I illustrates that the alias-reduction performance of the polynomial basis functions is superior to the traditional tabulated windowed basis function approach, the table-based algorithm is not completely useless. Publication II shows that by sampling a parametric window function or by optimizing the table values using a perceptually informed objective function, aliasing distortion can be reduced significantly also in the table-based oscillator algorithm.

Publication III: "Optimized polynomial spline basis function design for quasi-bandlimited classical waveform synthesis"

In Publication III, the optimization procedure presented in Publication II is applied to polynomial basis-function designs. Compared to the integrated interpolation polynomials in Publication I, the resulting polynomials are shown to yield even better alias reduction than the interpolation polynomials. The waveforms generated by the polynomial BLIT and BLEP algorithms are analyzed and found to be alias-free at all fundamental frequencies typically used in music. In fact, the algorithms are shown to produce alias-free waveforms that have only one component in the hearing range in the extreme cases at very high fundamental frequency.

Publication IV: “Nonlinear-phase basis function generators for quasi-bandlimited waveform synthesis”

Most of the basis function designs, including those introduced in Publications I–III, for quasi-bandlimited oscillator algorithms are linear-phase functions. Publication IV introduces a novel nonlinear-phase approach to the basis-function design. The proposed approach transforms an analog prototype filter to a set of parallel digital IIR filters that sample the impulse response of the prototype filter with arbitrary time shifts. The resulting digital system is triggered with short burst-like signals that can be approximated efficiently with low-order polynomials. By choosing the prototype filter properly, the nonlinear-phase approach can provide excellent alias-reduction performance with low computational complexity.

Publication V: “Filter-based alias reduction for digital classical waveform synthesis”

Publication V proposes new linear post-processing approaches with which alias distortion can be reduced. The proposed filters can be applied to the output of any oscillator algorithm to improve the alias-reduction performance. By filtering the waveform with a highpass filter, alias distortion can be suppressed below the fundamental frequency. However, in order to get the best overall alias-reduction performance the waveform needs to be filtered with an IIR comb filter.

Publication VI: “Audibility of aliasing distortion in sawtooth signals and its implications to oscillator algorithm design”

Before Publication VI, the audibility of aliasing distortion produced by the different oscillator algorithms were analyzed using computational measures and informal listening tests. In Publication VI, the threshold of audibility of aliasing in a trivially sampled sawtooth signal is sought with a formal listening test.

The previously used computational measures are analyzed in the light of the results obtained from the test. In addition, generic thresholds of the audibility of aliasing distortion are proposed to help with the oscillator algorithm design.

Publication VII: "Discrete-time modelling of the Moog sawtooth oscillator waveform"

Apart from a couple of special cases, the objective of oscillator algorithms has been the ideal textbook waveforms. However, the waveforms produced by an analog synthesizer differ from these ideal signals. In Publication VII, two alternative parametric approaches for the modeling of the sawtooth waveform generated by the MiniMoog analog synthesizer are proposed. Of the proposed alternatives, the second approach that filters the output of an antialiasing oscillator algorithm can also be applied to other waveforms.

7. Conclusions

This thesis presented the recent development in the oscillator algorithm design for virtual analog synthesizers. A special focus was on time-varying filter-based algorithms that yield efficient algorithms and great alias reduction. In addition, some problems on the audibility of aliasing distortion and the modeling of the actual analog waveforms were addressed.

Even though the oscillator algorithms discussed in this thesis seem to offer alternative implementations, the problem of finding a computationally efficient antialiasing oscillator algorithm is not completely solved yet. As discussed by Pekonen and Välimäki [2011], there is no optimal oscillator algorithm that has all three desirable properties:

1. it generates alias-free signals in the range of musically interesting fundamental frequencies, such as from 20 Hz to 8 kHz,

2. it is computationally efficient and has low memory consumption, and

3. it does not require a division by a time-varying parameter, like the fundamental frequency.

For instance, the algorithms proposed in Publications I and III have the first two properties. However, the computation of the fractional delay value (see Publication I, Eq. (13)) requires a division by the fundamental frequency. On the other hand, many of the ad-hoc algorithms (see Section 3.4) may have the third property, but they do not fulfill the first or the second requirement.

In addition, audibility of aliasing distortion is still a topic that has not been investigated thoroughly. Publication VI was the first ever publication that dealt with the topic, but it focused only on the trivially sampled sawtooth. The other waveforms (rectangular pulse wave with different pulse width, triangular wave, and asymmetric triangular wave) as well as other aliasing patterns, e.g., different spectral tilts, need to be studied to find general thresholds for audibility of aliasing. Moreover, the study in Publication VI used a few discrete fundamental frequencies. The data between these frequencies also need to be analyzed. Similarly, the effects of different modulations applied to the fundamental frequency, e.g., glissando and vibrato, on audibility requires additional studies. Likewise, the threshold of audibility of aliasing should be investigated also for special oscillator effects like supersaw and hard/soft sync [Stilson, 2006; Välimäki and Huovilainen, 2006; Nam et al., 2010; Kleimola et al., 2010; Timoney et al., 2012].

Another topic that will gain interest in source-signal research is the modeling of the output waveforms of analog synthesizers. This research is needed for realistic virtual analog synthesis modeling. So far, only the sawtooth waveform of the MiniMoog Voyager has been modeled (Kleimola et al. [2010] and Publication VII). In the future, different analog synthesizers will be analyzed. In addition to extending the selection of modeled synthesizers, the models will most likely use signal-based and circuit-based techniques, just like with the synthesizer filters [Stilson and Smith, 1996b; Fontana, 1997; Huovilainen, 2004; Stilson, 2006; Välimäki and Huovilainen, 2006; Hélie, 2006; Civolani and Fontana, 2008; Stinchcombe, 2008; Hélie, 2010; Fontana and Civolani, 2010; Huovilainen, 2010; Hélie, 2011; Germain, 2011; Zambon and Fontana, 2011; Smith, 2012; Daly, 2012; Parker and D'Angelo, 2013]. The first approach tries to model the waveform of the oscillator output while the second technique models the oscillator circuitry.

In addition to the actual models of the analog synthesizer oscillators, the perceptual aspect of the differences between the analog and modeled waveforms should also be investigated. A particularly interesting issue would be to find the threshold at which the modeling accuracy is sufficient. In addition, the oscillator effects that are used in analog oscillators would require listening tests to verify the perceptually correct behavior of the models.

Bibliography

Abramowitz, M. and Stegun, I. A., editors (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. United States Department of Commerce, Washington, D.C., 10th edition.

Alles, H. G. (1980). Music synthesis using real time digital techniques. *Proceedings of the IEEE*, 68(4):436–449.

Ambrits, D. and Bank, B. (2013). Improved polynomial transition regions algorithm for alias-suppressed signal synthesis. In *Proceedings of the Sound and Music Computing Conference (SMC 2013)*, pages 561–568, Stockholm, Sweden.

Bateman, W. (1980). *Introduction to Computer Music*. John Wiley & Sons, Inc., New York, NY.

Brandenburg, K. (1987). Evaluation of quality for audio encoding at low bit rates. In *Proceedings of the 82nd AES Convention*, London, UK. Preprint 2433.

Brandenburg, K. and Sporer, T. (1992). 'NMR' and 'Masking Flag': Evaluation of quality using perceptual criteria. In *Proceedings of the 11th International AES Conference on Test and Measurement*, pages 169–179, Portland, OR.

Brandt, E. (2001). Hard sync without aliasing. In *Proceedings of the 2001 International Computer Music Conference (ICMC'01)*, pages 365–368, Havana, Cuba.

Burk, P. (2004). Band limited oscillators using wave table synthesis. In Greenebaum, K. and Barzel, R., editors, *Audio Anecdotes II — Tools, Tips, and Techniques for Digital Audio*, pages 37–53. A. K. Peters Ltd., Wellesley, MA.

Cahill, T. (1897). Art of and apparatus for generating and distributing music electronically. U.S. patent 580,035.

Carlson, A. B., Crilly, P. B., and Rutledge, J. C. (2002). *Communication Systems: An Introduction to Signals and Noise in Electrical Communication*. McGraw-Hill, New York, NY, 5th intl. edition.

Chamberlin, H. (1985). *Musical Applications of Microprocessors*. Hayden Book Company, Hasbrouck Heights, NJ, 2nd edition.

Chaudhary, A. (1998). Band-limited simulation of analog synthesizer modules by additive synthesis. In *Proceedings of the 105th AES Convention*, San Francisco, CA. Preprint 4779.

Chidlaw, R. and Muha, R. (2004). Oscillator providing waveform having dynamically continuously variable waveshape. U.S. patent 6,806,413.

Civolani, M. and Fontana, F. (2008). A nonlinear digital model of the EMS VCS3 voltage-controlled filter. In *Proceedings of the 11th International Conference on Digital Audio Effects (DAFx-08)*, pages 35–42, Espoo, Finland.

Daly, P. (2012). A comparison of virtual analogue Moog VCF models. Master's thesis, University of Edinburgh, Edinburgh, UK.

De Sanctis, G. and Sarti, A. (2010). Virtual analog modeling in the wave-digital domain. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(4):715–727.

Deslauriers, G. and Leider, C. (2009). A bandlimited oscillator by frequency-domain synthesis for virtual analog applications. In *Proceedings of the 127th AES Convention*, New York, NY. Preprint 7923.

Egan, J. P. and Hake, H. W. (1950). On the masking pattern of a simple auditory stimulus. *Journal of the Acoustical Society of America*, 22(5):622–630.

Erkut, C., Välimäki, V., Karjalainen, M., and Penttinen, H. (2008). Physics-based sound synthesis. In Polotti, P. and Rocchesso, D., editors, *Sound to Sense, Sense to Sound: A State of the Art in Sound and Music Computing*, chapter 8, pages 303–343. Sound and Music Computing Network.

Farrow, C. W. (1988). A continuously variable digital delay element. In *Proceedings of the IEEE Symposium on Circuits and Systems*, volume 3, pages 2641–2645, Espoo, Finland.

Fontana, F. (1997). Preserving the structure of the Moog VCF in the digital domain. In *Proceedings of the 1997 International Computer Music Conference (ICMC'97)*, pages 291–294, Thessaloniki, Greece.

Fontana, F. and Civolani, M. (2010). Modeling of the EMS VCS3 voltage-controlled filter as a nonlinear filter network. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(4):760–772.

Franck, A. (2008). Efficient algorithms and structures for fractional delay filtering based on lagrange interpolation. *Journal of the Audio Engineering Society*, 56(12):1036–1056.

Franck, A. and Välimäki, V. (2012). Higher-order integrated wavetable synthesis. In *Proceedings of the 15th International Conference on Digital Audio Effects (DAFx-12)*, pages 245–252, York, UK.

Franck, A. and Välimäki, V. (2013). Higher-order integrated wavetable and sampling synthesis. *Journal of the Audio Engineering Society*, 61(9):624–636.

Frei, B. (2002). Digital sound generation. Online publication, available at `http://www.icst.net/research/projects/digital-sound-generation/` (last viewed on February 26, 2014).

Germain, F. G. (2011). A nonlinear analysis framework for electronic synthesizer circuits. Master's thesis, McGill University, Montréal, QC, Canada. Available online `http://digitool.library.mcgill.ca/R/?func=dbin-jump-full&object_id=104845&local_base=GEN01-MCG02` (last viewed on February 26, 2014).

Green, D. M. (1967). Additivity of masking. *Journal of the Acoustical Society of America*, 41(6):1517–1525.

Greenwood, D. D. (1961). Auditory masking and the critical band. *Journal of the Acoustical Society of America*, 33(4):484–502.

Hammond, L. (1934). Electronic musical instrument. U.S. patent 1,956,350.

Hélie, T. (2006). On the use of Volterra series for real-time simulations of weakly nonlinear analog audio devices: Application to the Moog ladder filter. In *Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-06)*, pages 7–12, Montréal, QC, Canada.

Hélie, T. (2010). Volterra series and state transformation for real-time simulations of audio circuits including saturations: Application to the Moog ladder filter. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(4):747–759.

Hélie, T. (2011). Lyapunov stability analysis of the Moog ladder filter and dissipativity aspects in numerical solutions. In *Proceedings of the 14th International Conference on Digital Audio Effects (DAFx-11)*, pages 45–52, Paris, France.

Humes, L. E. and Jesteadt, W. (1989). Models of the additivity of masking. *Journal of the Acoustical Society of America*, 85(3):1285–1294.

Huovilainen, A. (2004). Non-linear digital implementation of the Moog ladder filter. In *Proceedings of the 7th International Conference on Digital Audio Effects (DAFx-04)*, pages 61–64, Naples, Italy.

Huovilainen, A. (2010). Design of a scalable polyphony-MIDI synthesizer for a low-cost DSP. Master's thesis, Aalto University School of Science and Technology, Espoo, Finland.

Huovilainen, A. and Välimäki, V. (2005). New approaches to digital subtractive synthesis. In *Proceedings of the 2005 International Computer Music Conference (ICMC'05)*, pages 399–402, Barcelona, Spain.

International Telecommunication Union (1998). Method for objective measurement of perceptual audio quality. Recommendation ITU-R BS.1387.

Ishibashi, M. (1987). Electronic musical instrument. U.S. patent 4,658,691.

Kleimola, J. (2005). Design and implementation of a software sound synthesizer. Master's thesis, TKK Helsinki University of Technology, Espoo, Finland.

Kleimola, J. (2008). Audio synthesis by bitwise logical modulation. In *Proceedings of the 11th International Conference on Digital Audio Effects (DAFx-08)*, pages 67–70, Espoo, Finland.

Kleimola, J. (2013). *Nonlinear Abstract Sound Synthesis Algorithms*. Doctoral dissertation, Aalto University School of Electrical Engineering, Espoo, Finland.

Kleimola, J., Lazzarini, V., Timoney, J., and Välimäki, V. (2010). Phaseshaping oscillator algorithms for musical sound synthesis. In *Proceedings of the 7th Sound and Music Computing Conference (SMC 2010)*, pages 94–101, Barcelona, Spain.

Kleimola, J., Lazzarini, V., Timoney, J., and Välimäki, V. (2011a). Aspects of second-order feedback AM synthesis. In *Proceedings of the 2011 International Computer Music Conference (ICMC'11)*, pages 92–95, Huddersfield, UK.

Kleimola, J., Lazzarini, V., Timoney, J., and Välimäki, V. (2011b). Feedback amplitude modulation synthesis. *EURASIP Journal on Advances in Signal Processing*, 2011, Article ID 434378, 18 pages.

Kleimola, J., Lazzarini, V., Timoney, J., and Välimäki, V. (2011c). Vector phaseshaping synthesis. In *Proceedings of the 14th International Conference on Digital Audio Effects (DAFx-11)*, pages 233–240, Paris, France.

Kleimola, J., Pekonen, J., Penttinen, H., Välimäki, V., and Abel, J. S. (2009). Sound synthesis using an allpass filter chaing with audio-rate coefficient modulation. In *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx-09)*, pages 305–312, Como, Italy.

Kleimola, J. and Välimäki, V. (2012). Reducing aliasing from synthetic audio signals using polynomial transition regions. *IEEE Signal Processing Letters*, 19(2):67–70.

Kreyszig, E. (1999). *Advanced Engineering Mathematics*. John Wiley & Sons, Inc., New York, NY, 8th edition.

Laakso, T. I., Välimäki, V., Karjalainen, M., and Laine, U. K. (1996). Splitting the unit delay — tools for fractional delay filter design. *IEEE Signal Processing Magazine*, 13(1):30–60.

Lane, J. (1990). Pitch detection using a tunable IIR filter. *Computer Music Journal*, 14(3):46–59.

Lane, J., Hoory, D., Martinez, E., and Wang, P. (1997). Modeling analog synthesis with DSPs. *Computer Music Journal*, 21(4):23–41.

Laroche, J. (2007). On the stability of time-varying recursive filters. *Journal of the Audio Engineering Society*, 55(6):460–471.

Lazzarini, V. and Timoney, J. (2010a). New perspectives on distortion synthesis for virtual analog oscillators. *Computer Music Journal*, 34(1):28–40.

Lazzarini, V. and Timoney, J. (2010b). Theory and practice of modified frequency modulation synthesis. *Journal of the Audio Engineering Society*, 58(6):459–471.

Lazzarini, V., Timoney, J., Kleimola, J., and Välimäki, V. (2009a). Five variations on a feedback theme. In *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx-09)*, pages 139–145, Como, Italy.

Lazzarini, V., Timoney, J., and Lysaght, T. (2007). Adaptive FM synthesis. In *Proceedings of the 10th International Conference on Digital Audio Effects (DAFx-07)*, pages 21–26, Espoo, Finland. Revised and extended version published as "The generation of natural-synthetic spectra by means of adaptive frequency modulation" in *Computer Music Journal*,32(2):12–32, Summer 2008.

Lazzarini, V., Timoney, J., Pekonen, J., and Välimäki, V. (2009b). Adaptive phase distortion synthesis. In *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx-09)*, pages 28–35, Como, Italy.

Lazzaro, J. and Wawrzynek, J. (2004). Subtractive synthesis without filters. In

Greenebaum, K. and Barzel, R., editors, *Audio Anecdotes II — Tools, Tips, and Techniques for Digital Audio*, pages 55–63. A. K. Peters Ltd., Wellesley, MA.

Leary, A. B. and Bright, C. T. (2009). Bandlimited digital synthesis of analog waveforms. U.S. patent 7,589,272.

Lisle, R. J. and McDonald, B. S. (1993). Method and apparatus for selective reduction of upper harmonic content in digital synthesizer excitation signals. U.S. patent 5,194,684.

Lowenfels, D. (2003). Virtual analog synthesis with a time-varying comb filter. In *Proceedings of the 115th AES Convention*, New York, NY. Preprint 5960.

Lufti, R. A. (1983). Additivity of simultaneous masking. *Journal of the Acoustical Society of America*, 73(1):262–267.

Lufti, R. A. (1985). A power-law transformation predicting masking by sounds with complex spectra. *Journal of the Acoustical Society of America*, 77(6):2128–2136.

Martenot, M. L. E. (1931). Electronic musical instrument. U.S. patent 1,824,402.

Massie, D. C. (1998). Wavetable sampling synthesis. In Kahrs, M. and Brandenburg, K., editors, *Applications of Digital Signal Processing to Audio and Acoustics*, pages 311–341. Kluwer Academic Publishers, Norfolk, MA.

Moog, R. A. (1964). Voltage-controlled electronic music modules. In *Proceedings of the 16th AES Convention*, New York, NY. Preprint 346. Reprinted in *Journal of the Audio Engineering Society*, 13(3):200–206, July 1965.

Moore, B. C. J. (1997). *An Introduction to the Psychology of Hearing*, pages 49–53, 93–94, and 103–114. Academic Press, London, UK, 4th edition.

Moore, F. R. (1990). *Elements of Computer Music*, pages 44–48. Prentice-Hall, Englewood Cliffs, NJ.

Moorer, J. A. (1976). The synthesis of complex audio spectra by means of discrete summation formulas. *Journal of the Audio Engineering Society*, 24(9):717–727. Published also as "The synthesis of complex audio spectra by means of discrete summation formulae," report STAN-M-5,Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, CA, December 1975.

Nam, J., Välimäki, V., Abel, J. S., and Smith, J. O. (2009). Alias-free virtual analog oscillators using feedback delay loops. In *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx-09)*, pages 347–352, Como, Italy.

Nam, J., Välimäki, V., Abel, J. S., and Smith, J. O. (2010). Efficient antialiasing oscillator algorithms using low-order fractional delay filters. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(4):773–785.

Olson, H. F., Belar, H., and Timmens, J. (1955). Electronic music synthesis. *Journal of the Acoustical Society of America*, 27(3):595–612.

Parker, J. and D'Angelo, S. (2013). A digital model of the Buchla lowpass-gate. In *Proceedings of the 16th International Conference on Digital Audio Effects (DAFx-13)*, pages 278–285, Maynooth, Ireland.

Pekonen, J. (2007). Computationally efficient music synthesis — methods and sound design. Master's thesis, TKK Helsinki University of Technology, Espoo, Finland.

Pekonen, J. (2008). Coefficient-modulated first-order allpass filter as distortion effect. In *Proceedings of the 11th International Conference on Digital Audio Effects (DAFx-08)*, pages 83–87, Espoo, Finland.

Pekonen, J. and Välimäki, V. (2011). The brief history of virtual analog synthesis. In *Proceedings of the 6th Forum Acusticum*, pages 461–466, Aalborg, Denmark.

Pekonen, J., Välimäki, V., Abel, J. S., and Smith, J. O. (2009). Spectral delay filters with feedback and time-varying coefficients. In *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx-09)*, pages 157–164, Como, Italy.

Pekonen, J., Välimäki, V., Nam, J., Abel, J. S., and Smith, J. O. (2010). Variable fractional delay filters in bandlimited oscillator algorithms for music synthesis. In *Proceedings of the 2010 International Conference on Green Circuits and Systems (ICGCS2010)*, pages 148–153, Shanghai, China.

Puckette, M. (2007). *The Theory and Technique of Electronic Music*, chapter 10.4, pages 307–313. World Scientific Publishing Co., Hackensack, NJ.

Regalia, P. A. and Mitra, S. K. (1987). Tunable digital frequency response equalization filters. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(1):118–120.

Roads, C. (1998). *The Computer Music Tutorial*. The MIT Press, Cambridge, MA.

Rodet, X. (1984). Time-domain formant-wave-function synthesis. *Computer Music Journal*, 8(3):9–14.

Rodet, X. and Depalle, P. (1992). Spectral envelopes and inverse FFT synthesis. In *Proceedings of the 93rd AES Convention*, Paris, France. Preprint 3393.

Schimmel, J. (2012). Audible aliasing distortion in digital audio synthesis. *Radioengineering*, 21(1):56–62.

Schoffhauzer, P. (2007). Synthesis of quasi-bandlimited analog waveforms using frequency modulation. Online publication, available at `http://scp.web.elte.hu/papers/synthesis1.pdf` (last viewed on February 26, 2014).

Smith, J. O. (1996). Physical modeling synthesis update. *Computer Music Journal*, 20(2):44–56.

Smith, J. O. (2012). Signal processing libraries for FAUST. In *Proceedings of the 2012 Linux Audio Conference (LAC 2012)*, pages 153–161, Stanford, CA.

Stilson, T. (2006). *Efficiently-Variable Non-Oversampling Algorithms for Virtual-Analog Music Synthesis — A Root-Locus Perspective*. Doctoral dissertation, Stanford University, Stanford, CA.

Stilson, T. and Smith, J. O. (1996a). Alias-free digital synthesis of classic analog waveforms. In *Proceedings of the 1996 International Computer Music Conference (ICMC'96)*, pages 332–335, Hong Kong, China.

Stilson, T. and Smith, J. O. (1996b). Analyzing the Moog VCF with considerations for digital implementation. In *Proceedings of the 1996 International Computer Music*

Conference (ICMC'96), pages 398–401, Hong Kong, China.

Stinchcombe, T. E. (2008). Analysis of the Moog transistor ladder and derivative filters. Online publication, available at `http://www.timstinchcombe.co.uk/synth/Moog_ladder_tf.pdf` (last viewed on February 26, 2014).

Tassart, S. (2013a). Band-limited impulse train generation using sampled infinite impulse responses of analog filters. *IEEE Trans. Audio, Speech, and Lang. Process.*, 21(3):488–497.

Tassart, S. (2013b). Graphical equalization using interpolated filter banks. *Journal of the Audio Engineering Society*, 61(5):263–279.

Théremin, L. S. (1925). Method of and apparatus for the generation of sounds. U.S. patent 1,661,058.

Timoney, J., Lazzarini, V., Carty, B., and Pekonen, J. (2009a). Phase and amplitude distortion methods for digital synthesis of classic analogue waveforms. In *Proceedings of the 126th AES Convention*, Munich, Germany. Preprint 7792.

Timoney, J., Lazzarini, V., Hodgkinson, M., Kleimola, J., Pekonen, J., and Välimäki, V. (2012). Virtual analog oscillator hard synchronisation: Fourier series and an efficient implementation. In *Proceedings of the 15th International Conference on Digital Audio Effects (DAFx-12)*, pages 217–224, York, UK.

Timoney, J., Lazzarini, V., and Lysaght, T. (2008). A modified FM synthesis approach to bandlimited signal generation. In *Proceedings of the 11th International Conference on Digital Audio Effects (DAFx-08)*, pages 27–33, Espoo, Finland.

Timoney, J., Lazzarini, V., Pekonen, J., and Välimäki, V. (2009b). Spectrally rich phase distortion sound synthesis using an allpass filter. In *Proceedings of the 2009 IEEE International Conference in Acoustics, Speech, and Signal Processing (ICASSP'09)*, pages 293–296, Taipei, Taiwan.

Tomisawa, N. (1981). Tone production method for an electronic musical instrument. U.S. patent 4,249,447.

Välimäki, V. (1995). *Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters*. Doctoral dissertation, report no. 37, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland.

Välimäki, V. (2005). Discrete-time synthesis of the sawtooth waveform with reduced aliasing. *IEEE Signal Processing Letters*, 12(3):214–217.

Välimäki, V. and Huovilainen, A. (2006). Oscillator and filter algorithms for virtual analog synthesis. *Computer Music Journal*, 30(2):19–31.

Välimäki, V. and Huovilainen, A. (2007). Antialiasing oscillators in subtractive synthesis. *IEEE Signal Processing Magazine*, 24(2):116–125.

Välimäki, V., Nam, J., Smith, J. O., and Abel, J. S. (2010). Alias-suppressed oscillators based on differentiated polynomial waveforms. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(4):786–798.

Välimäki, V., Pakarinen, J., Erkut, C., and Karjalainen, M. (2006). Discrete-time modelling of musical instruments. *Reports on Progress in Physics*, 69(1):1–78.

Wegel, R. L. and Lane, C. E. (1924). The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear. *Physical Review*, 23(2):266–285.

Winham, G. and Steiglitz, K. (1970). Input generators for digital sound synthesis. *Journal of the Acoustical Society of America*, 47(2B):665–666.

Zambon, S. and Fontana, F. (2011). Efficient polynomial implementation of the EMS VCS3 filter. In *Proceedings of the 14th International Conference on Digital Audio Effects (DAFx-11)*, pages 287–290, Paris, France.

Zwicker, E. and Fastl, H. (1990). *Psychoacoustics — Facts and Models*, pages 15–19 and 56–102. Springer-Verlag, Berlin, Germany.

Errata

Publications I, II, and VII have an error in their first reference. The referred paper has three authors (Olson et al. [1955]), not two as in the aforementioned publications.

Publication V

The first expression of Equation (2) was originally presented by Lane [1990], and it was first utilized by Lane et al. [1997] in virtual analog oscillator algorithms. The article misses citations to those publications.

Publication VII

The caption of Figure 11 refers erroneously to the third-order B-spline BLEP method. The waveforms plotted in Figures 11(e) and (f) are obtained from the fourth-order B-spline BLEP algorithm. The article text refers to the correct algorithm.

Aalto-DD 26/2014

Digital modeling of the subtractive sound synthesis principle used in analog synthesizers has been a popular research topic in the past few years. In subtractive sound synthesis, a spectrally rich oscillator signal is filtered with a time-varying filters. The trivial digital implementation of the oscillator waveforms typically used in this synthesis method suffers from disturbing aliasing distortion. This thesis presents efficient filter-based algorithms that produce these waveforms with reduced aliasing. In addition, perceptual aspects of audibility of aliasing and modeling of analog synthesizer oscillator output signals are addressed.

ISBN 978-952-60-5588-6
ISBN 978-952-60-5586-2 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934
ISSN 1799-4942 (pdf)

Aalto University
School of Electrical Engineering
Department of Signal Processing and Acoustics
www.aalto.fi

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

DOCTORAL
DISSERTATIONS