PULSETABLE SYNTHESIS OF WIND INSTRUMENT TONES

Christian Dittmar

Fraunhofer Institute for Integrated Circuits
Erlangen, Germany
christian.dittmar@iis.fraunhofer.de

Simon Schwär, Manuel Peters, Stefan Balke, Meinard Müller *

International Audio Laboratories Erlangen
Erlangen, Germany
meinard.mueller@audiolabs-erlangen.de

ABSTRACT

We revisit pulsetable synthesis, an efficient technique for generating plausible and expressive wind instrument tones. Based on the principles of pulse forming theory, this method models sound production as the periodic repetition of shaped pulses characterizing the target instruments' spectral envelope. In this approach, single-cycle waveforms, referred to as pulses, are stored in pulsetables indexed by their corresponding fundamental frequency. During synthesis, the pulses are read from these tables to form a periodic waveform, which is further shaped by time-varying low-pass filtering, amplification, and reverberation. These processes are guided by control signal contours that describe how fundamental frequency, brightness, and loudness evolve over time. Through case studies with real-world wind instrument recordings, we show how the interplay between these control signals gives rise to articulations such as attack transients, vibrato, and growl. Finally, we discuss the potential of this framework for integration into Differentiable Digital Signal Processing (DDSP) models, where neural networks could learn synthesis parameters directly from training data.

1. INTRODUCTION

Wind instruments from the brass, single-reed, and double-reed families share a common sound production mechanism in which stable tones with fundamental frequency F_0 and period duration $T_0 = {F_0}^{-1}$ emerge from the interaction between lip or reed excitation at the mouthpiece and the tube resonances of the instrument body into which the mouthpiece feeds. The effective tube length can be varied by applying different fingerings and helps to stabilize oscillations close to its resonant frequencies. By adjusting airflow and embouchure, players can further modify F_0 drastically (e.g., locking into higher harmonics of the tube resonances) or subtly (e.g., vibrato, pitch bending). The periodic opening and closing of lips or reeds produces pulses of width $\tau \ll T_0$. For many wind instruments, the pulse width stays nearly constant across a wide range of F_0, while the pulse shape varies from smooth to spiky depending on the applied force (see Fig. 2). Consequently, wind instrument tones exhibit clearly harmonic spectra whose brightness increases with dynamic level. Due to the continuous and nuanced control that players exert over their instrument, it is quite challenging to synthesize realistic wind instrument tones.

In this demo paper, we revisit pulsetable synthesis [1, 2], which shares a common foundation with pulse forming synthesis [3], formant wave synthesis [4], and pulsar synthesis [5]. Since the lineage of pulse-based synthesis techniques is comparatively obscure, we highlight in Sec. 2 connections to more well-established approaches in their scientific-historical context. In Sec. 3, we provide a brief overview of the method, focusing on the validity of the basic assumptions and suitability for wind instrument synthesis. Finally, in Sec. 4 we discuss the potential and remaining challenges of integration with Differentiable Digital Signal Processing (DDSP) paradigms.

2. RELATED WORK

In the 1960s, Frans Fransson simulated the pulse-like opening and closing of the double-reeds in the bassoon with electronic circuits and was able to generate spectra similar to those measured from acoustic instrument recordings [6]. A decade later, Jobst Fricke and Wolfgang Voigt came up with the basic principle of pulse forming synthesis [7, 3]: Given a constant pulse width τ and variable period duration T_0, the spectral envelope remains approximately constant during F_0 changes (see Fig. 2). Their findings laid the foundation for the analogue electronic wind instruments Martinetta[1] and Variophon[2]. In order to achieve natural articulations, these instruments constantly measured the players' blowing pressure and derived control signals for slight variations of pitch, brightness and loudness. Michael Oehler and Christoph Reuter investigated this approach in detail and coined the term micro-modulations [8, 9, 10]).

From the 1980s on, physical modeling of wind instruments became attractive due to the advances in digital signal processing [11, 12]. Notably, Ron W. Berry showed that it is also possible to implement physical models of wind instruments with analogue circuits [13]. Hardware instruments specialized on physical modeling, such as Yamaha's VL1[3] became commercially available in the 1990s. Today, it is mainly software plug-ins such as SWAM[4] and Pianoteq[5] that offer this type of synthesis. There is often a vast array of model parameters whose complex interplay influences the timbre of the resulting tones. In practice, it is not always clear how intuitive control parameters (e.g., blowing pressure) should be mapped to those model parameters [14] in the sense of micro-modulations.

As an alternative, data-driven approaches for wind instrument synthesis emerged from the 1990s on. Basic techniques such as

* The International Audio Laboratories Erlangen are a joint institution of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) and Fraunhofer Institute for Integrated Circuits IIS.

[1] https://muwiserver.univie.ac.at/martinetta
[2] http://www.variophon.de/vario_e.htm
[3] https://www.soundonsound.com/reviews/yamaha-vl1
[4] https://audiomodeling.com/swam-engine/
[5] https://www.modartt.com/pianoteq_overview

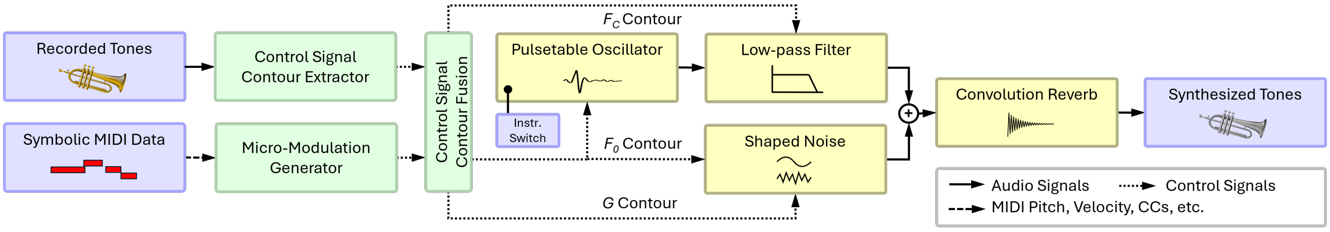

Figure 1: System overview: Blue boxes denote data input and output, green boxes relate to control signals and yellow are DSP blocks. The different color of the input and output instrument indicates that timbral alterations might be introduced during resynthesis, which can be exploited for deliberate timbre transfer [1].

wavetable synthesis [15], spectral and additive synthesis [16, 17, 18] were combined with the retrieval of sound or control signal snippets in larger corpora of instrument recordings. Today, some commercial software plug-ins (SampleModeling[6], AcousticSamples[7]) use variants of pitch-synchronous overlap-add (PSOLA [19]) in a similar fashion. Notably, the wind instrument models by Joel Blanco Berg[8], use a wavetable oscillator and low-pass filter quite similar to the principle proposed 30 years ago by Horner and Beauchamp [15].

In recent years, generative neural models in conjunction with modern training paradigms have achieved breakthrough results in the synthesis of music signals from textual descriptions [20] or symbolic music representations [21, 22]. The more specific task of wind instrument synthesis has been successfully tackled by DDSP-based approaches [23, 24]. We would like to highlight that the integration of additive synthesis with neural architectures had already been proposed 20 years earlier by Wessel et al. [17]. Harking back to the notion of micro-modulations, both Wessel and the DDSP authors found that a reduction to the three control parameters fundamental frequency, brightness and loudness covers a wide range of typical articulations found in instrument recordings. We applied the same set of control signals to pulsetable synthesis [1] and also integrated the method into a DDSP framework [2].

3. PULSETABLE SYNTHESIS

Pulsetable synthesis is based on concatenating short prototype waveforms at regular intervals [1], approximating the excitation in certain wind instruments. Variations in intensity are emulated through time-variant low-pass filtering [15]. Additive shaped noise models the players' breath-driven airflow, while a final convolution stage imprints the acoustic environment [2].

3.1. Pulse Shapes

As discussed in [3, 5, 9], a single-period pulse encodes the instantaneous spectral envelope, which correlates with the timbre of the resulting tone. In Fig. 2(a), we show three trumpet tones from the University of Iowa Musical Instrument Samples (MIS)[9] database, recorded in anechoic conditions. For visualization purposes, we selected pitches that are spaced three octaves apart, the lowest being B♭3 ($F_0 \approx 233$ Hz, $T_0 \approx 4.3$ ms). In the upper half of the plot, the tones are played at highest intensity (fortissimo, ***ff***), in the lower half they are played at lowest intensity (pianissimo, ***pp***), which is clearly visible by the different extent of their amplitude envelopes. In Fig. 2(b), we zoom into the waveforms at the level of single-period pulses. The bold curves show their pulse shapes, the semi-transparent curves show how the original waveform continues by periodic repetition. Note that the main energy is concentrated at the initial impulse (up until $\tau \approx 1.5$ ms), whose shape stays relatively constant throughout the octaves.

It can also be seen that the pulse shapes of the ***pp*** tones appear like smoothed, less spiky versions of the ***ff*** pulse shapes. This is supported by the spectrum plots in Fig. 2(c), where we overlay the bold curve of the spectral envelope with the semi-transparent spectrum corresponding to the steady-state oscillation of the tones. For better visibility, we use a pitch-scaled frequency axis with ticks indicating the pitches corresponding to the F_0 of the three selected tones. It can clearly be seen that the ***ff*** spectra exhibit more pronounced upper harmonics in comparison to the ***pp*** spectra. This corroborates the previous findings that intensity and brightness are proportional and low-pass filtering can effectively emulate the transition from high to low intensity [15].

3.2. System Overview

The core DSP blocks of pulsetable synthesis are depicted as yellow boxes in Fig. 1 and have already been described in [1]. The pulsetable oscillator generates the pulse train, selecting the appropriate pulse shape according to the target F_0. As we have shown in Sec. 3.1, some wind instruments (especially from the brass family) exhibit almost no change in pulse shape across their entire pitch range. In those cases, it is not strictly necessary to populate the pulsetable with one pulse shape for each chromatic pitch. Instead, we can reduce the number of entries, in the extreme case to just a single pulse shape. The time-variant low-pass filter which is applied to the pulse train is vital for emulating tones of lower intensity. As we will show in Sec. 3.3, micro-modulations of the low-pass filter cutoff F_c also give rise to typical articulations, such as attack transients.

Going beyond the original proposal from [1], we introduce a shaped noise component [23, 2] to emulate the turbulent, intermittent airstream running through the tube resonator. More precisely, bandpass filtered white noise is amplitude-modulated with a pulse train of quickly decaying, sawtooth-shaped pulses that are synchronized to the pulsetable oscillator signal. Although this ad-hoc design is not based on a systematic analysis of real-world recordings, we still found it beneficial for adding a certain "raspiness" to the pulsetable tones, especially in the case of low intensity (i.e., stronger low-pass filtering) and low pitched notes. The final reverb block is then used to give a more realistic impression of the

[6]https://www.samplemodeling.com/products/
[7]https://www.acousticsamples.net/Bundles/vhornsbundles
[8]https://lydoel.gumroad.com/
[9]https://theremin.music.uiowa.edu/MIS.html

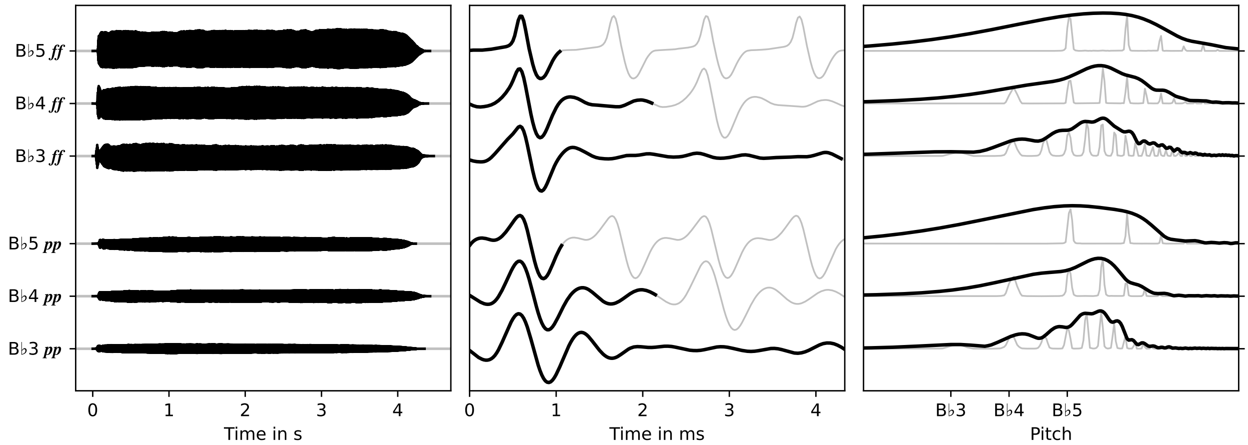

Figure 2: Real-world pulse shapes of trumpet tones B♭3, B♭4, B♭5, recorded in anechoic conditions, at dynamic levels ***pp*** and ***ff***. The subfigures show: **(a)** Waveform of all tones, entire duration from onset to offset. **(b)** Zoom into the waveforms, magnifying the period duration of B♭3 ($T_0 \approx 4.3$ ms). **(a)** Corresponding spectral envelope (bold black) and harmonic spectra (light gray).

instrument being played within an acoustic environment (e.g., a rehearsal room or concert hall). While arbitrary impulse responses could be applied here, we also made use of learned room impulse responses that had been jointly estimated with core waveform generation parameters in a DDSP fashion [23, 2].

3.3. Control Signals for Micro-Modulations

Wind instrument players exert continuous control over pitch, brightness and loudness, with transitions between notes strongly influenced by their context. In contrast to that, the core DSP blocks described in Sec. 3.2 would only generate static tones at a certain pitch and brightness level that sound completely synthetic and inanimate. In order to mimic expressive articulations and to achieve plausible synthesis results, subtle and interdependent micro-modulations need to be applied to the control signal contours of fundamental frequency F_0, low-pass cutoff F_c, and overall gain G. To support this, we refer to the perceptual listening tests with synthesized bassoon and oboe tones in [8, 9]. For both instruments, test listeners rated stimuli with vibrato-related micro-modulations as most natural.

In this demo paper, we indicate the system components related to micro-modulations as green boxes in Fig. 1. By means of tracking instantaneous frequency and magnitude of selected harmonics, we can extract the trajectories of F_0, F_c and G from real-world recordings. We refer to our accompanying webpage[10] for a more detailed description of the underlying feature extraction procedure. In addition, we can generate control signal trajectories in a rule-based fashion from an aligned symbolic transcription (e.g., MIDI). Fusing both paths together, we can combine and manipulate control signal contours (e.g., remove fluctuations, transpose pitch) as shown in Fig. 3. There, we illustrate a typical phenomenon related to note onsets on brass instruments: rapid modulation of all three control signal contours produces inharmonic side-bands similar to those in amplitude or frequency modulation. Since they only occur for a very short duration, they are perceived as attack transients. Deliberately removing those micro-modulations leads to an overly soft attack as shown in Fig. 3(b). Further examples and details are provided on our accompanying website[10].

4. CONCLUSIONS AND OUTLOOK

In this demo paper, we showed how pulsetable synthesis can efficiently generate lively and natural wind instrument tones. A remaining difficulty is the robust, automatic extraction of suitable pulse shapes and control signal contours from real-world recordings, especially under reverberant conditions. Major parts of the pulsetable method were successfully integrated into a DDSP-like framework in [2]. This served as a proof-of-concept that pulse shapes can be learned from un-annotated recordings in a data-driven manner. Future work should focus on extending that setup by means of fully differentiable, time-variant low-pass and band-pass filtering for the pulse train, respectively shaped noise. Not only is this approach more efficient than the bank of sinusoidal oscillators and band-pass filters used in DDSP [23], it could also enable us to jointly learn all relevant control signal contours directly from real-world instrument recordings.

5. ACKNOWLEDGMENTS

This work was partly funded by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) under grant numbers 500643750 (MU 2686/15-1) and 555525568 (MU 2686/18-1). We would like to thank Christoph Reuter as well as Roger Dannenberg for insightful discussions and valuable comments that helped us situate the literature.

6. REFERENCES

[1] C. Dittmar, J. Zeitler, S. Balke, S. Schwär, and M. Müller, “Pulse-it: Lightweight and expressive synthesis of wind instrument playing,” in *Late-Breaking Demos of the Intl. Society for Music Information Retrieval Conf. (ISMIR)*, Daejeon, Republic of Korea, 2025.

[2] S. Schwär, C. Dittmar, S. Balke, and M. Müller, “Differentiable pulsetable synthesis for wind instrument modeling,”

[10] `https://www.audiolabs-erlangen.de/resources/MIR/2026_DittmarSBM_PulsetableSynthesis_DAFx`

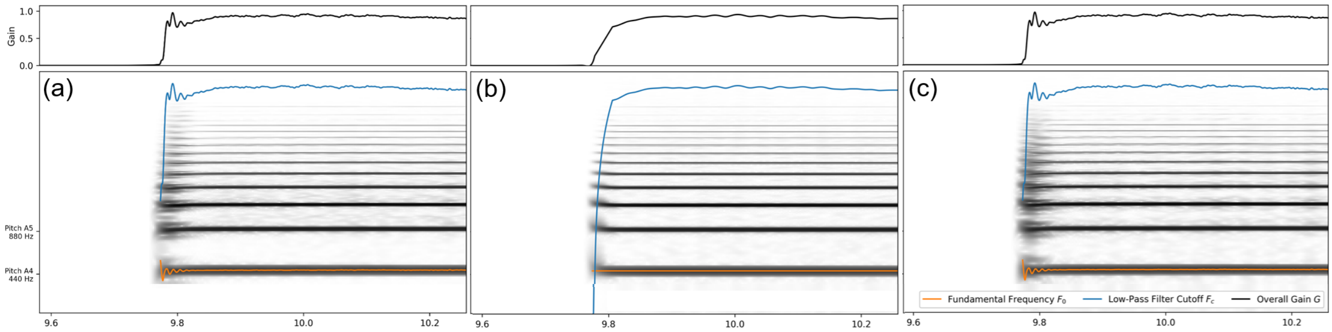

Figure 3: Visualization of the interplay between the control signal contours F_0, F_c and G during the attack transient around 9.8 seconds in **(a)** real-world, anechoic recording of a B♭4 trumpet tone , **(b)** pulsetable synthesis with flattened and smoothed control signal contours, and **(c)** pulsetable synthesis with unmodified control signal contours.

in *Proc. of the IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Barcelona, Spain, 2026, pp. 14 792–14 796.

[3] J. P. Fricke, “Formantbildende Impulsfolgen bei Blasinstrumenten,” in *Fortschritte der Akustik: Plenarvorträge und Kurzreferate der Tagung der Deutschen Arbeitsgemeinschaft für Akustik (DAGA)*, Braunschweig, Germany, 1975, pp. 407–411.

[4] X. Rodet, “Time-domain formant-wave-function synthesis,” *Computer Music Journal*, vol. 8, no. 3, pp. 9–14, 1984.

[5] C. Roads, “Introduction to pulsar synthesis,” *The Journal of the Acoustical Society of America*, vol. 109, no. 5_Supplement, pp. 2401–2401, 2001.

[6] F. Fransson, “The source spectrum of double-reed woodwind instruments,” *Speech Transmission Laboratory Quarterly Progress and Status Report*, vol. 7, no. 4, pp. 35–37, 1966.

[7] W. Voigt, “Untersuchungen zur Formantbildung in Klängen von Fagott und Dulzianen,” Doctoral Dissertation, Universität Köln, 1975.

[8] M. Oehler and C. Reuter, “Vibrato experiments with bassoon sounds by means of the digital pulse forming synthesis and analysis framework,” in *Proc. of the Audio Engineering Socienty (AES) Convention*, New York, USA, 2007.

[9] M. Oehler, “Die digitale Impulsformung als Werkzeug für die Analyse und Synthese von Blasinstrumentenklängen,” Doctoral Dissertation, Hochschule für Musik und Theater Hannover, 2007.

[10] C. Reuter and H. Schefczik, “A Martinetta sounds better – 50 years of pulse forming synthesis,” in *Proc. of the Joint Conf. DAS|DAGA Annual Conf. on Acoustics*, 2025, pp. 664–667.

[11] K. Karplus and A. Strong, “Digital synthesis of plucked-string and drum timbres,” *Computer Music Journal*, vol. 7, no. 2, pp. 43–55, 1983.

[12] D. H. Keefe, “Physical modeling of wind instruments,” *Computer Music Journal*, vol. 16, no. 4, pp. 57–73, 1992.

[13] R. W. Berry, “Experiments in computer controlled acoustic modelling (A step backwards??),” in *Proc. of the Intl. Computer Music Conf. (ICMC)*, Cologne, Germany, September 1988, pp. 334–348.

[14] W. D’haes and X. Rodet, “A new estimation technique for determining the control parameters of a physical model of a trumpet,” in *Proc. of the Intl. Conf. on Digital Audio Effects (DAFx)*, London, UK, 2003.

[15] A. Horner and J. Beauchamp, “Synthesis of trumpet tones using a wavetable and a dynamic filter,” *Journal of the Audio Engineering Society*, vol. 43, no. 10, pp. 799–812, 1995.

[16] I. Derenyi and R. B. Dannenberg, “Synthesizing trumpet performances,” in *Proc. of the Intl. Computer Music Conf. (ICMC)*, Ann Arbor, Michigan, USA, 1998, pp. 490–496.

[17] D. Wessel, C. Drame, and M. Wright, “Removing the time axis from spectral model analysis-based additive synthesis: Neural networks versus memory-based machine learning,” in *Proc. of the Intl. Computer Music Conf. (ICMC)*, Ann Arbor, Michigan, USA, 1998, pp. 62–65.

[18] N. Hu, “Automatic construction of synthetic musical instruments and performers,” Doctoral Dissertation, Carnegie Mellon University, 2018.

[19] U. Zölzer, Ed., *DAFx: Digital Audio Effects*, 2nd ed. Chichester, UK: John Wiley & Sons, 2011.

[20] J. Copet, F. Kreuk, I. Gat, T. Remez, D. Kant, G. Synnaeve, Y. Adi, and A. Défossez, “Simple and controllable music generation,” in *Proc. of the Annual Conf. on Neural Information Processing Systems (NeurIPS)*, New Orleans, LA, USA, 2023, pp. 47 704–47 720.

[21] E. Cooper, X. Wang, and J. Yamagishi, “Text-to-speech synthesis techniques for midi-to-audio synthesis,” in *Proc. of the ISCA Speech Synthesis Workshop (SSW)*, Budapest, Hungary, 2021, pp. 130–135.

[22] B. Maman, J. Zeitler, M. Müller, and A. H. Bermano, “Multi-aspect conditioning for diffusion-based music synthesis: Enhancing realism and acoustic control,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 33, pp. 68–81, 2025.

[23] J. H. Engel, L. H. Hantrakul, C. Gu, and A. Roberts, “DDSP: Differentiable digital signal processing,” in *Proc. of the Intl. Conf. on Learning Representations (ICLR)*, 2020.

[24] Y. Wu, E. Manilow, Y. Deng, R. Swavely, K. Kastner, T. Cooijmans, A. C. Courville, C. A. Huang, and J. H. Engel, “MIDI-DDSP: detailed control of musical performance via hierarchical modeling,” in *Proc. of the Intl. Conf. on Learning Representations (ICLR)*, 2022.