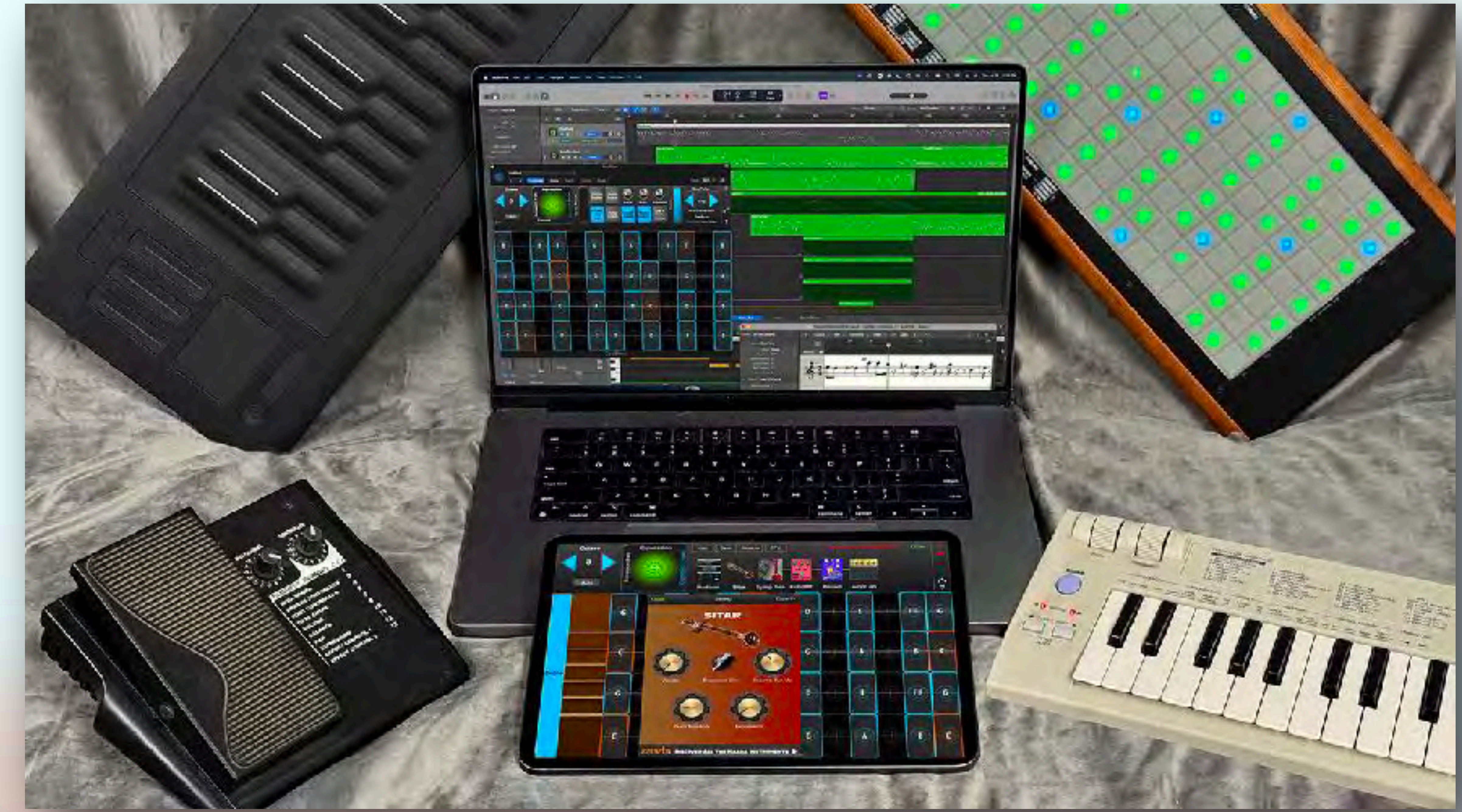
Disrupting the Compositional Workflow

From Sample Libraries to Expressive AI-Calibrated Physical Models

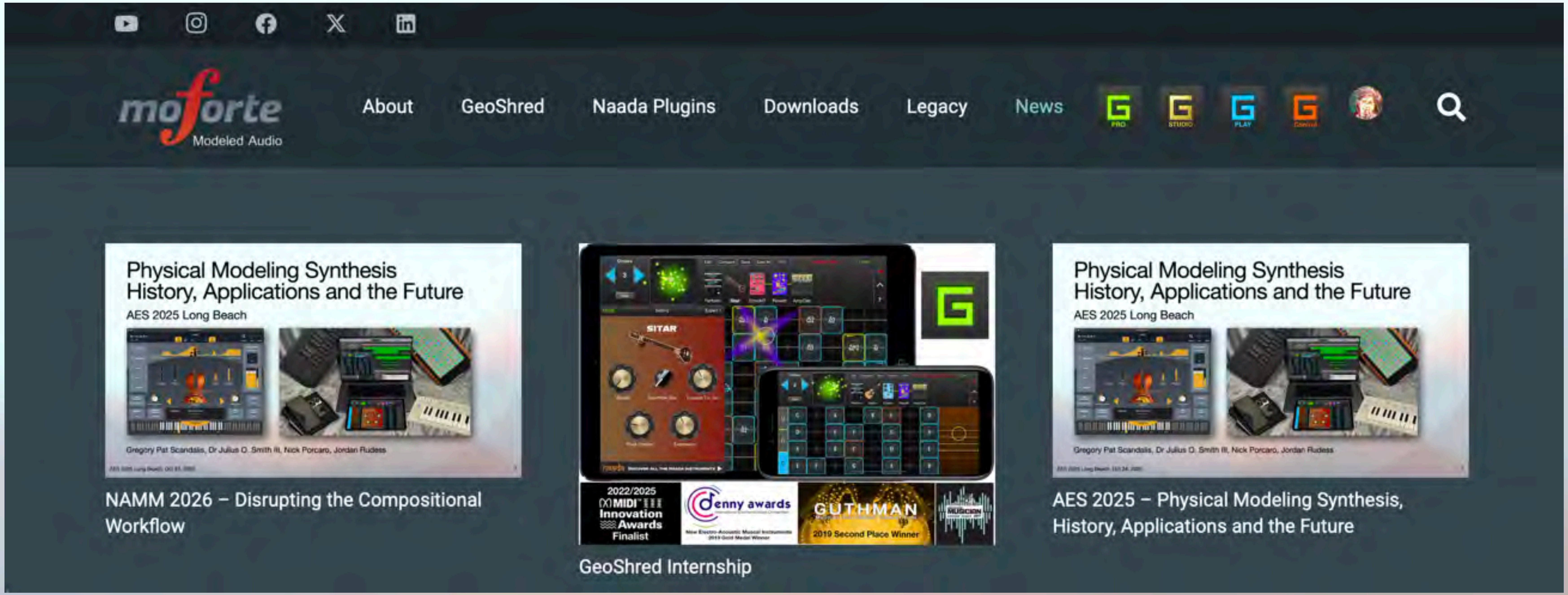
A3E Workshop/NAMM 2026 - Jan 23, 2026



Gregory Pat Scandalis, Dr Julius O. Smith III, Nick Porcaro, Jordan Rudess

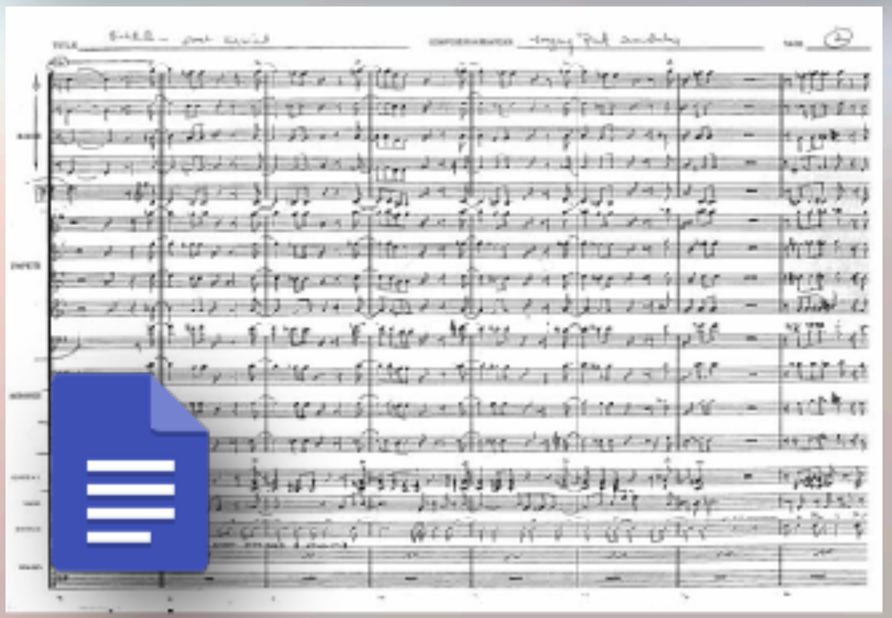
This Presentation Can be Found at:

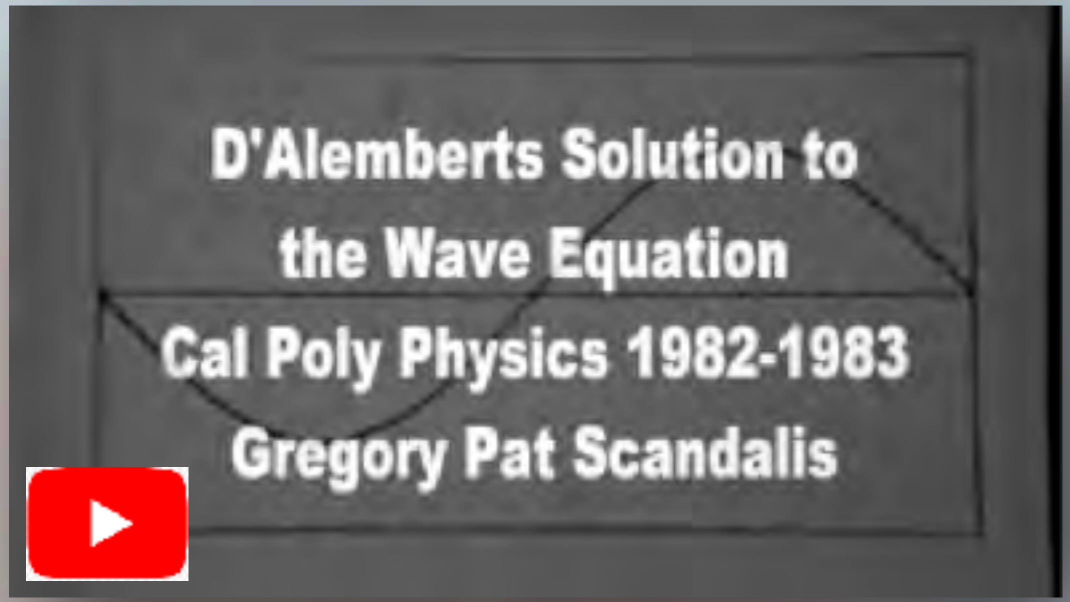
http://www.moforte.com/news



About Pat

- 43 years in the Silicon Valley as an Engineer
- Built my first monophonic electronic instrument from a Radio Shack P-Box kit in 1970
- Gigged with an Arp Avatar guitar synth (1978)
- Guitar Player for Weird Al Band (1980)
- Computer modeling of vibrating strings and membranes for senior thesis in Physics (1982)
- Big Band arranging for Cal Poly Jazz Band (1982/83)
- Researcher in Physical Modeling at Stanford/CCRMA (1994)
- Ran Liquid Audio, first Legal Music Download Company
- CEO/CTO of moForte
- Chairman of the MPE Subcommittee MIDI Association, Co-chair IASIG AIWG





CCRMA

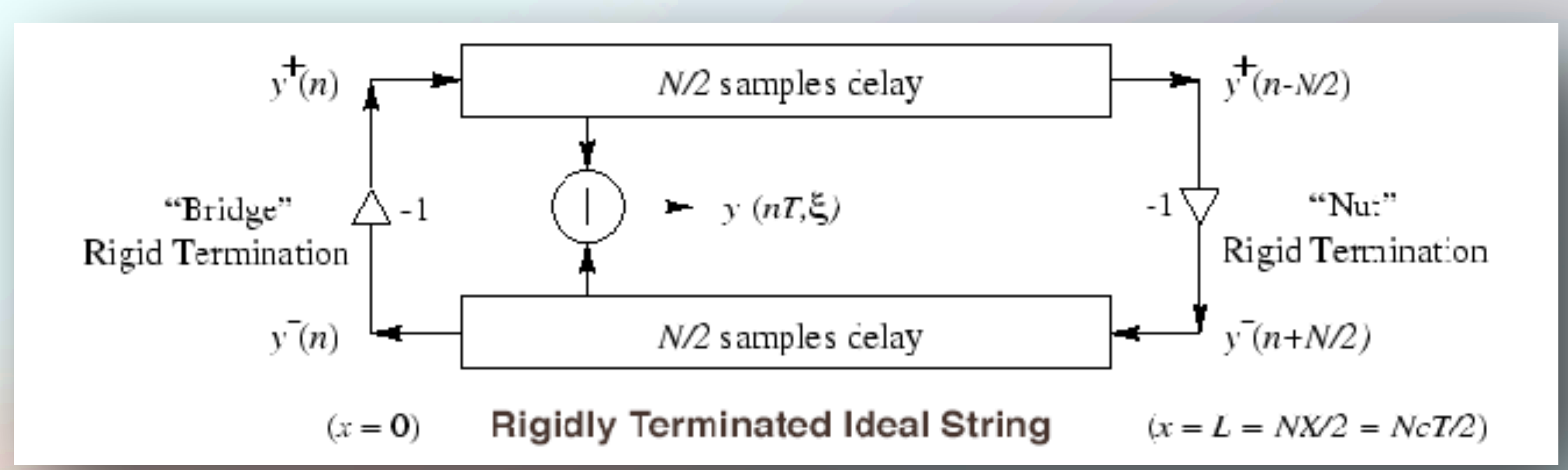
Orchestral Sample Libraries

- Orchestral Sample Libraries are for the most part static recordings.
- Expression is simulated with Key Switches
- Composers/Orchestrators chose versions of Samples instead of expressively performing an instrument

A Shift

Emerging AI technologies make it possible to **calibrate** expressive, physical models to specific instruments which can be used in as an alternative to orchestral sample libraries.

- Physical Models will be **calibrated** to specific instruments
- Replace Key Switches with model controls that respond like the real instruments
- **Keep the human in the compositional and performance workflow**



Outline

- A Brief History of Physical Modeling Synthesis
- Demos of Physical Models
- MPE is an Enabling Technology For Composition/Orchestration
- Full Musical Ensembles Using Physical Models
- The NEAR Future - AI Calibration, MIDI 2, Orchestral Articulation, Virtual Performers
- Questions!

Physical Modeling Collaborators

wizdom music

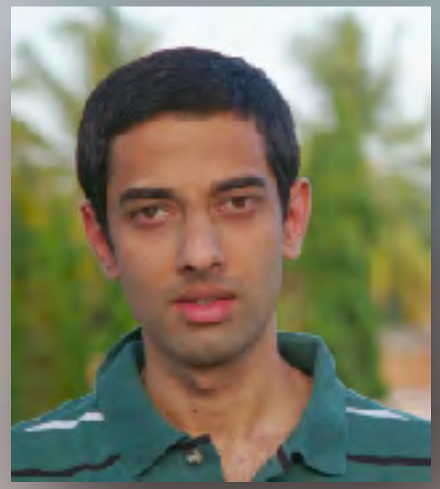
moforte

Audio Modeling

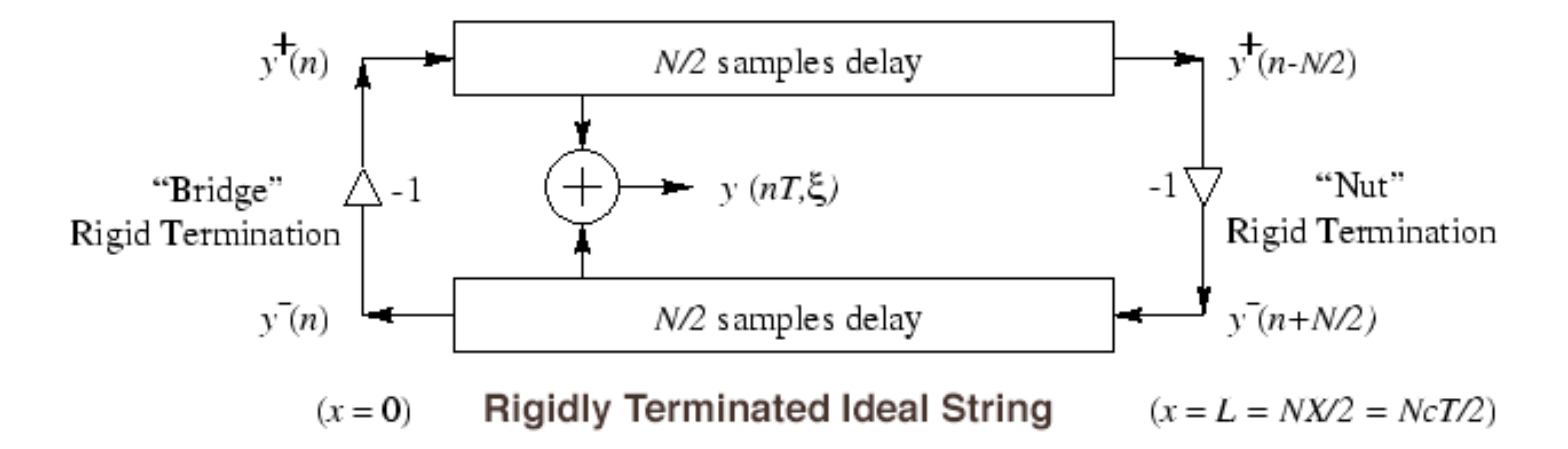
GeoShred is a collaboration between Rock Star and mobile music innovator Jordan Rudess, Stanford/CCRMA Professor Dr. Julius O. Smith III, Nick Porcaro, Pat Scandalis

Additional models developed by Audio Modeling/ SWAM (Stefano Lucato, Lele Parravicini) and AccelMatrix/Naada (Suthambhara Nagaraj)

naada



A Brief History of Physical Modeling Synthesis

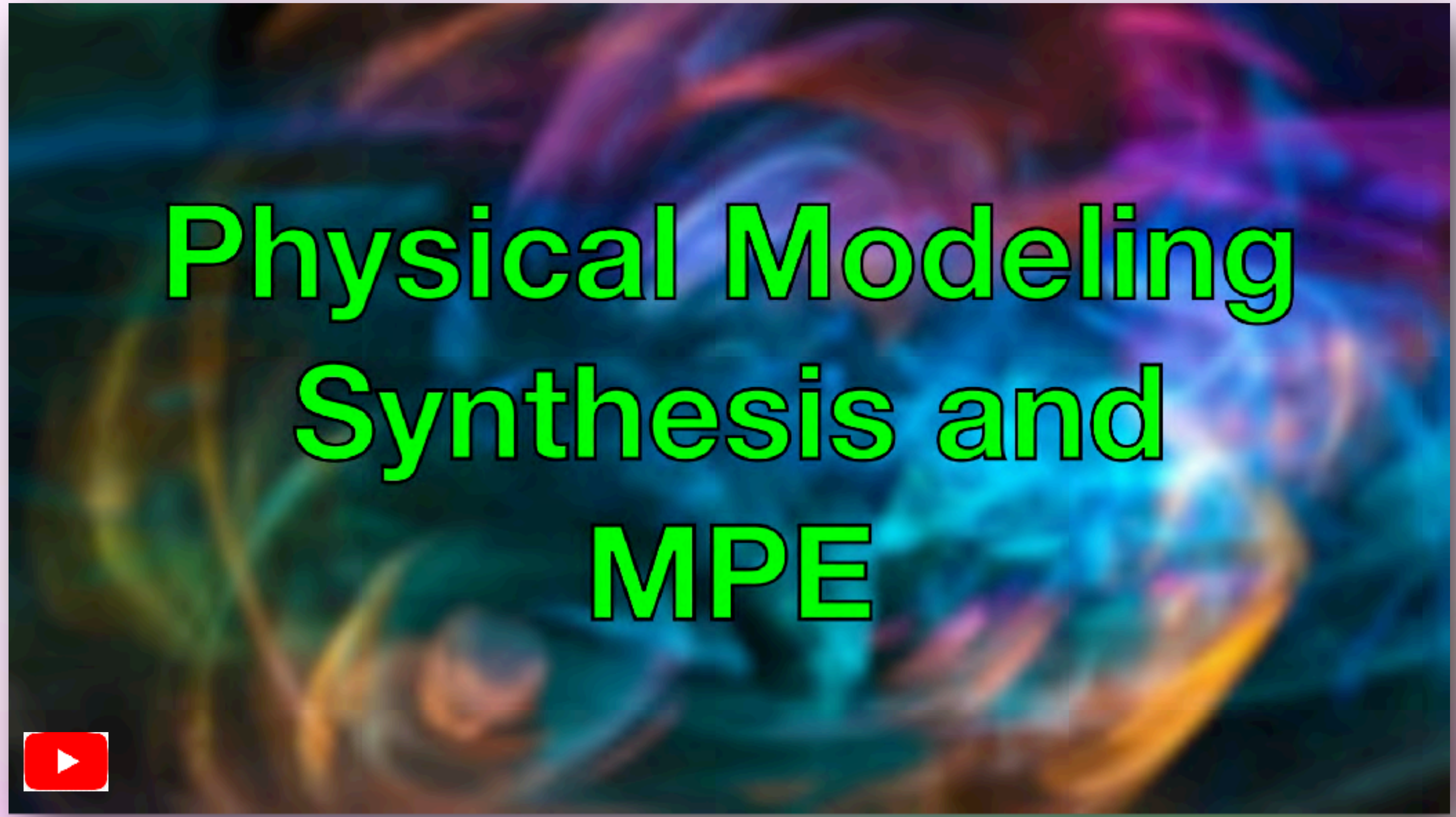


The Story

We find ourselves in a place where any of us can be Jimi Hendrix with just a small device in the palm of our hands. It's a fun and deeply technical topic drawing on many fields including physics, acoustics, digital signal processing and music.

In 1994, PM was poised to be the "Next Big Thing". **Why it's back, the future and what role might AI play!**

Physical Modeling Demos



As performing musicians, our role is to transform emotions and feelings into musical expression. Physical modeling (PM) instrument models give us expressive controls that enhance our ability to convey emotion to an audience.

What is Physical Modeling Synthesis?

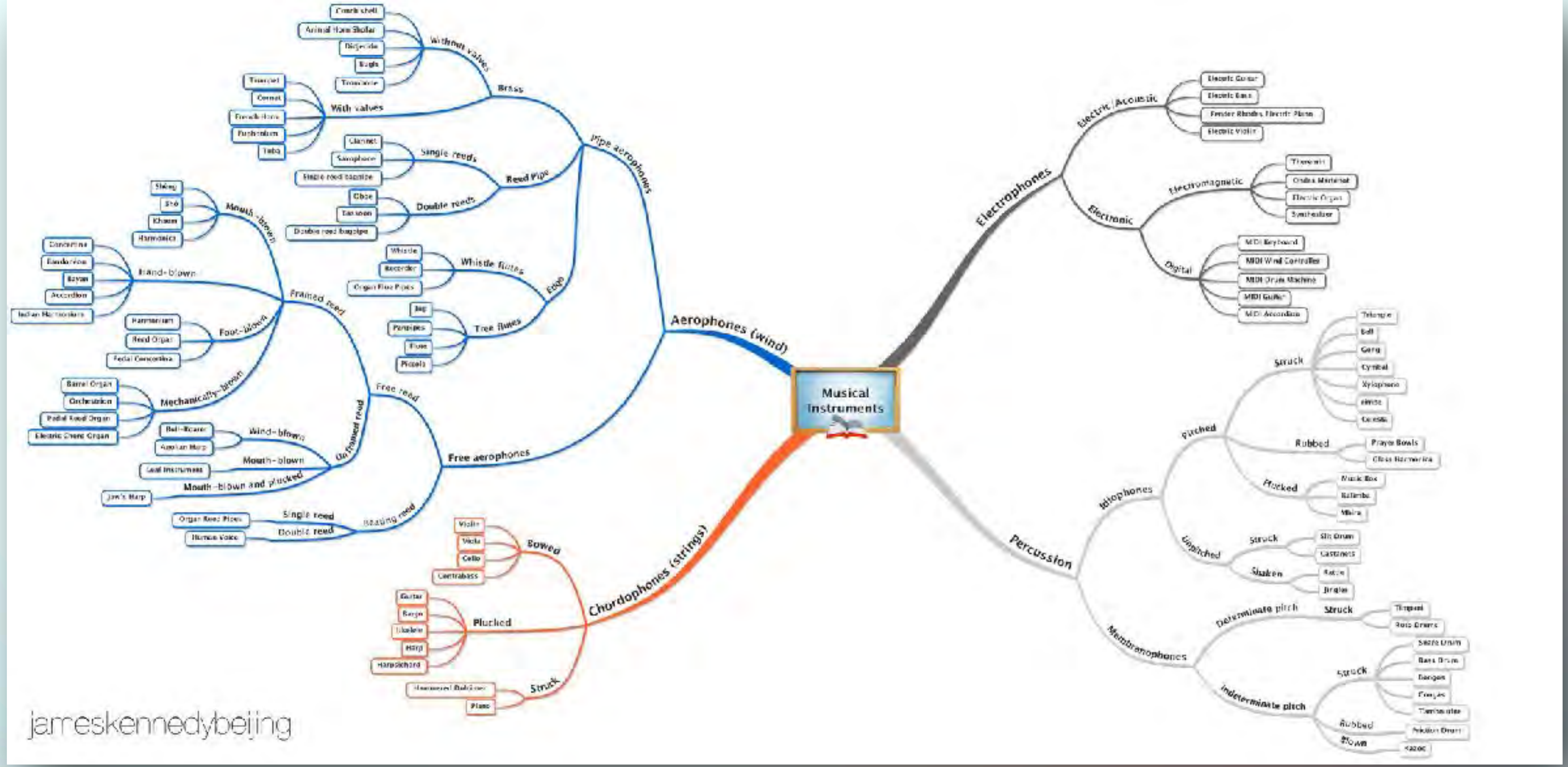
- **NOT AI !**
- Methods in which a sound is generated using a mathematical model of the physical source of sound.
- Any gestures that are used to interact with a real physical system can be mapped to parameters yielded an interactive and expressive performance experience.
- **Physical modeling is a collection of different techniques specific to each sound generation process.**

$$\frac{\partial^2 y}{\partial t^2} = \frac{1}{v_w^2} \frac{\partial^2 y}{dt^2}$$

Physics + Math
物理 + 數學

Taxonomy of Modeling Areas

Hornbostel–Sachs Classification



- Chordaphones - Guitars
- Aerophones - Woodwinds
- Membranophones - Drums

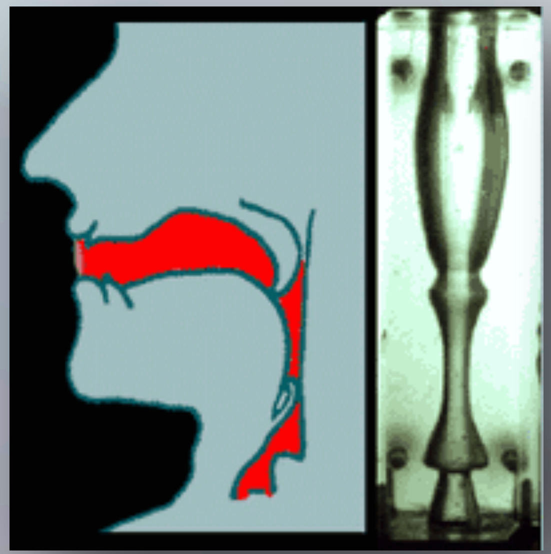
- Idiophones - Mallet Instruments
- Electrophones - Virtual Analog
- Game Sounds
- Human Voice

Early Mechanical Voice Synthesis

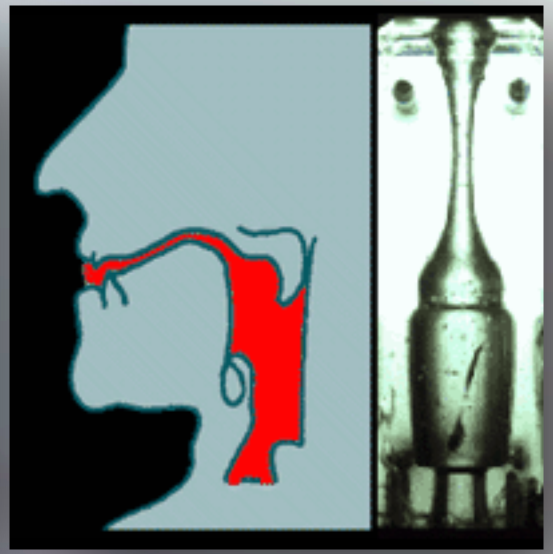
- 1000 -1200 ce - Speech Machines, Brazen Heads
- 1791 - Wolfgang Von Kempelin, speaking machine.
- 1857 - Joseph Faber, Euphonia (pictured)

Its been know for a long time that the vocal tract can be modeled with a bellows, a reed, a number of different size resonators and special elements for the tongue, the mouth. See Exploratorium Vocal Vowels.

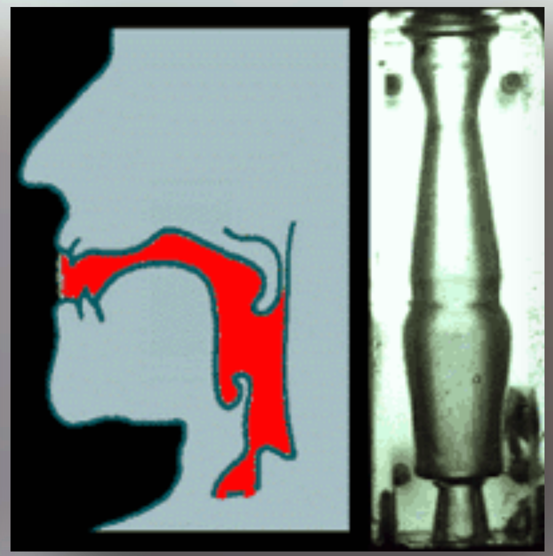
AH



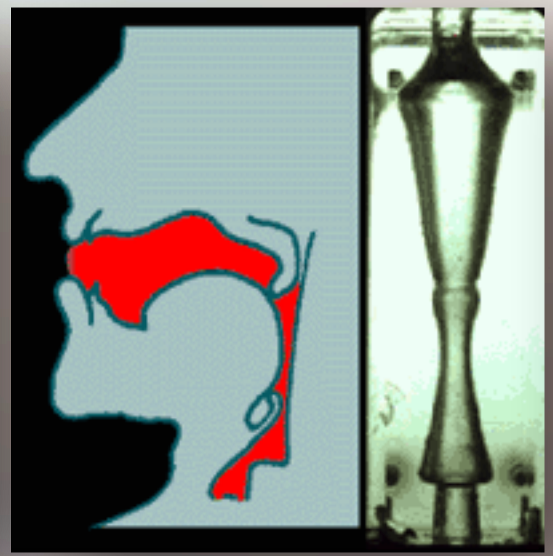
EE



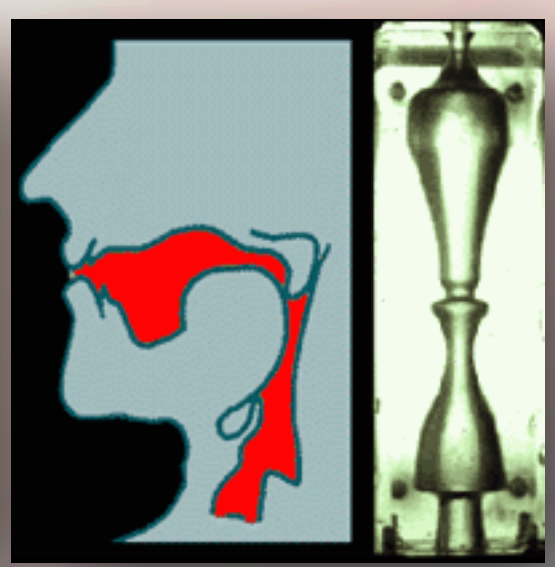
EH



OH



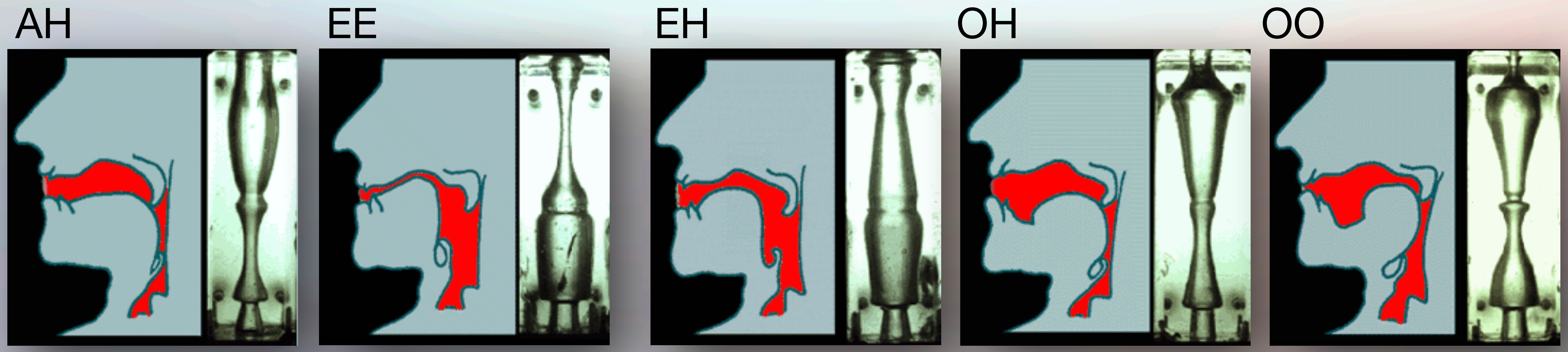
OO



Early Mechanical Voice Synthesis

- 1000 -1200 ce - Speech Machines, Brazen Heads
- 1791 - Wolfgang Von Kempelin, speaking machine.
- 1857 - Joseph Faber, Euphonia (pictured)

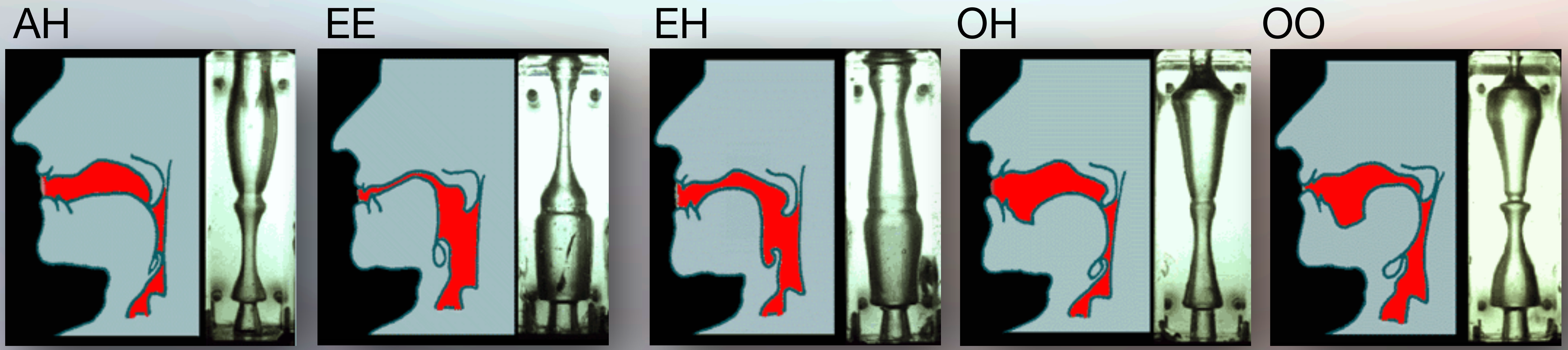
Its been know for a long time that the vocal tract can be modeled with a bellows, a reed, a number of different size resonators and special elements for the tongue, the mouth. See Exploratorium Vocal Vowels.



Early Mechanical Voice Synthesis

- 1000 -1200 ce - Speech Machines, Brazen Heads
- 1791 - Wolfgang Von Kempelin, speaking machine.
- 1857 - Joseph Faber, Euphonia (pictured)

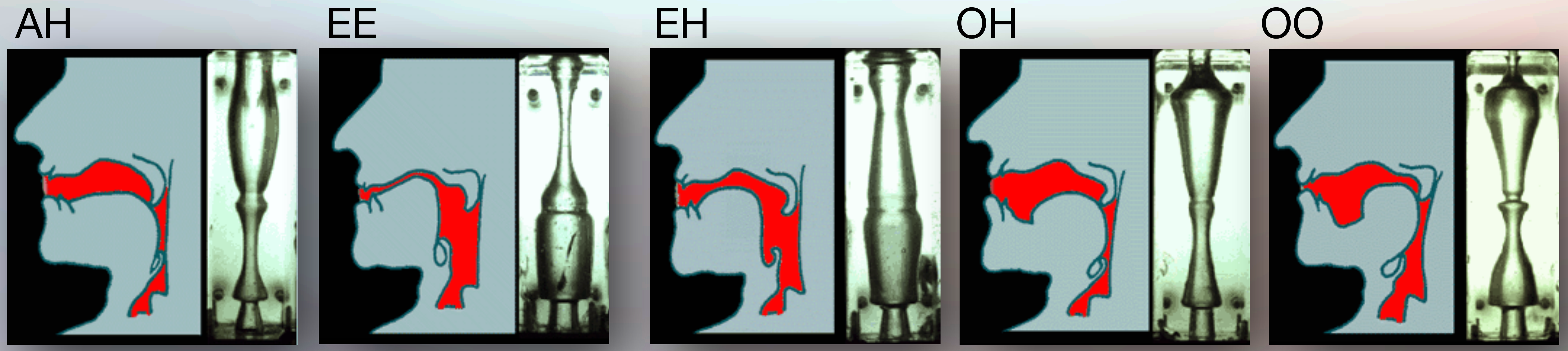
Its been know for a long time that the vocal tract can be modeled with a bellows, a reed, a number of different size resonators and special elements for the tongue, the mouth. See Exploratorium Vocal Vowels.



Early Mechanical Voice Synthesis

- 1000 -1200 ce - Speech Machines, Brazen Heads
- 1791 - Wolfgang Von Kempelin, speaking machine.
- 1857 - Joseph Faber, Euphonia (pictured)

Its been know for a long time that the vocal tract can be modeled with a bellows, a reed, a number of different size resonators and special elements for the tongue, the mouth. See Exploratorium Vocal Vowels.

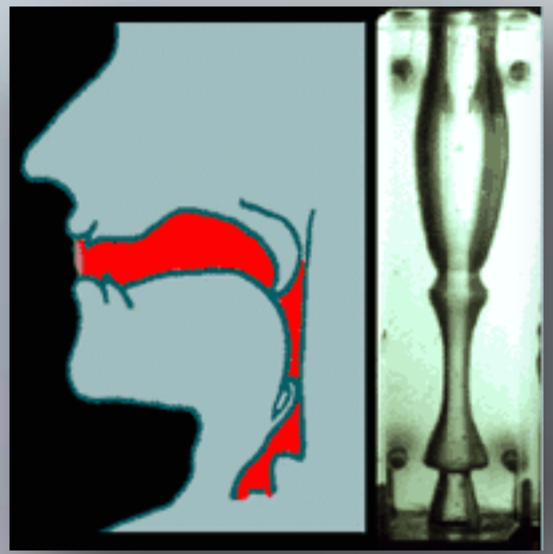


Early Mechanical Voice Synthesis

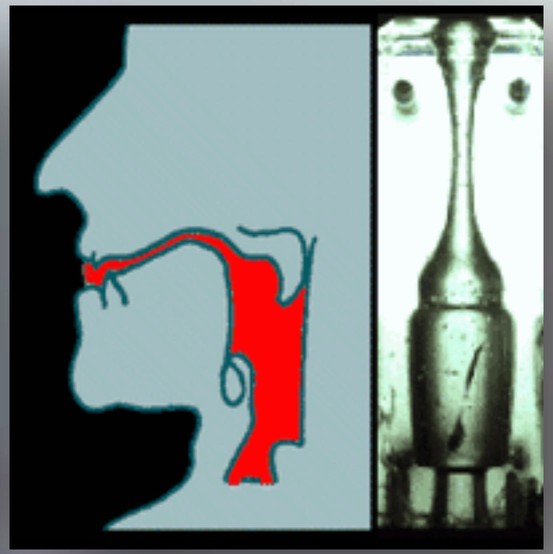
- 1000 -1200 ce - Speech Machines, Brazen Heads
- 1791 - Wolfgang Von Kempelin, speaking machine.
- 1857 - Joseph Faber, Euphonia (pictured)

Its been know for a long time that the vocal tract can be modeled with a bellows, a reed, a number of different size resonators and special elements for the tongue, the mouth. See Exploratorium Vocal Vowels.

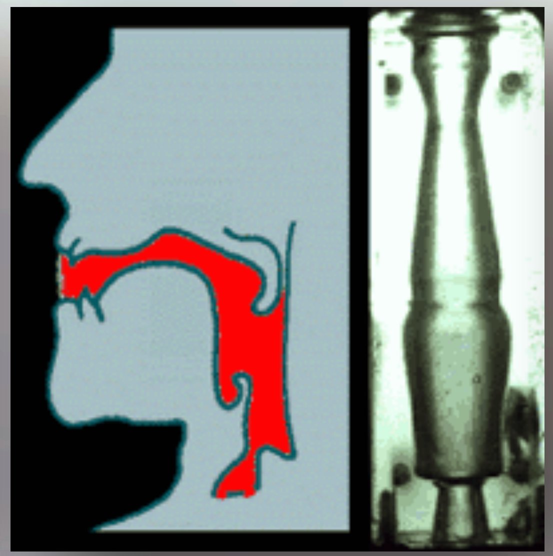
AH



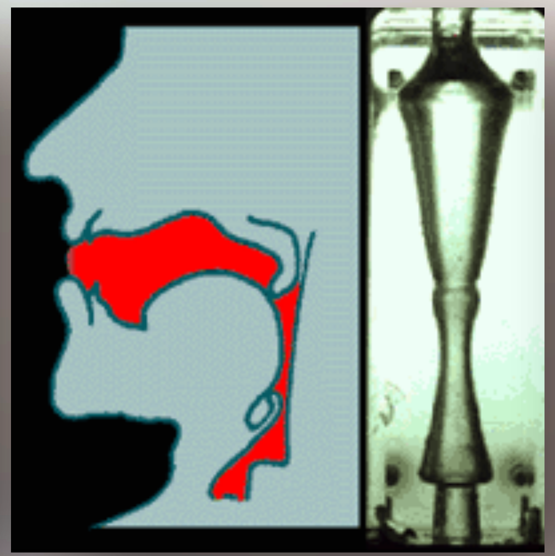
EE



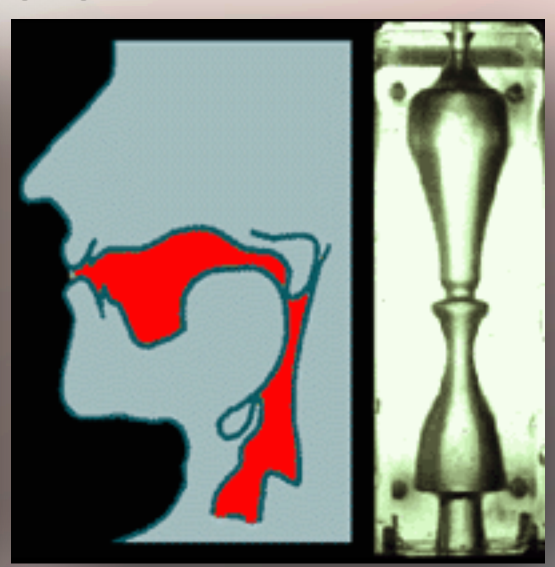
EH



OH



OO



The Voder (1937-39) - Homer Dudley

- Analog Electronic Speech Synthesis
- Analog model of the vocal tract. Vacuum Tubes!
- Developed from research on voice compression at Bell Labs.
- Featured at the 1939 Worlds fair

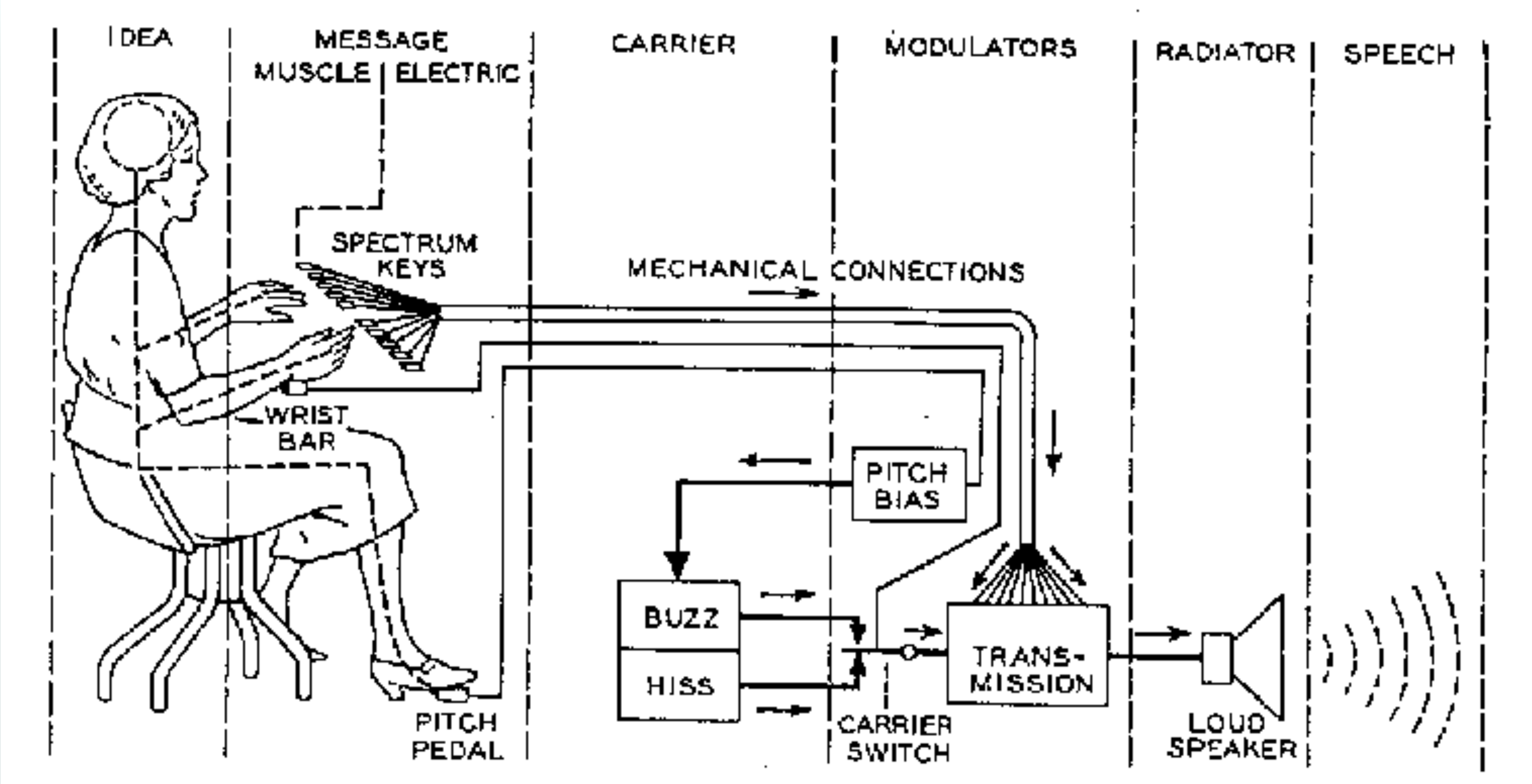
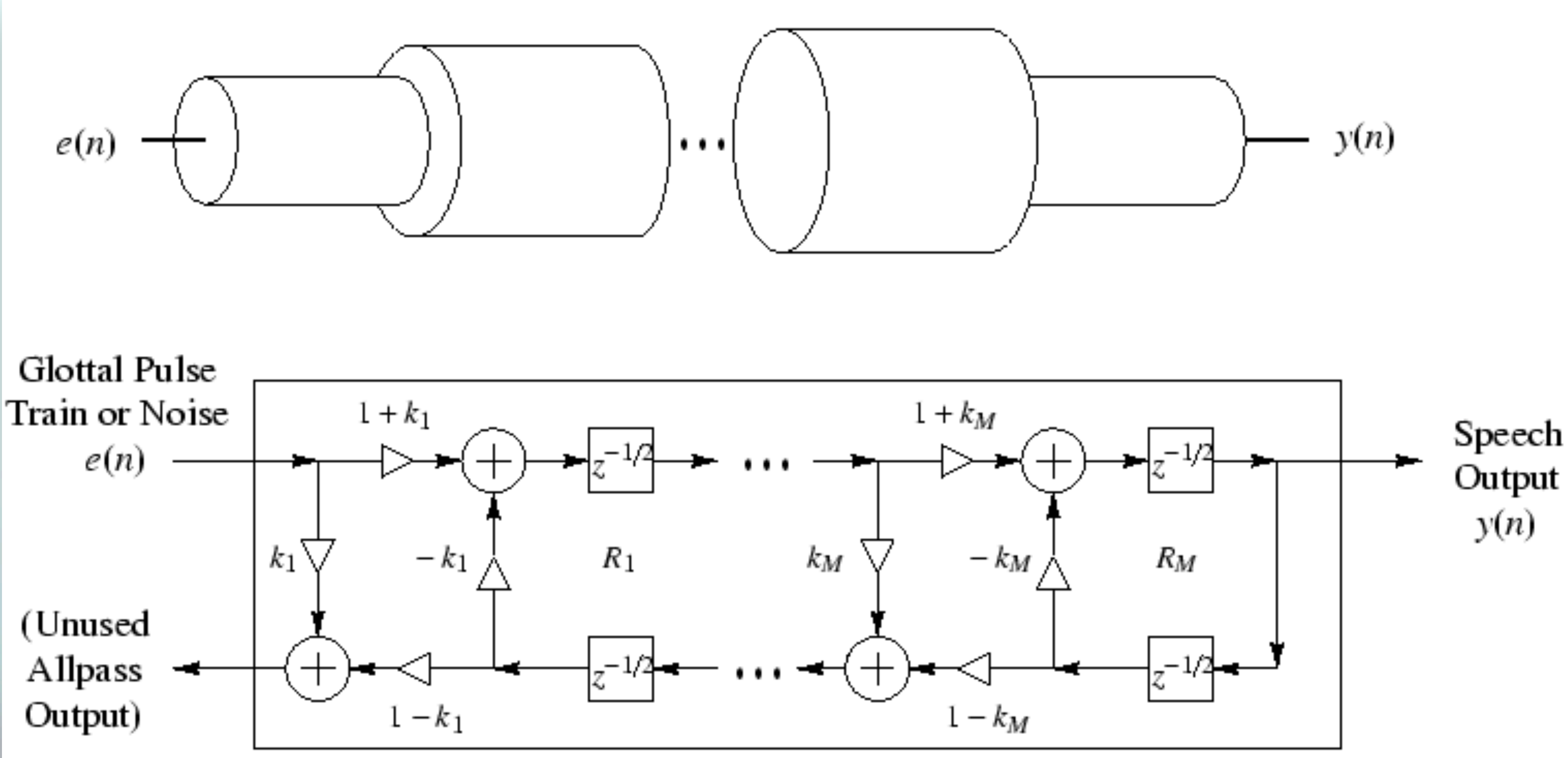


Fig. 8—Schematic circuit of the voder.

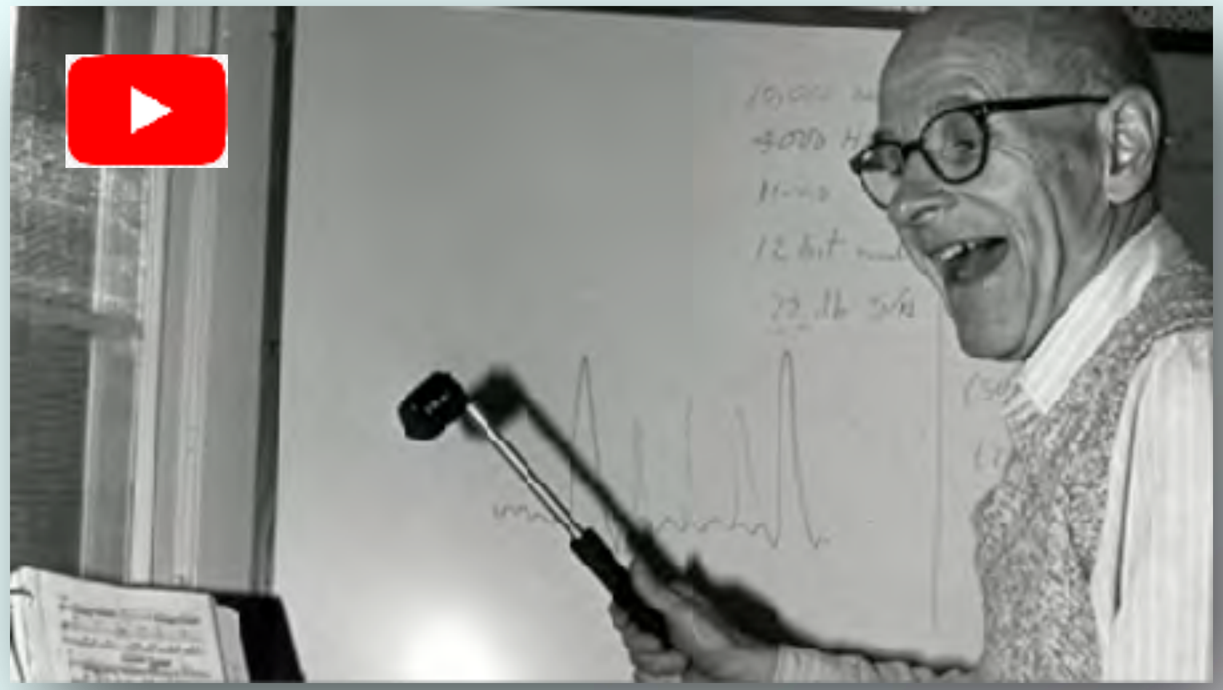
Kelly-Lochbaum Vocal Tract Model (1961)



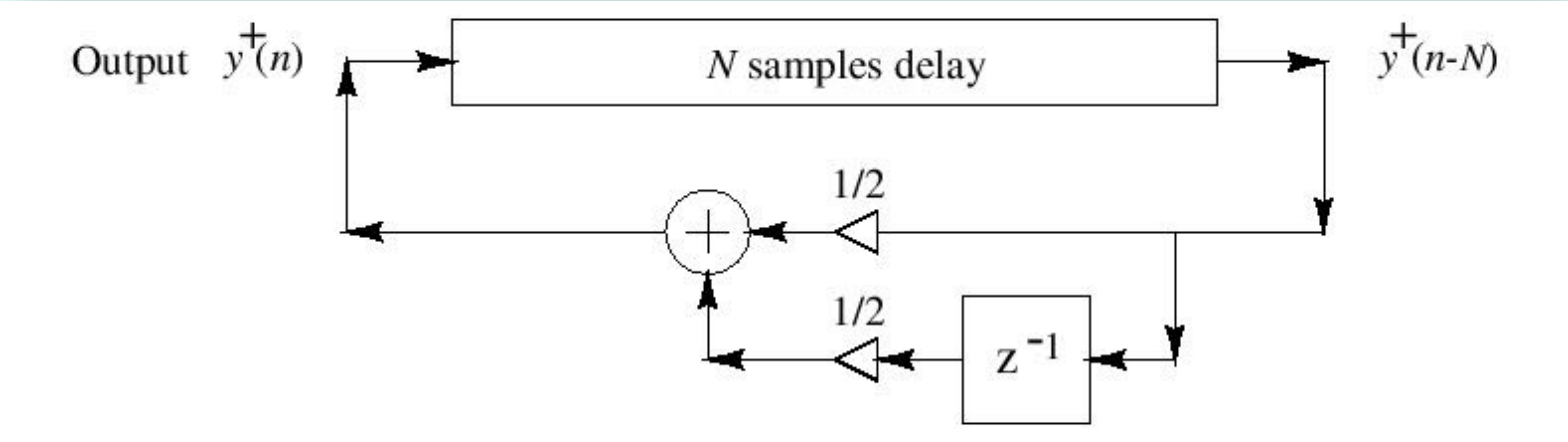
Kelly-Lochbaum Vocal Tract Model (Piecewise Cylindrical)

Daisy Bell (1961)

- Daisy Bell
- Vocal part by Kelly and Lochbaum (1961)
- Musical accompaniment by Max Mathews
- Computed on an IBM 704
- Based on Russian speech-vowel data from Gunnar Fant's book
- Probably the first digital physical-modeling synthesis sound example by any method
- Inspired Arthur C. Clarke to adapt it for "2001: A Space Odyssey" the Hal 9000's "first song"

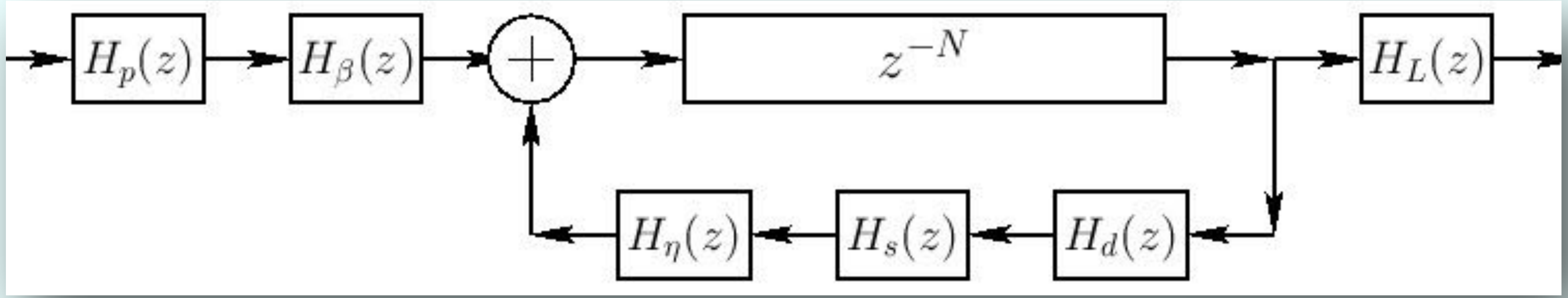


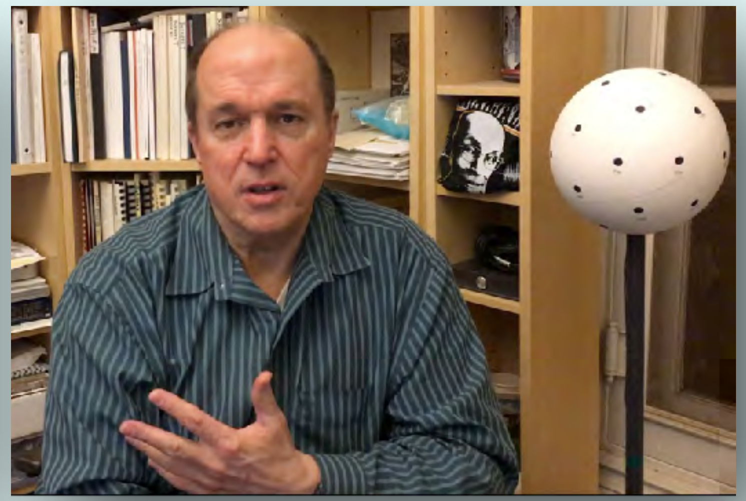
Karplus-Strong (KS) Algorithm (1978)



- Discovered (1978) as "self-modifying wavetable synthesis"
- Wavetable is preferably initialized with random numbers
- Licensed to Mattel
- The first musical use of the algorithm was in the work "*May All Your Children Be Acrobats*" written in 1981 by David A. Jaffe.

EKS Algorithm (Jaffe-Smith 1983)





$$
\begin{aligned}
H_p(z) &= \frac{1-p}{1-p\,z^{-1}} = \text{pick-direction lowpass filter} \\
H_\beta(z) &= 1 - z^{-\lfloor \beta N + 1/2 \rfloor} = \text{pick-position comb filter, } \beta \in (0,1) \\
H_d(z) &= \text{string-damping filter (one/two poles/zeros typical)} \\
H_s(z) &= \text{string-stiffness allpass filter (several poles and zeros)} \\
H_\eta(z) &= -\frac{\eta(N) - z^{-1}}{1 - \eta(N)\,z^{-1}} = \text{first-order string-tuning allpass filter} \\
H_L(z) &= \frac{1-R_L}{1-R_L\,z^{-1}} = \text{dynamic-level lowpass filter}
\end{aligned}
$$

- Musical Example “Silicon Valley Breakdown” (Jaffe 1992)
- Musical Example BWV-1041 (used to intro the NeXT machine 1988)

The KS and EKS Papers Were Published Simultaneously in the Computer Music Journal (CMJ) (1983)

Kevin Karplus
Computer Science Department
Cornell University
Ithaca, New York 14853

Alex Strong
Computer Science Department
Stanford University
Stanford, California 94305

Digital Synthesis of Plucked-String and Drum Timbres

Introduction

There are many techniques currently used for digital music synthesis, including frequency modulation (FM) synthesis, waveshaping, additive synthesis, and subtractive synthesis. To achieve rich, natural sounds, all of them require fast arithmetic capability, such as is found on expensive computers or digital synthesizers. For musicians and experimenters without access to these machines, musically interesting digital synthesis has been almost impossible.

The techniques described in this paper can be implemented quite cheaply on almost any computer. Real-time synthesis implementations have been done for Intel 8080A (by Alex Strong), Texas Instruments TMS9900 (by Kevin Karplus), and SC/MP (by Mike Plass) microprocessors. David Jaffe and Julius Smith have programmed the Systems Concept Digital Synthesizer at the Center for Computer Research in Music and Acoustics (CCRMA) to perform several variants of the algorithms (Jaffe and Smith 1983).

Not only are the algorithms simple to implement in software, but hardware realizations are easily done. The authors have designed and tested a custom *n*-channel metal-oxide semiconductor (*n*MOS) chip (the Digitar chip), which computes 16 independent notes, each with a sampling rate of 20 KHz.

Despite the simplicity of the techniques, the sound is surprisingly rich and natural. When the plucked-string algorithm was compared with additive synthesis at Bell Laboratories, it was found that as many as 30 sine wave oscillators were needed to produce a similarly realistic timbre (Sleator 1981). The entire plucked-string algorithm requires only as much computation as one or two sine wave oscillators.

The parameters available for control are pitch, amplitude, and decay time. The pitch is specified by an integer that is approximately the period of the sound, in samples (periodicity parameter p). Amplitude is specified as the initial peak amplitude A. Decay time is determined by the pitch and by a decay stretch factor S.

The algorithms in this paper lack the versatility of FM synthesis, additive synthesis, or subtractive synthesis. They are, however, cheap to implement, easy to control, and pleasant to hear. For musicians interested primarily in performing and composing music, rather than designing instruments, these algorithms provide a welcome new technique. For those interested in instrument design, they open a new field of effective techniques to explore.

Wavetable Synthesis

One standard synthesis technique is the *wavetable synthesis* algorithm. It consists of repeating a number of samples over and over, thus producing a purely periodic signal. If we let Y_t be the value of the t^{th} sample, the algorithm can be written mathematically as

$$Y_t = Y_{t-p}.$$

The parameter p is called the *wavetable length* or *periodicity parameter*. It represents the amount of memory needed and the period of the tone (in sam-

This research was supported in part by the Fannie and John Hertz Foundation.

Computer Music Journal, Vol. 7, No. 2,
Summer 1983, 0148-9267/83/020043-13 $04.00/0,

David A. Jaffe and Julius O. Smith
Center for Computer Research in Music and Acoustics (CCRMA)
Stanford University
Stanford, California 94305

Extensions of the Karplus-Strong Plucked-String Algorithm

Introduction

In 1960, an efficient computational model for vibrating strings, based on physical resonating, was proposed by McIntyre and Woodhouse (1960). This model plays a crucial role in their recent work on bowed strings (McIntyre, Schumacher, and Woodhouse 1981; 1983), and methods for calibrating the model to recorded data have been developed (Smith 1983).

Independently, in 1978, Alex Strong devised an efficient special case of the McIntyre-Woodhouse string model that produces remarkably rich and realistic timbres despite its simplicity (Karplus and Strong 1983). Since then, Strong and Kevin Karplus have explored several variations and refinements of the algorithm, with an emphasis on small-system implementations. We have found that the Karplus-Strong algorithm can be used with equally impressive results on fast, high-power equipment. The availability of multiplies, for example, allows several modifications and extensions that increase its usefulness and flexibility. These extensions are described in this paper. The developments were motivated by musical needs that arose during the composition of *May All Your Children Be Acrobats* (1981) for computer-generated tape, eight guitars, and voice and *Silicon Valley Breakdown* (1982) for four-channel, computer-generated tape, both written by David Jaffe. Our theoretical approach and the extensions based on it have also been applied to the McIntyre-Woodhouse algorithm (Smith 1983).

The String-Simulation Algorithm

The Karplus-Strong plucked-string algorithm is presented in this issue of *Computer Music Journal*. From our point of view, the algorithm consists of a high-order *digital filter*, which represents the string, and a short *noise burst*, which represents the "pluck."[1] The digital filter is given by the difference equation

$$y_n = x_n + \frac{y_{n-N} + y_{n-(N+1)}}{2}, \qquad (1)$$

where x_n is the input signal amplitude at sample n, y_n is the output amplitude at sample n, and N is the (approximate) desired pitch period of the note in samples. The noise burst is defined by

$$x_n = \begin{cases} Au_n, & n = 0, 1, 2, \ldots, N-1 \\ 0, & n \geq N, \end{cases}$$

where A is the desired amplitude, and $u_n \in [-1,1]$ is the output of a random-number generator. The output y_n is taken beginning at time $n = N$ in our implementation.

Analysis of the String Simulator

Before proceeding to practical extensions of the algorithm, we will describe the theory on which many of them are based. Various concepts from digital filter theory are employed. For a tutorial introduction to digital filter theory, see the works by Smith (1982b) and Steiglitz (1974).

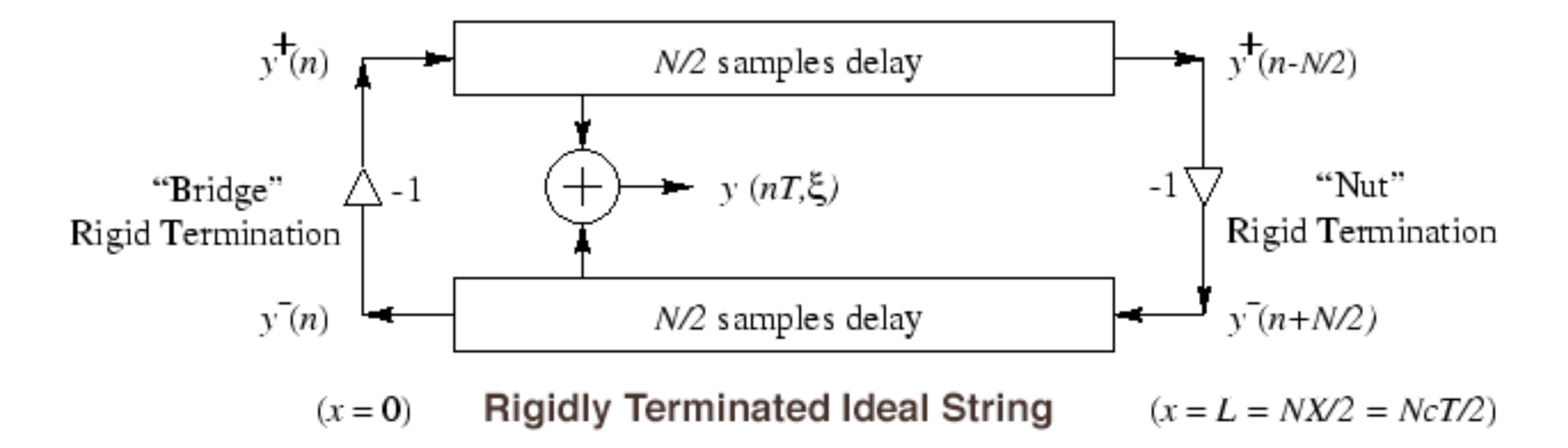
The input-output relation of Eq. (1) may be ex-

David A. Jaffe is also affiliated with the Music Department at Stanford University, and Julius O. Smith is also affiliated with the Electrical Engineering Department there.

Computer Music Journal, Vol. 7, No. 2,
Summer 1983, 0148-9257/83/020056-14 $04.00/0,

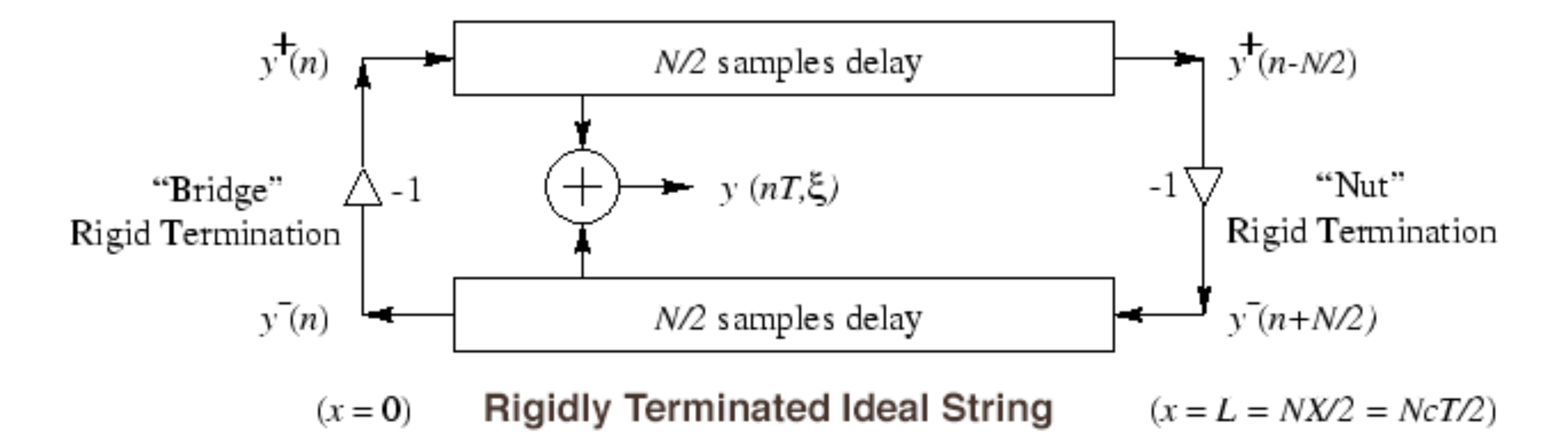
1. In some situations, the sound more closely resembles a string struck with a hammer or mallet than one plucked with a pick, but we will always use the term *pluck* when referring to the excitation.

Digital Waveguide Models (Smith 1985)



- Equivalent to d'Alembert's Solution to the Partial Differential Equation for a string (1747)
- Used for the Yamaha VL Family (1994)
- Shakuhachi, Tenor Sax

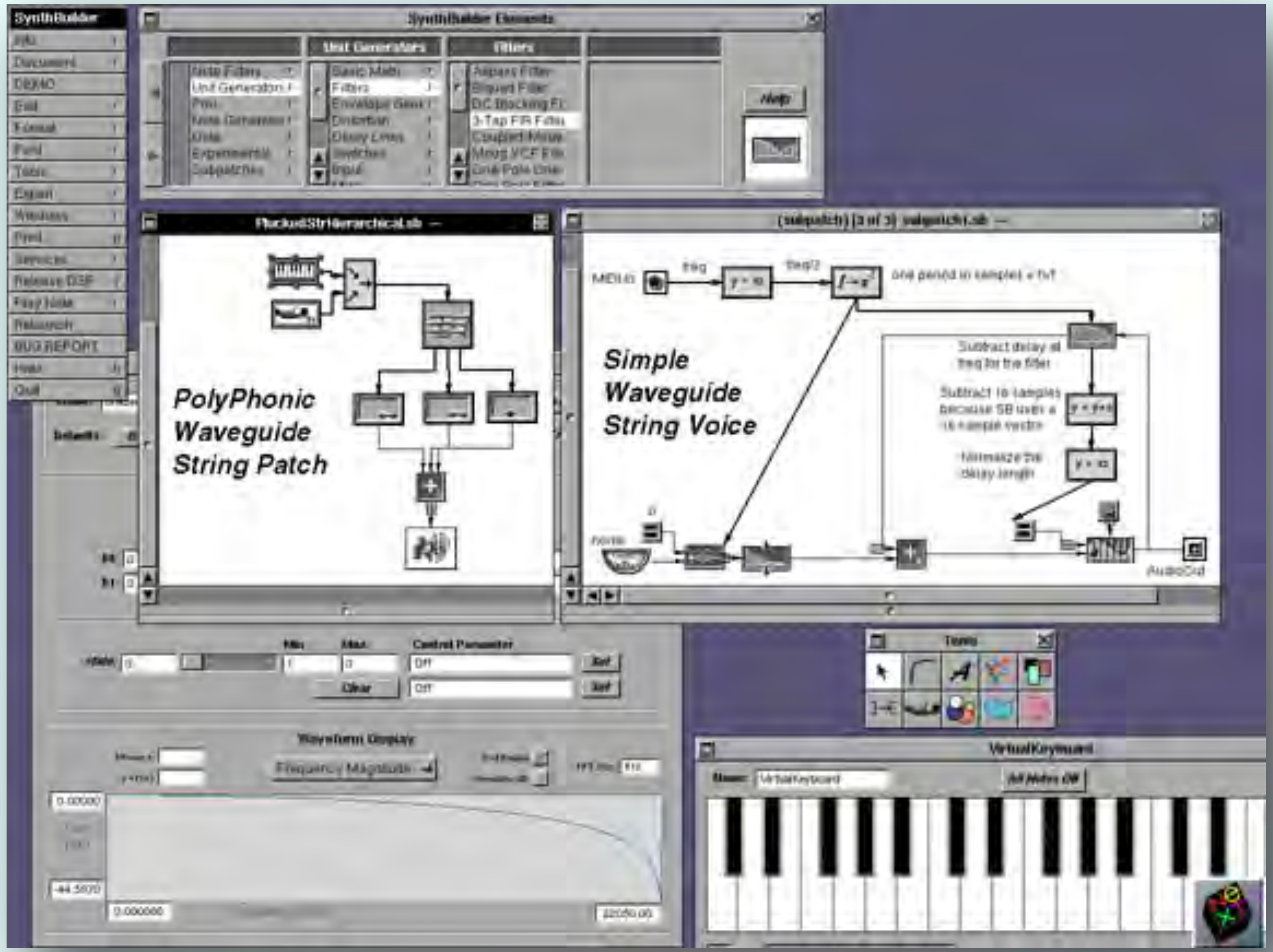
Digital Waveguide Models (Smith 1985)



- Equivalent to d'Alembert's Solution to the Partial Differential Equation for a string (1747)
- Used for the Yamaha VL Family (1994)
- Shakuhachi, Tenor Sax

SynthBuilder (Porcaro et al 1993-1997)

- SynthBuilder was a rapid-prototyping tool on the NeXT machine for the development of music synthesis and effects patches. Initially for the 56k DSP and later for SynthServer/ SynthScript.
- Leveraged the NeXT Music Kit and the source code for the NeXT Draw Program.
- It played a major role in the development of physical models including Coupled Mode Synthesis (Van Duyne), Virtual Analog (Stilson, Smith) Sondius Program.
- SynthBuilder was written by Nick Porcaro with significant contributions from David Jaffe and Pat Scandalis, Julius Smith, Tim Stinson and Scott Van Dyne.



1997 SynthBuilder won the Grand Prize in the Bourges International Music Software Competition

Coupled Mode Synthesis (CMS)
(Van Duyne 1996)

- Modeling of percussion sounds
- Modal technique with coupling
- Tibetan Bell Model
- Wind Chime Model
- Tubular Bells Model
- Percussion Ensemble
- Taiko Ensemble

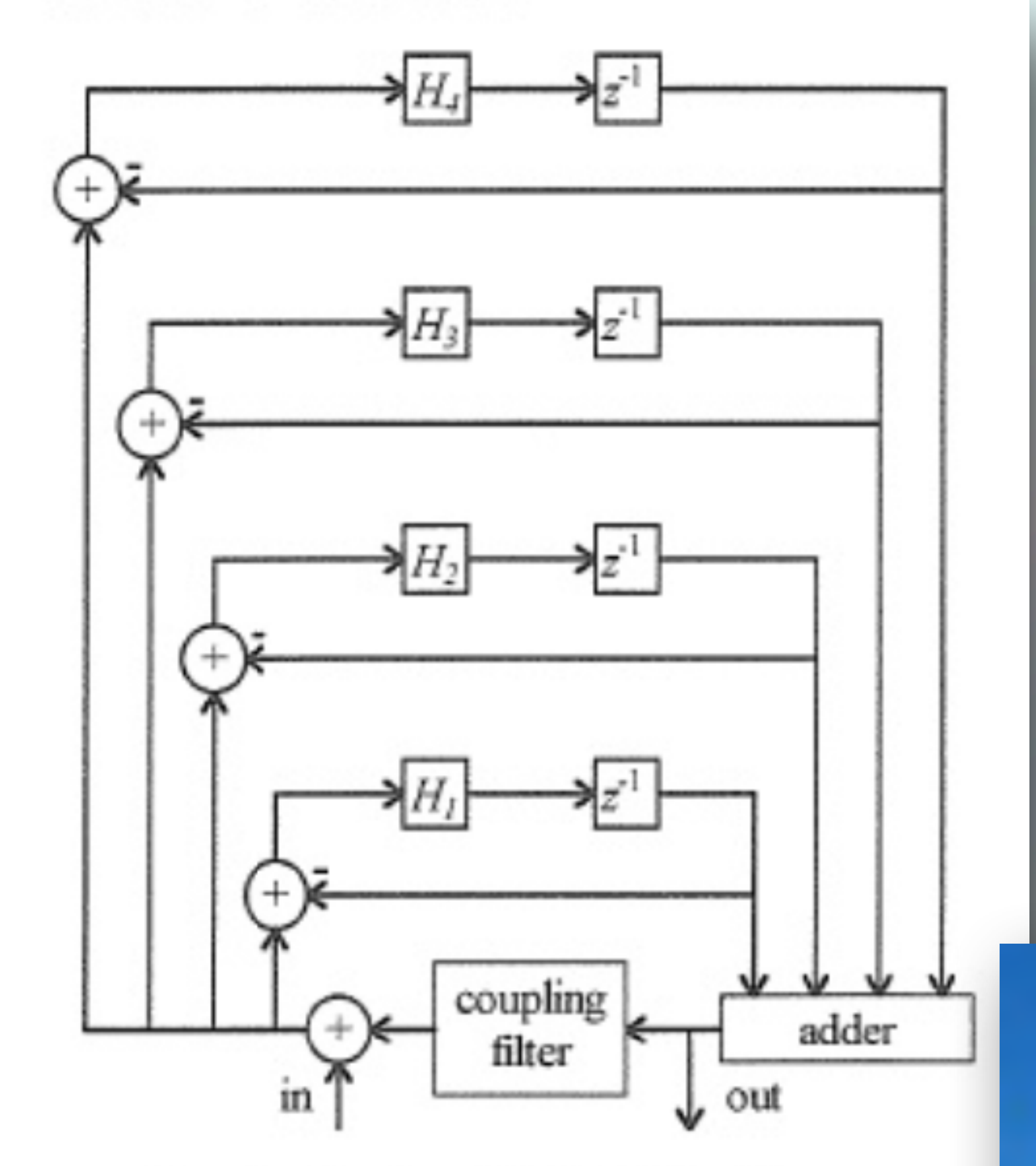


Figure 2: The Coupled Mode Filter Structure

Coupled Mode Synthesis (CMS)
(Van Duyne 1996)

- Modeling of percussion sounds
- Modal technique with coupling
- Tibetan Bell Model
- Wind Chime Model
- Tubular Bells Model
- Percussion Ensemble
- Taiko Ensemble

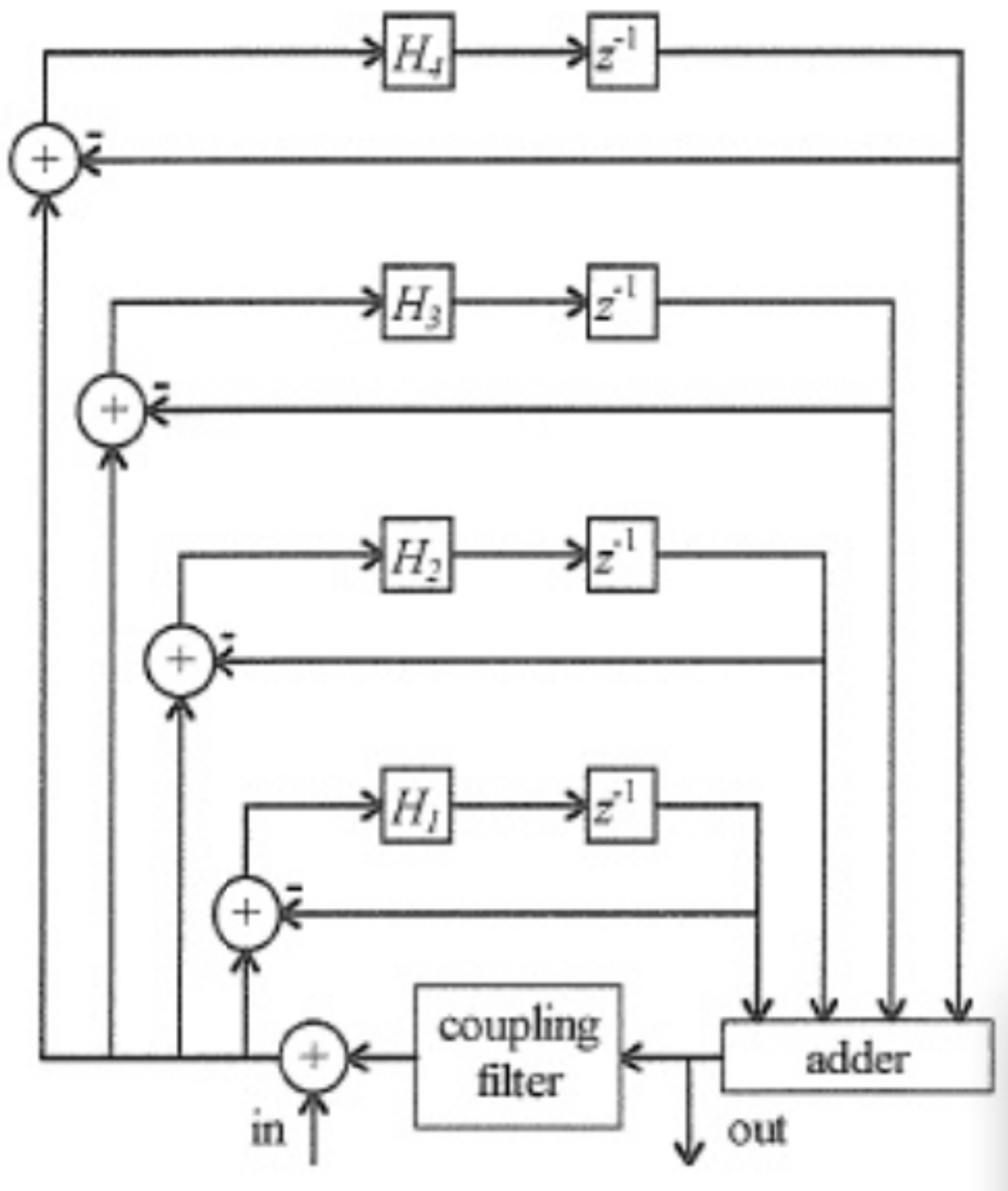


Figure 2: The Coupled Mode Filter Structure

Coupled Mode Synthesis (CMS)
(Van Duyne 1996)

- Modeling of percussion sounds
- Modal technique with coupling
- Tibetan Bell Model
- Wind Chime Model
- Tubular Bells Model
- Percussion Ensemble
- Taiko Ensemble

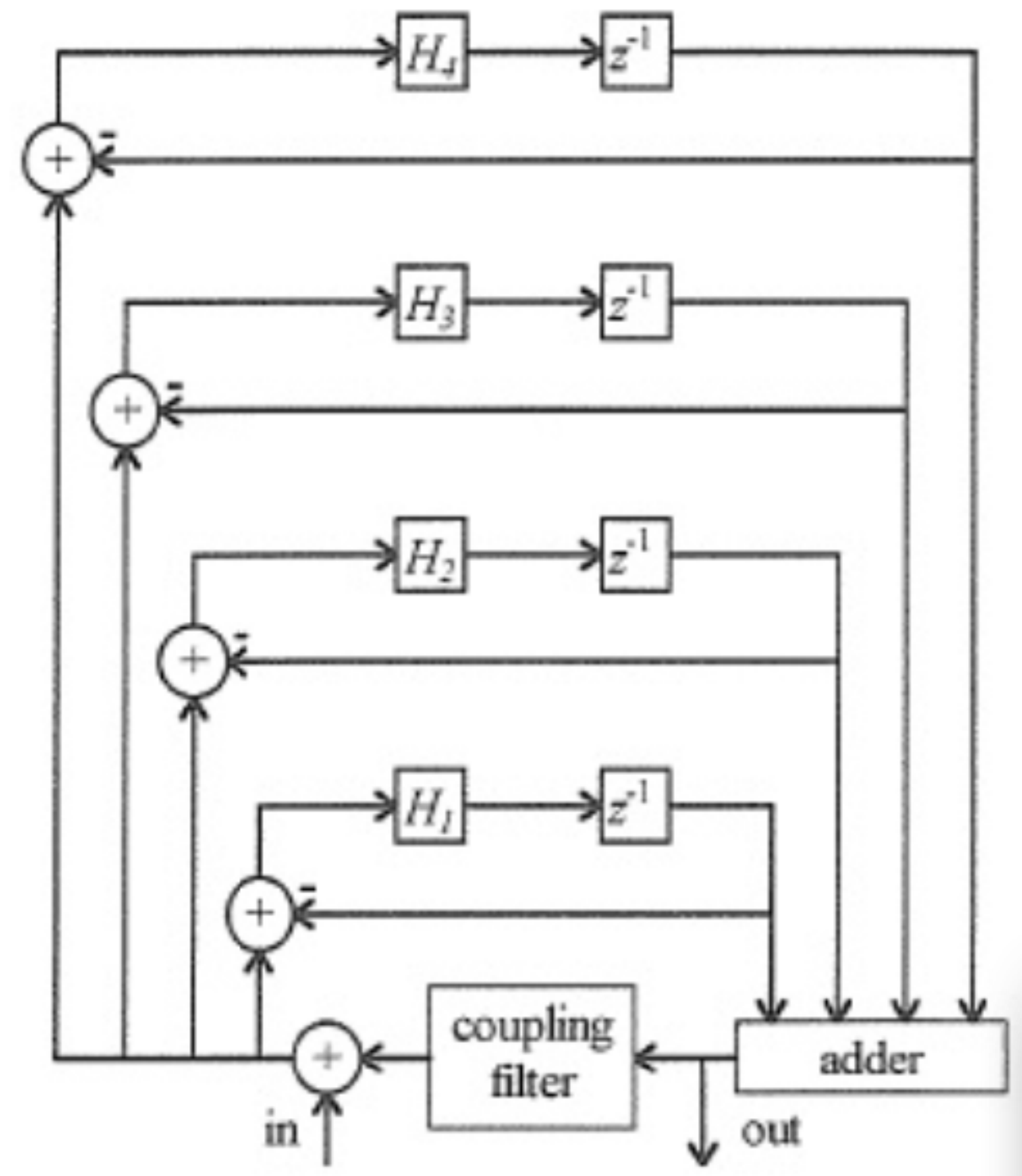


Figure 2: The Coupled Mode Filter Structure

Coupled Mode Synthesis (CMS)
(Van Duyne 1996)

- Modeling of percussion sounds
- Modal technique with coupling
- Tibetan Bell Model
- Wind Chime Model
- Tubular Bells Model
- Percussion Ensemble
- Taiko Ensemble

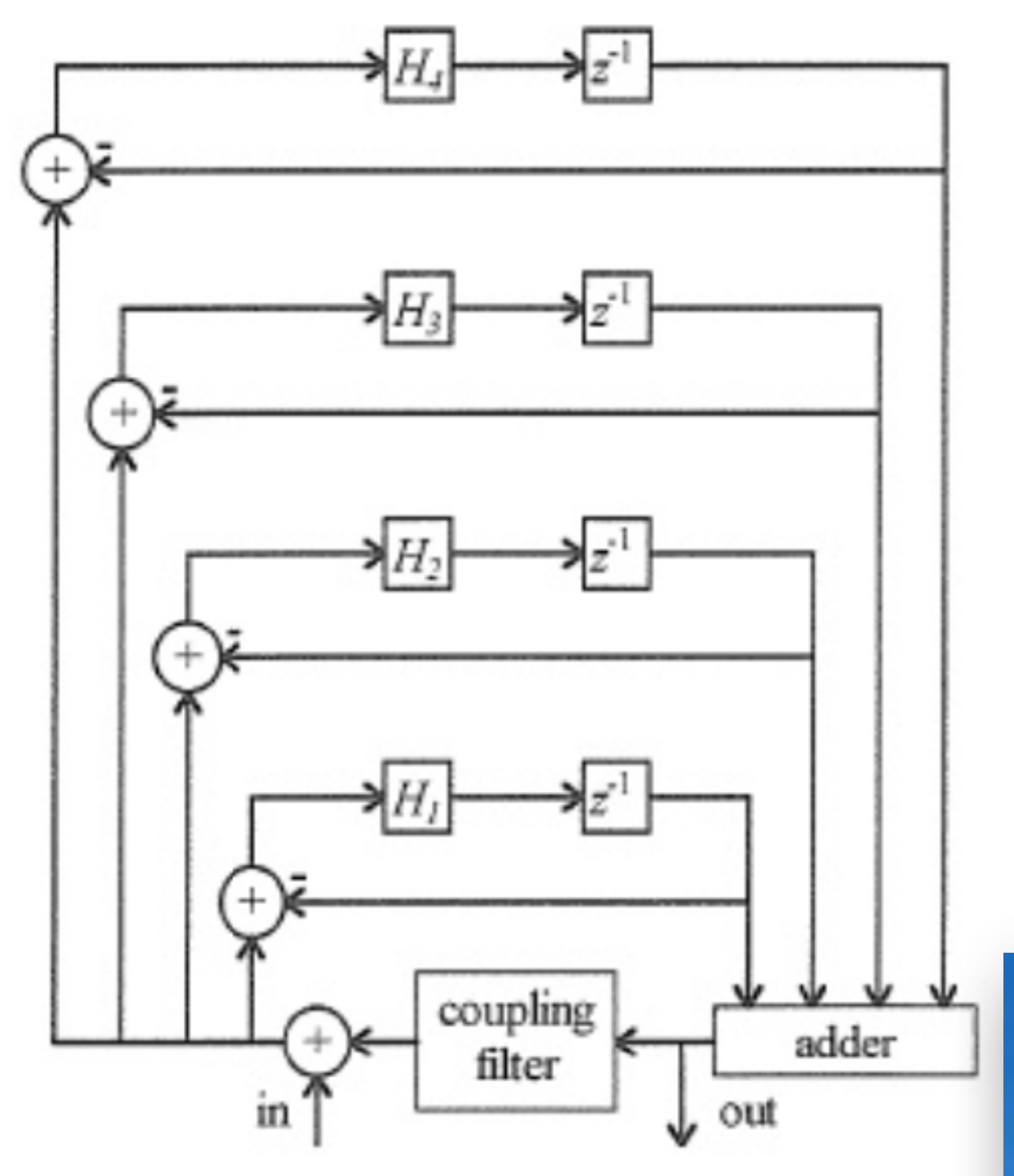


Figure 2: The Coupled Mode Filter Structure

Coupled Mode Synthesis (CMS) (Van Duyne 1996)

- Modeling of percussion sounds
- Modal technique with coupling
- Tibetan Bell Model
- Wind Chime Model
- Tubular Bells Model
- Percussion Ensemble
- Taiko Ensemble

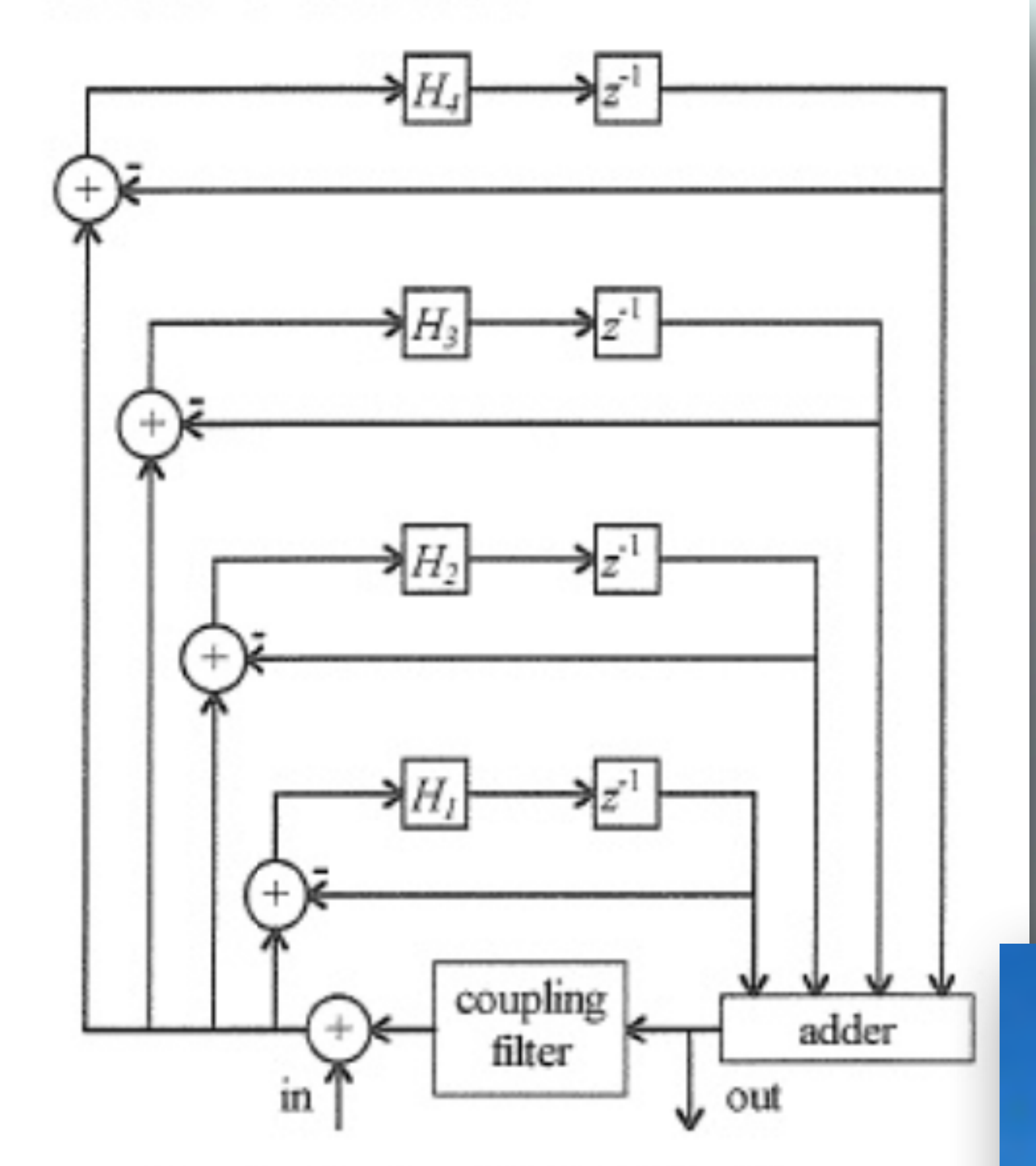
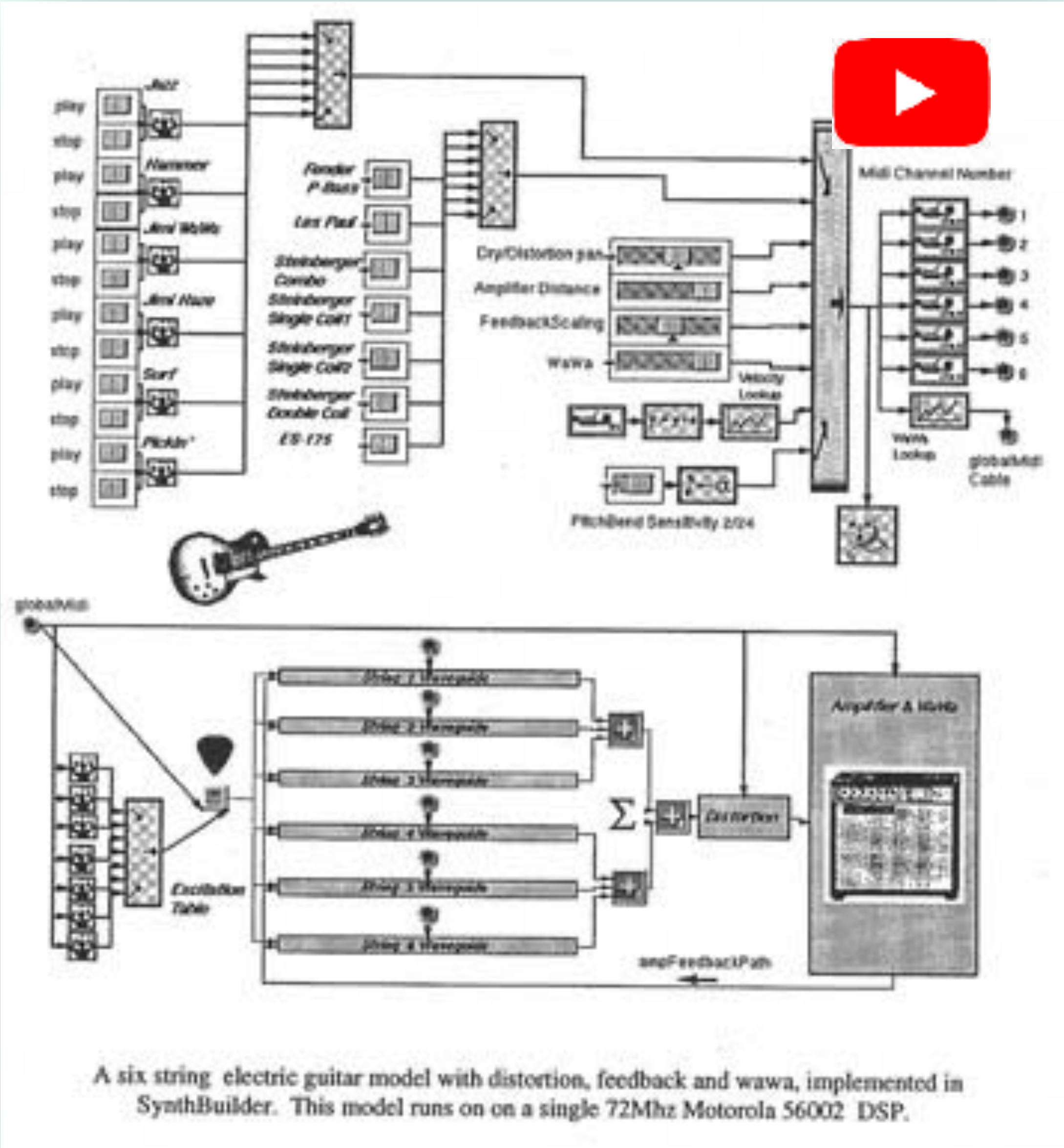
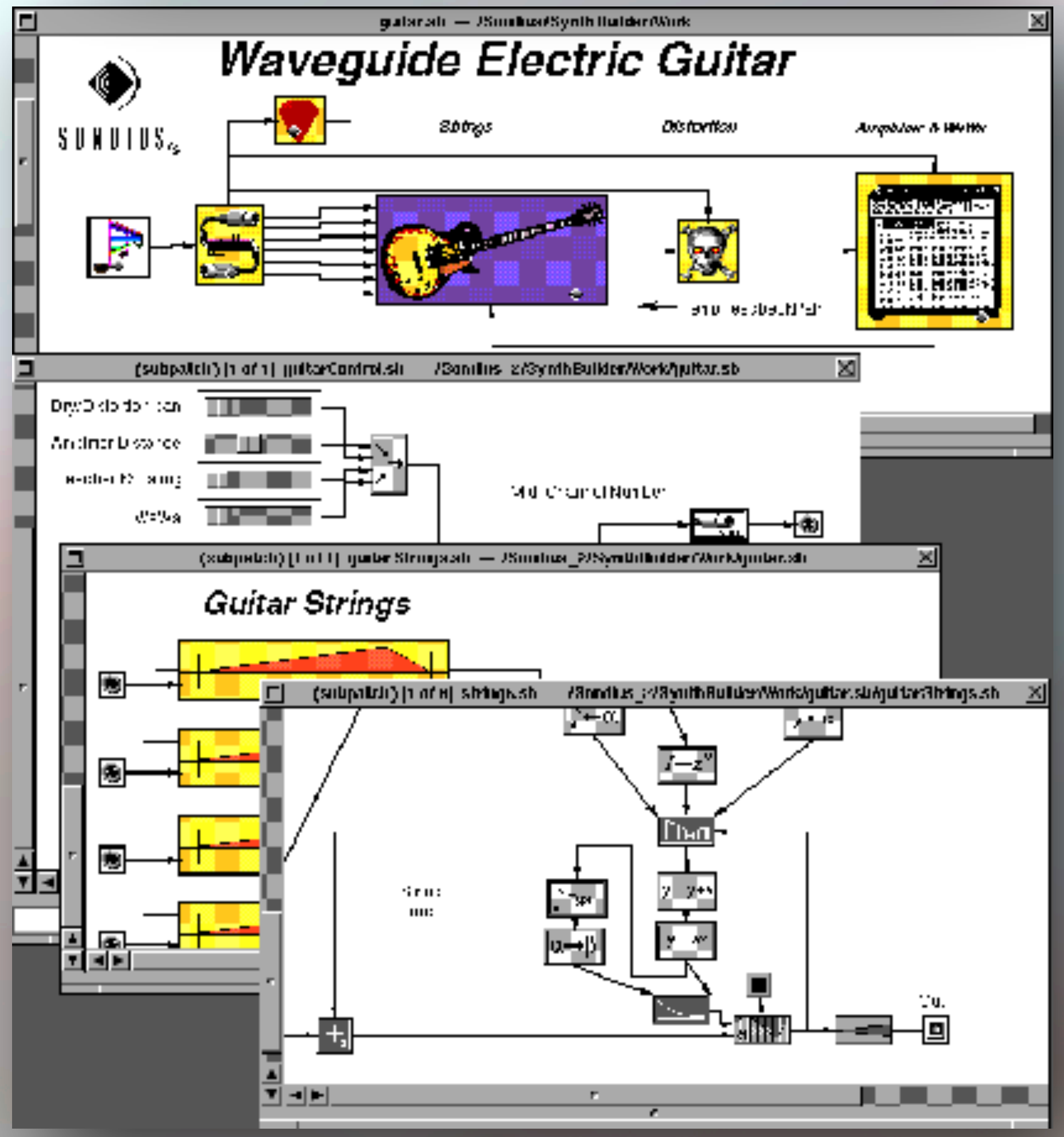


Figure 2: The Coupled Mode Filter Structure

Guitar Model (Scandalis 1996)

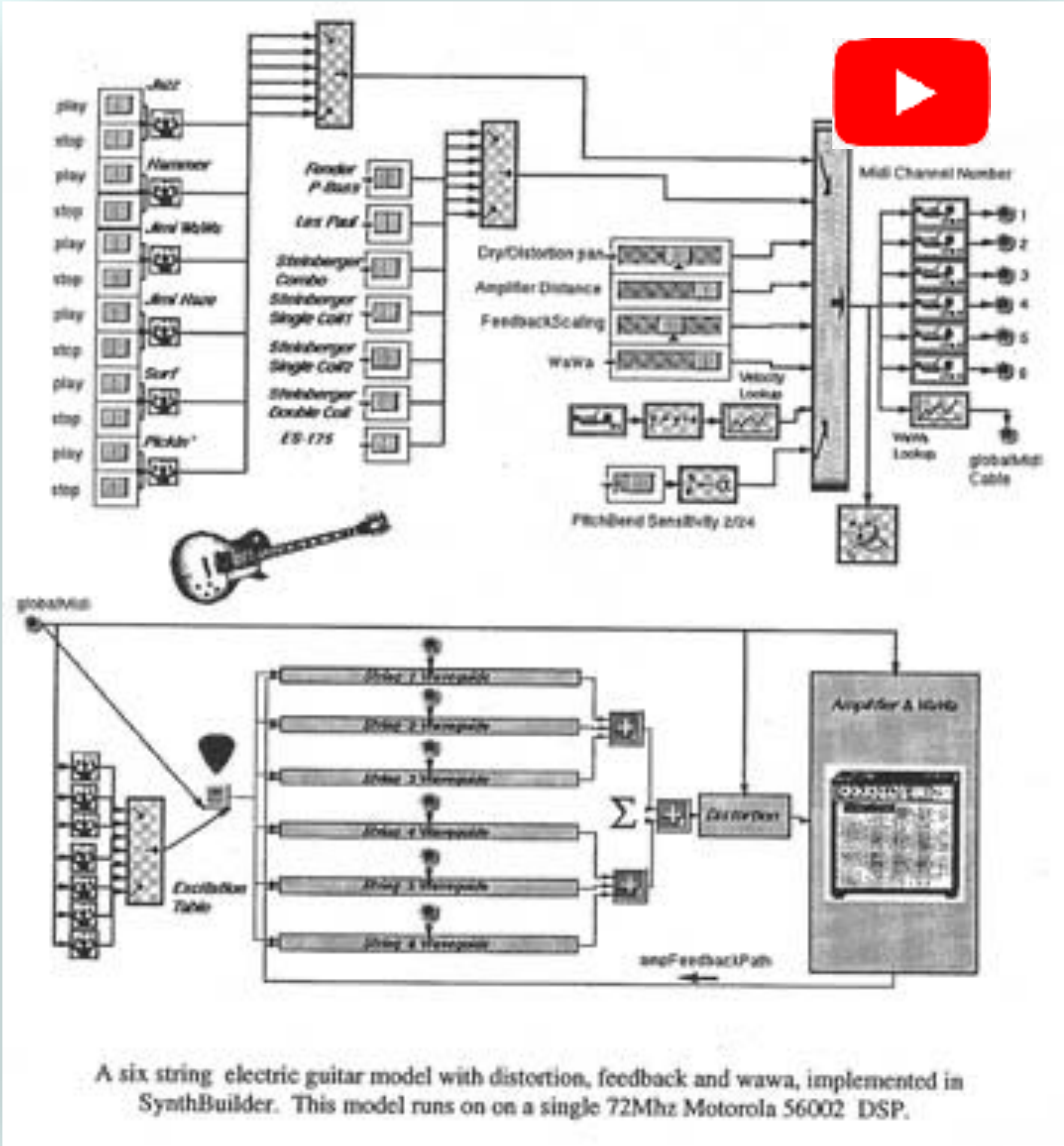
- Distortion, feedback and effects.
- Initial period excitations to capture the sound of different guitars.
- Controlled with Yamaha G10 guitar controller similar to today's MPE.

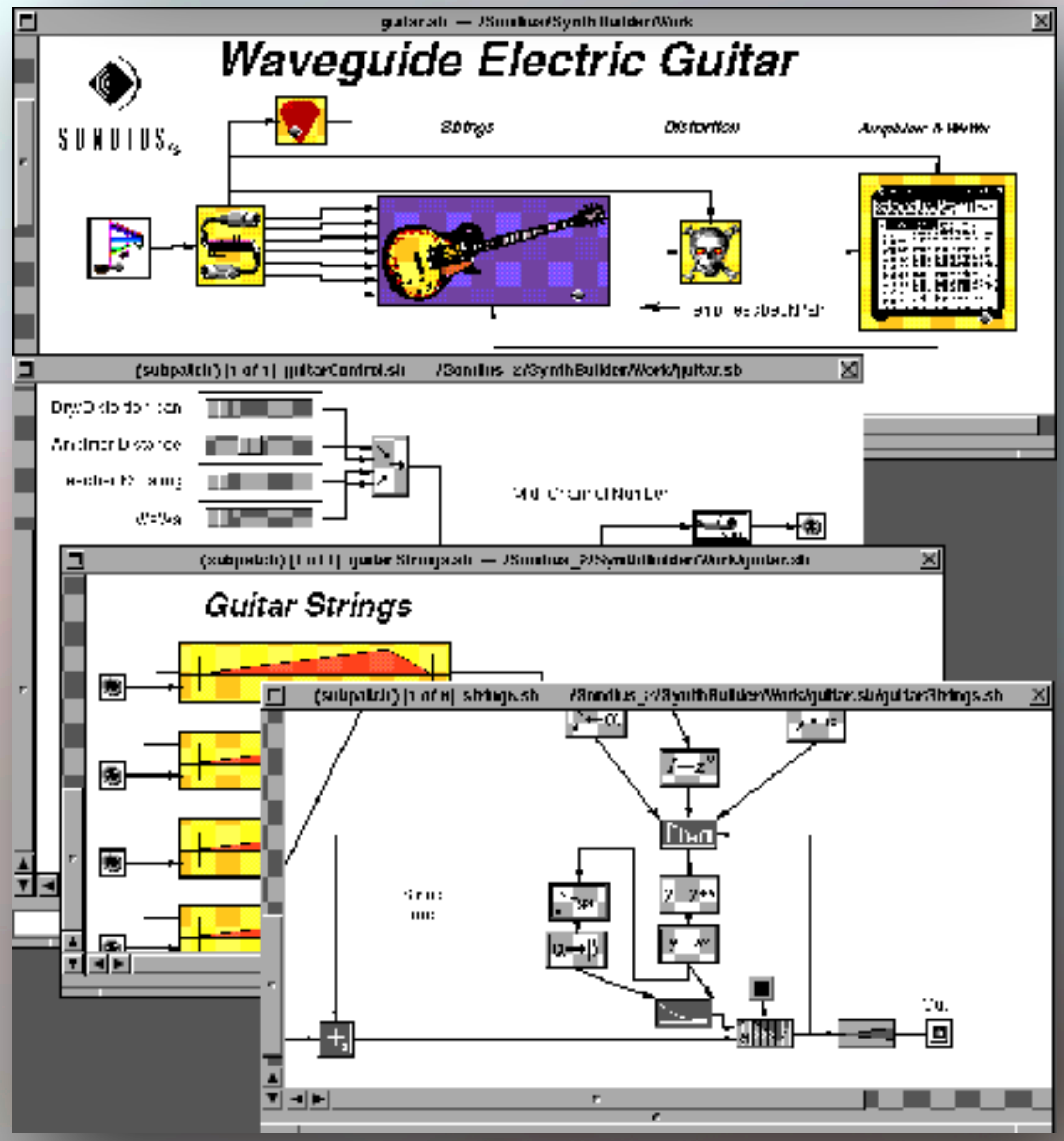




Guitar Model (Scandalis 1996)

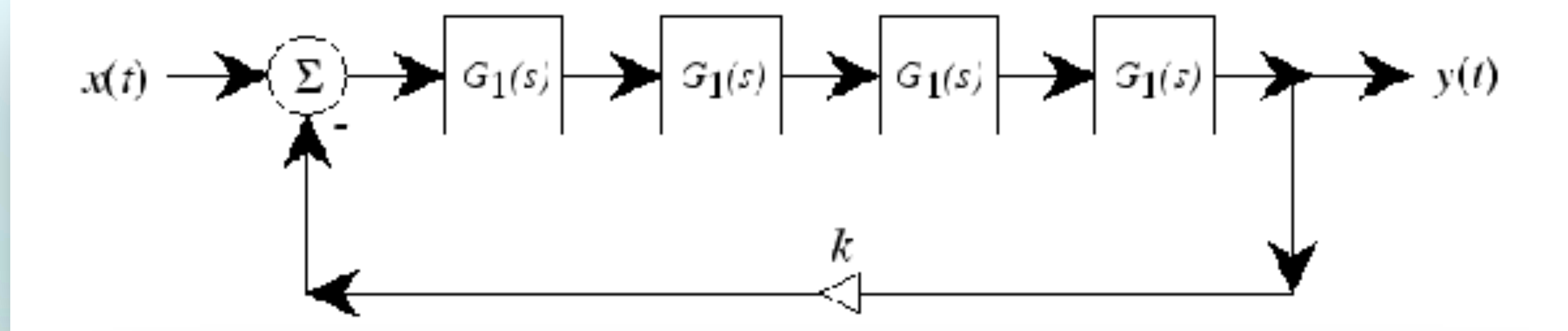
- Distortion, feedback and effects.
- Initial period excitations to capture the sound of different guitars.
- Controlled with Yamaha G10 guitar controller similar to today's MPE.

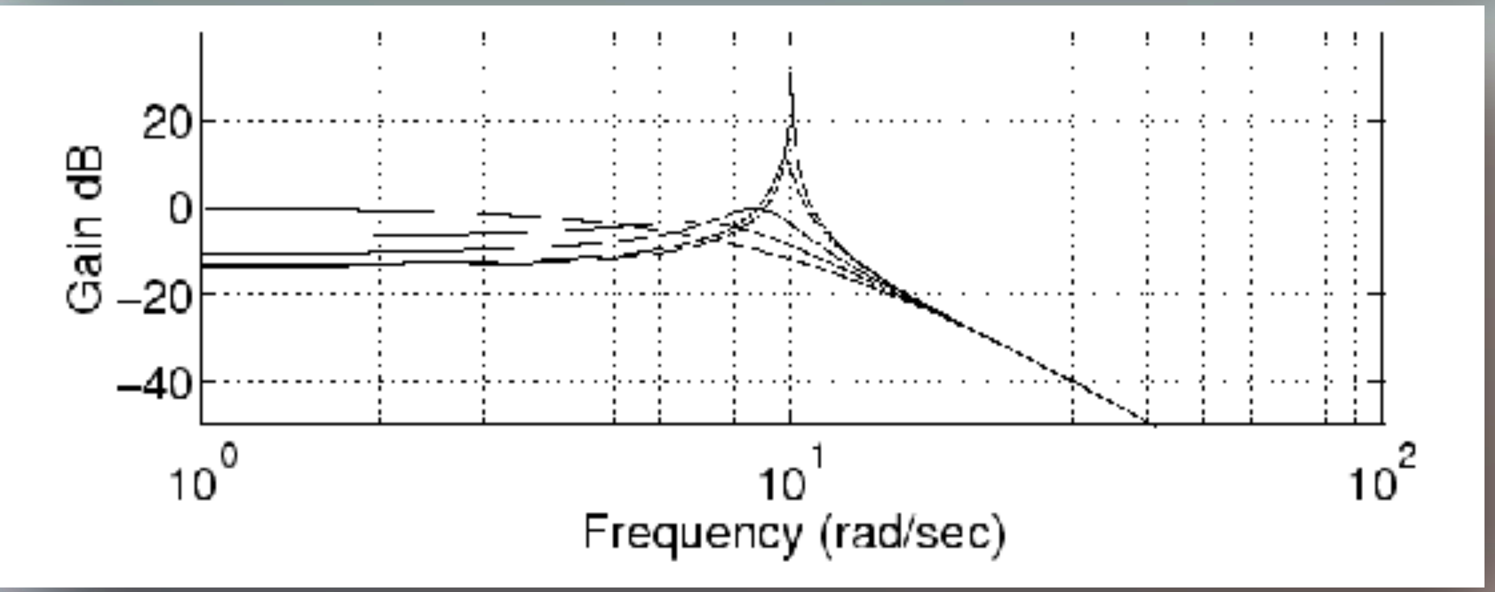


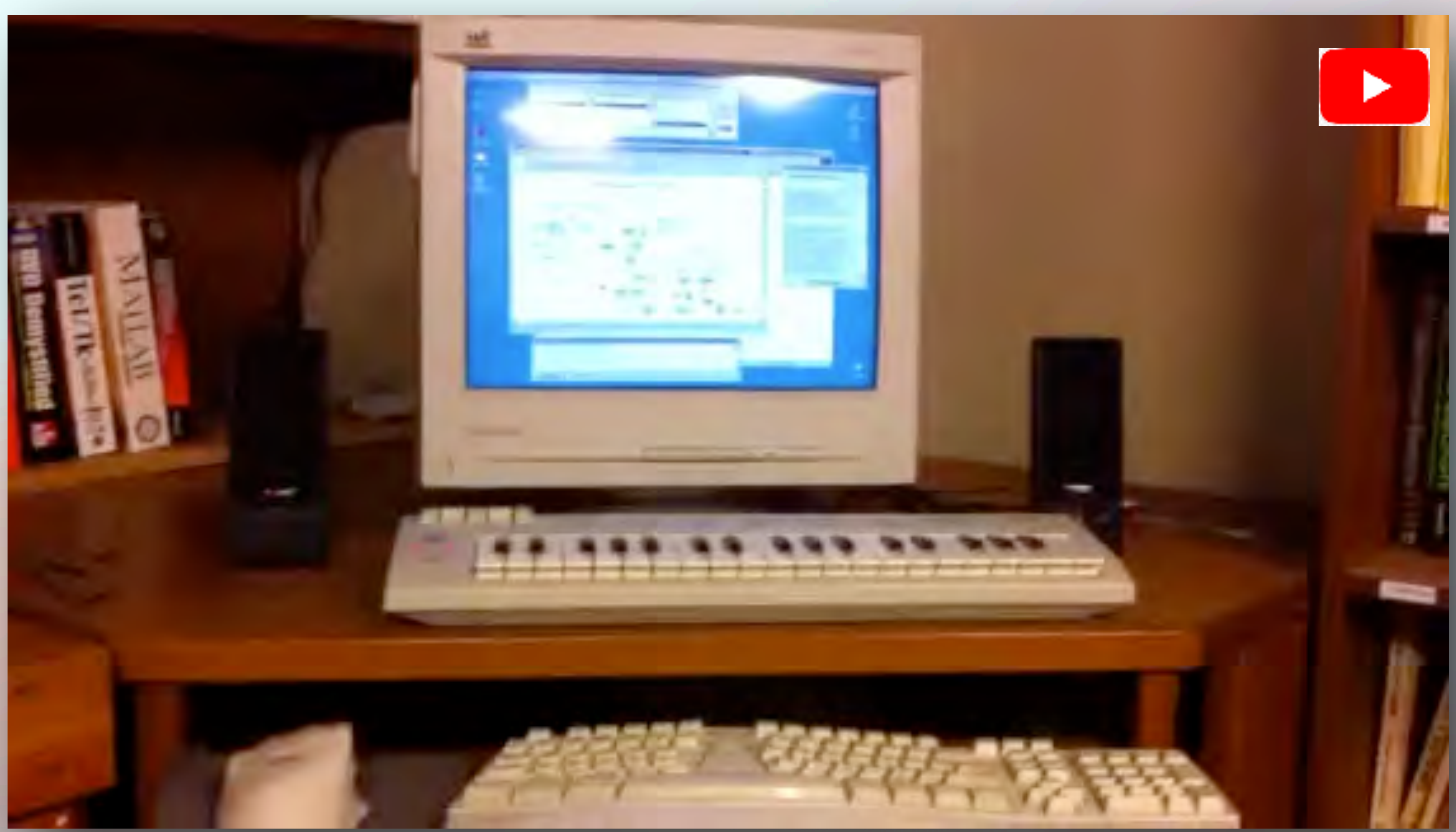


Virtual Analog (Stilson-Smith 1996)

- Alias-Free Digital Synthesis of Classic Analog Waveforms
- Digital implementation of the Moog VCF. Four identical one-poles in series with a feedback loop.
- Sounds great!

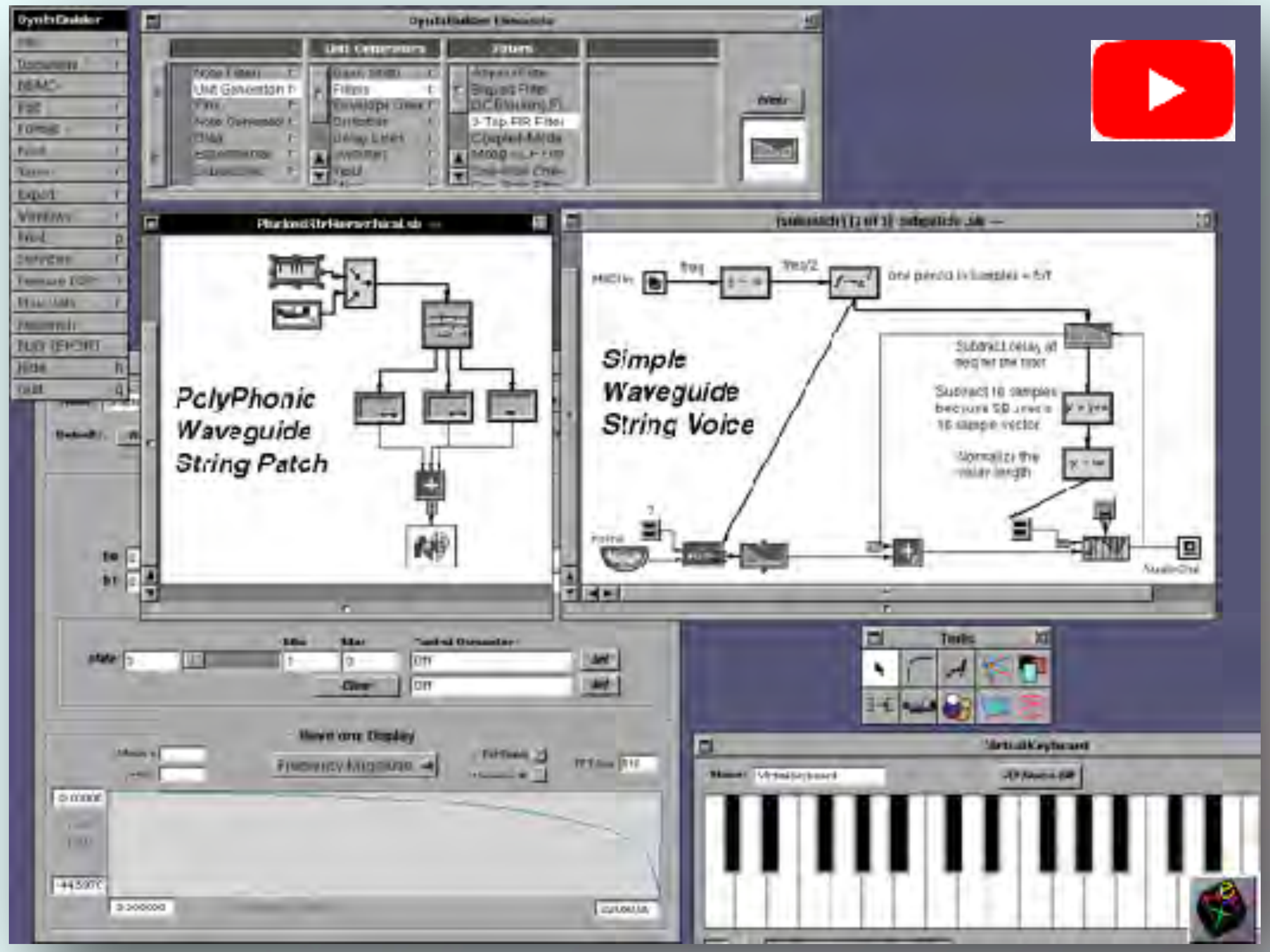


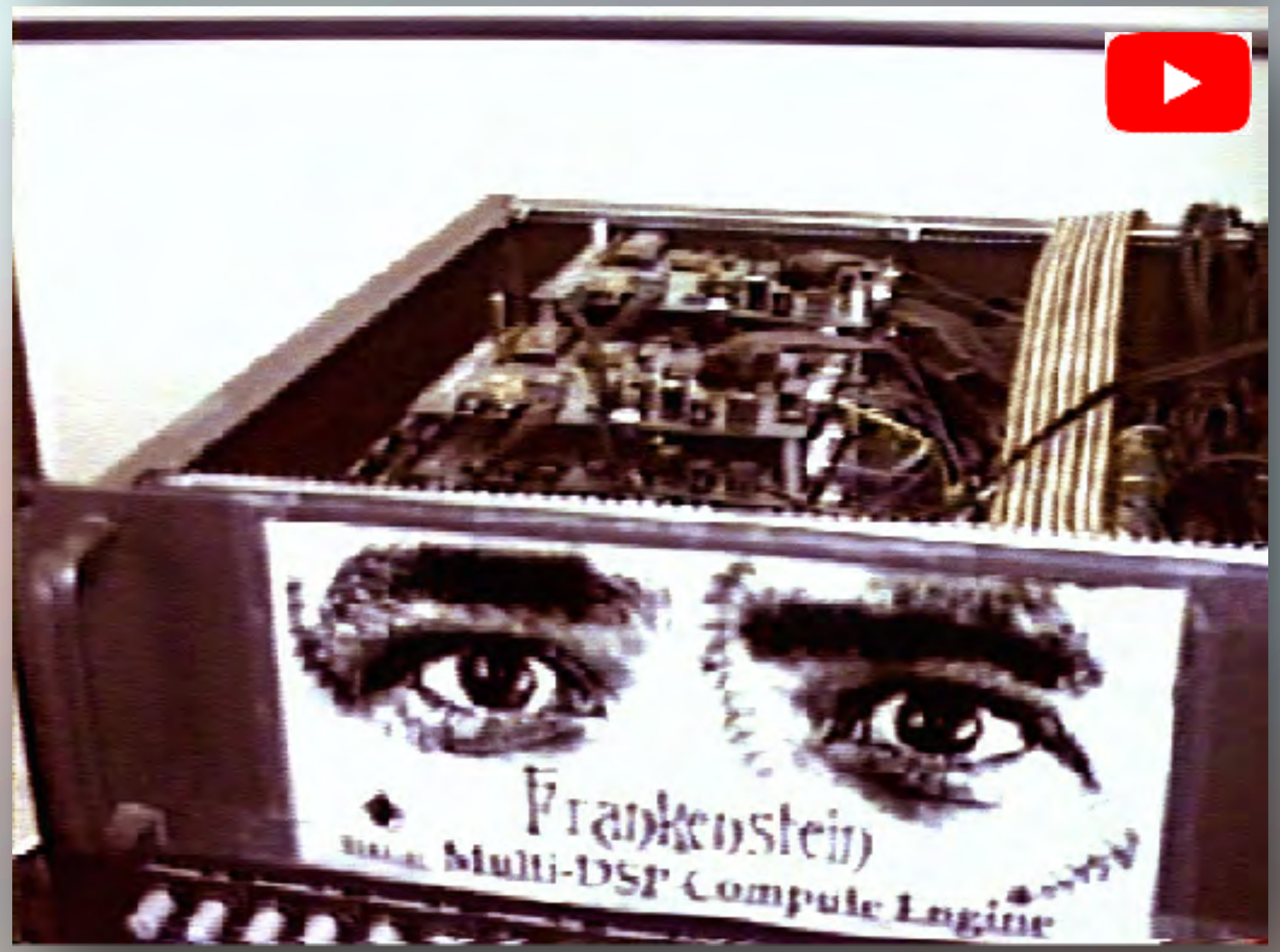




Full Ensembles all Physical Modeling (1997)

- Stanford OTL/CCRMA created the Sondius project to assist with commercializing physical modeling technologies.
- The result was a modeling tool, SynthBuilder, a DSP farm called Frankenstein, and a set of models covering about two thirds of the General MIDI set.
- Many modeling techniques were used including EKS, Waveguide, Commuted Synthesis, Coupled Mode Synthesis, Virtual Analog.

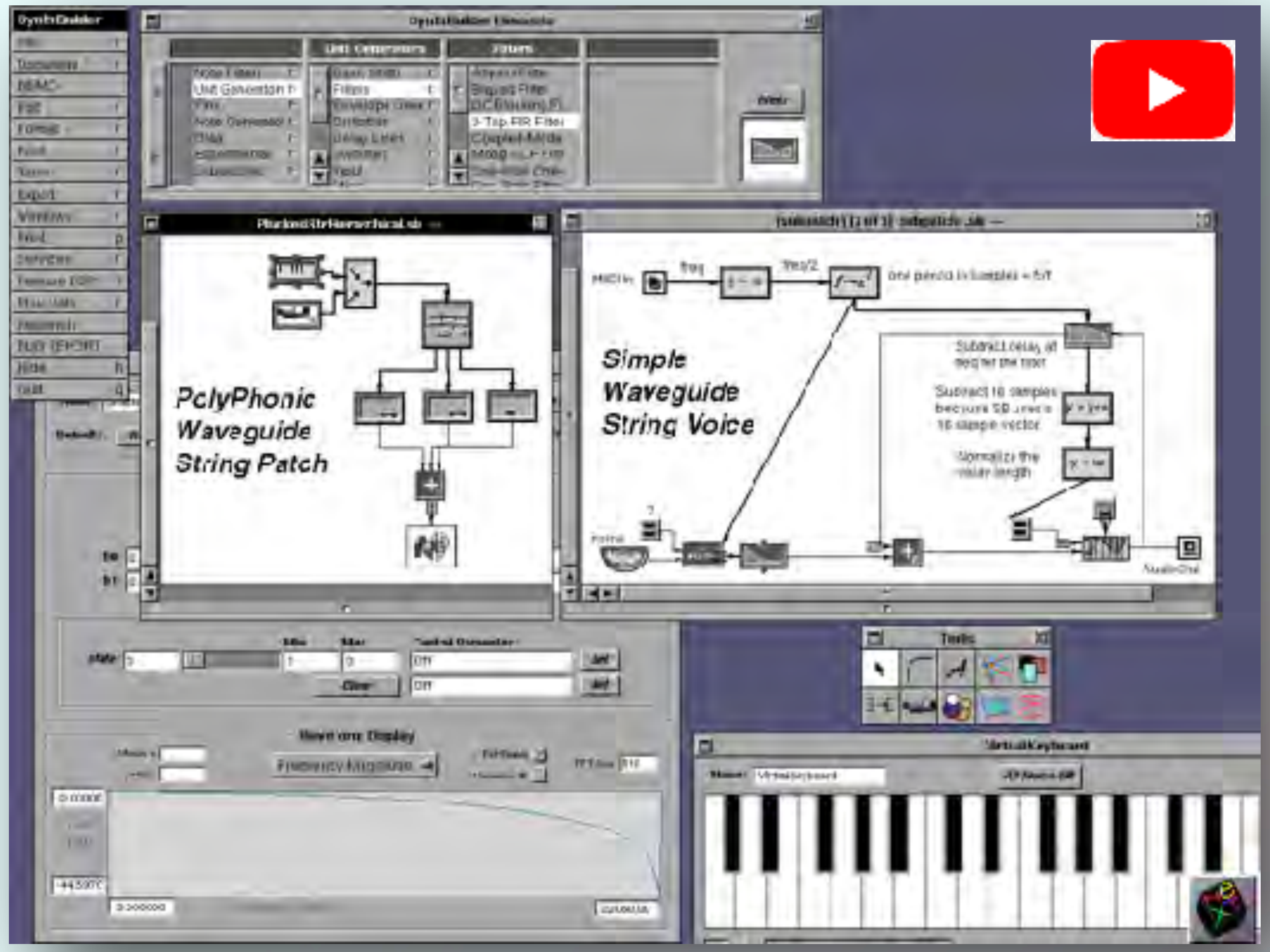


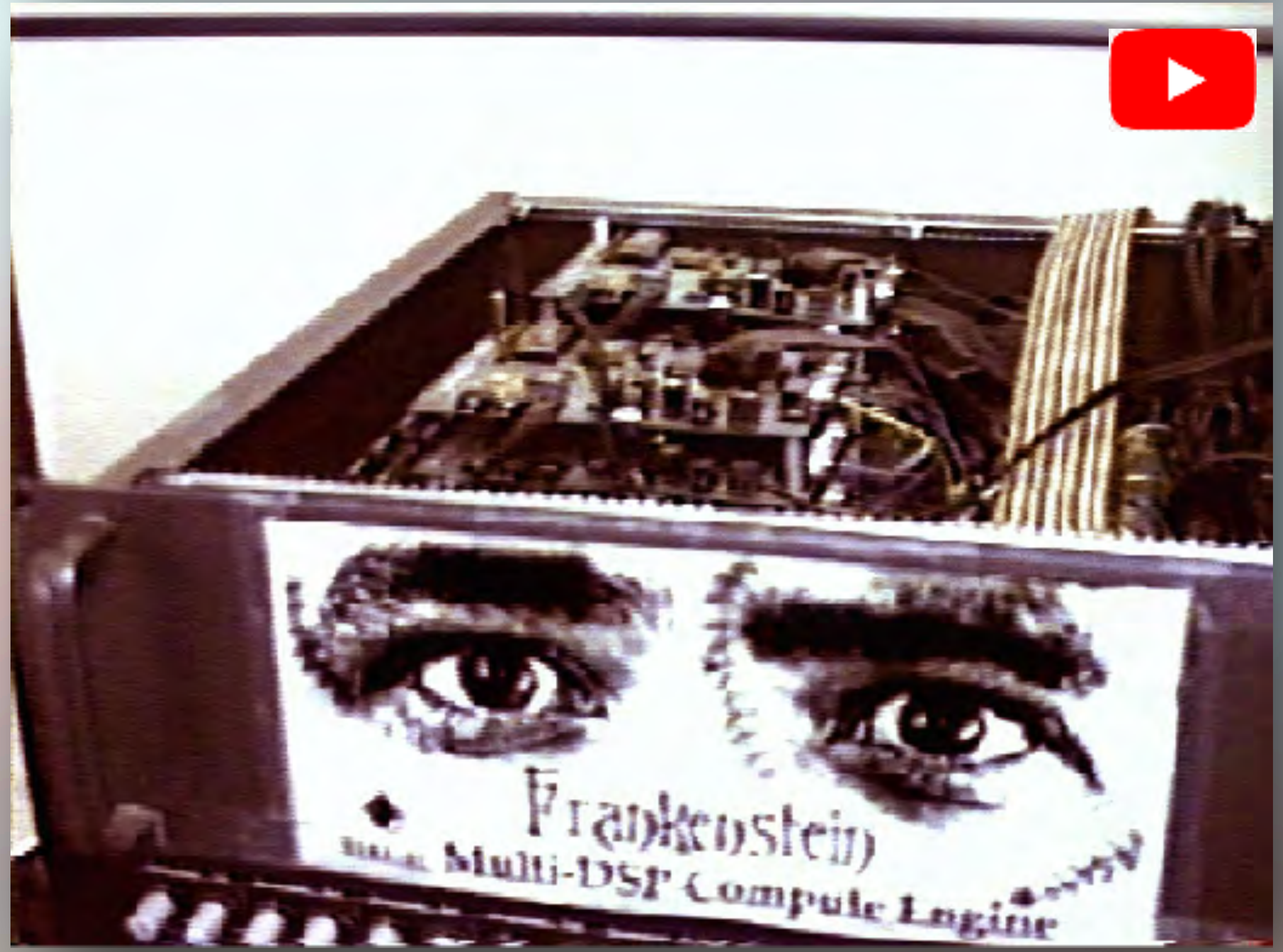


Full Ensembles all Physical Modeling (1997)

SONDIUS

- Stanford OTL/CCRMA created the Sondius project to assist with commercializing physical modeling technologies.
- The result was a modeling tool, SynthBuilder, a DSP farm called Frankenstein, and a set of models covering about two thirds of the General MIDI set.
- Many modeling techniques were used including EKS, Waveguide, Commuted Synthesis, Coupled Mode Synthesis, Virtual Analog.





First Generation PM Products

- Yamaha VL-1 + Chipsets (1994-2000)
- Korg SynthKit … Kronos (1994-present)
- Seer Systems Reality (1997)
- Aureal ASP 301 Chip (1995-1997)
- Staccato SynthCore Sondius Models (1997-2001)

In 1994 Physical Modeling Was Poised to be the “Next Big Thing”, So What Happened?

- Computationally expensive
- Developing the models is hard. Lots of research.
- Voicing PM is difficult (like FM), voicing samples is more direct.
- Controllers that could express multiple dimensions were not common.

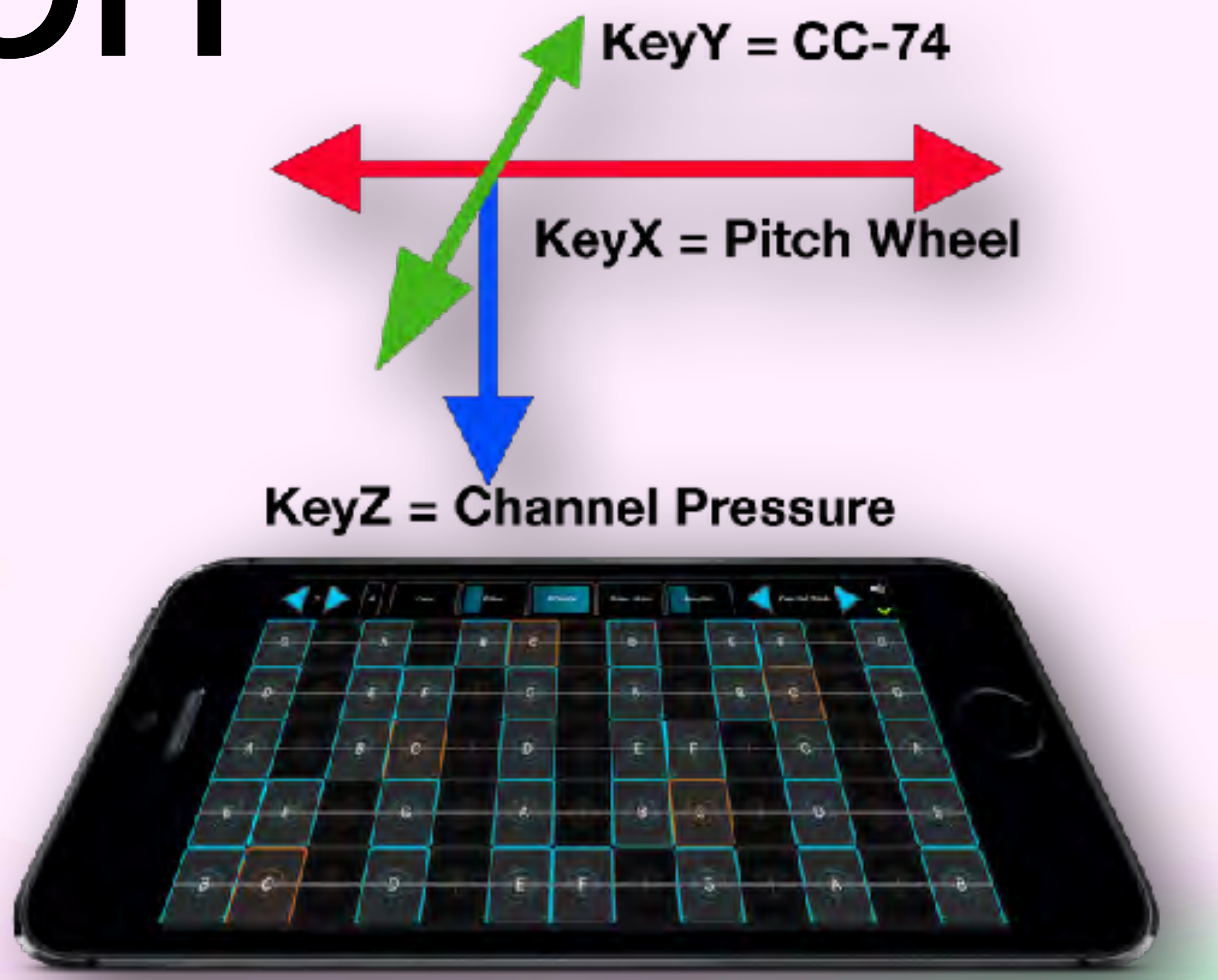
In 1994 Physical Modeling Was Poised to be the “Next Big Thing”, So What Happened?

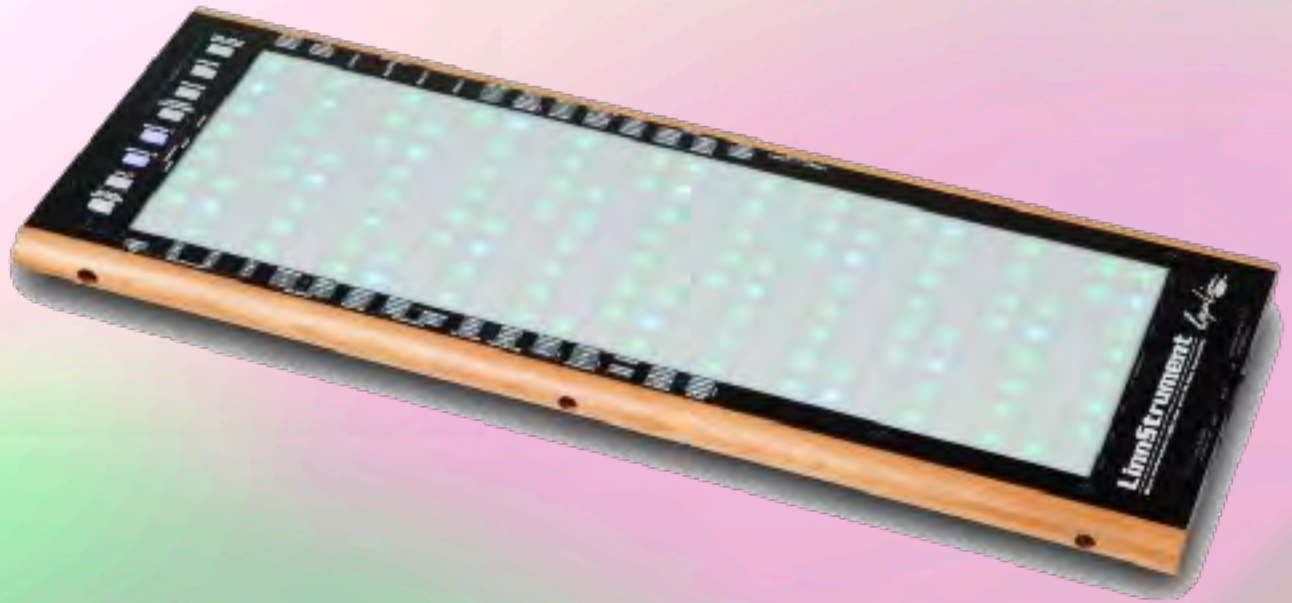
- By 1994, FM was the standard for PC Game Music. In part due to it’s small memory footprint.
- PM was seen by Yamaha as the successor to FM (John Chowning’s pioneer FM patent was expiring).
- The cost of memory starting plummeting in 1996. Sampling became common.
- Some expressivity could be achieved with extensively interpolated samples.
- **Sampling Won!**

Why is PM Back?

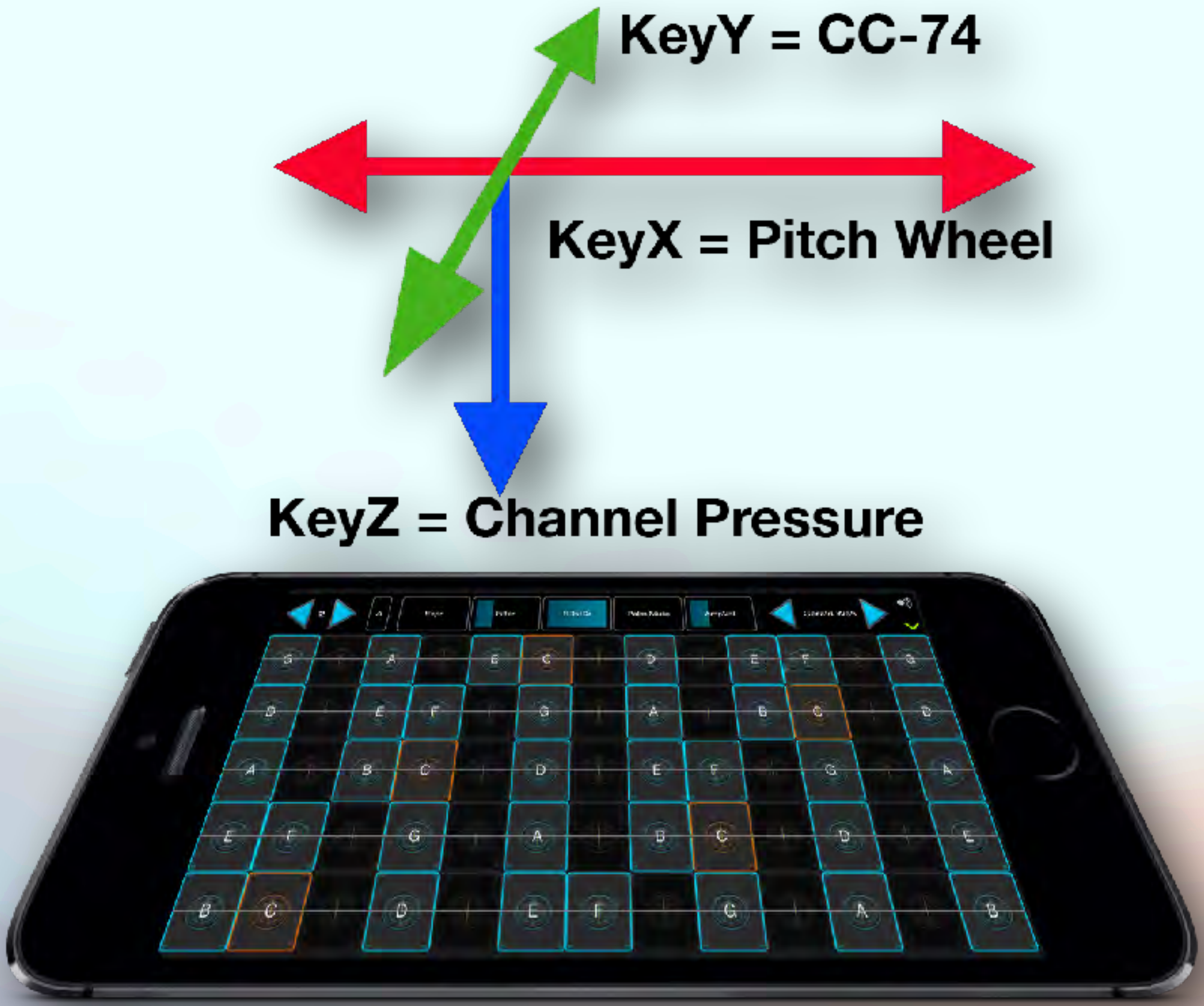
- **Compute Power** - A DSP Farm is no longer required.
- **Lots of Models Now** - Models require research to create and calibrate.
- **MPE** - There is a new generation of polyphonic expressive controllers based on the MIDI MPE spec.
- **Calibration -** Emerging research to calibrate physical models to specific instruments

MPE is an Enabling Technology for Composition and Orchestration



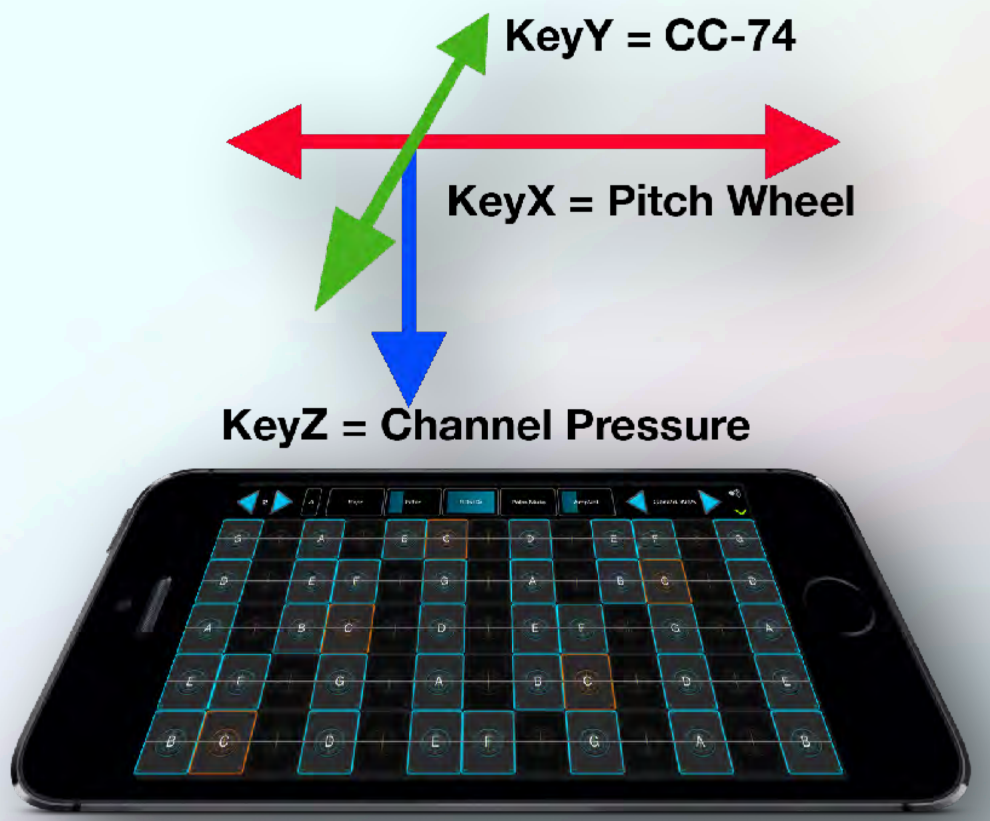


MPE = MIDI Polyphonic Expression
MPE + Modeling = BIG DEAL



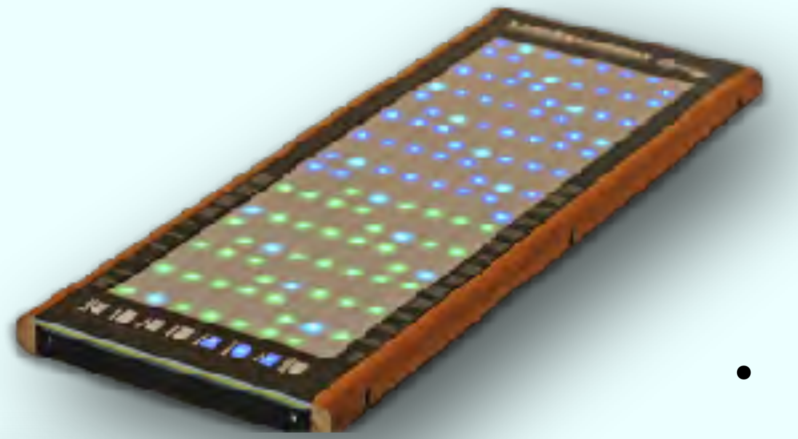
MPE in a Nutshell

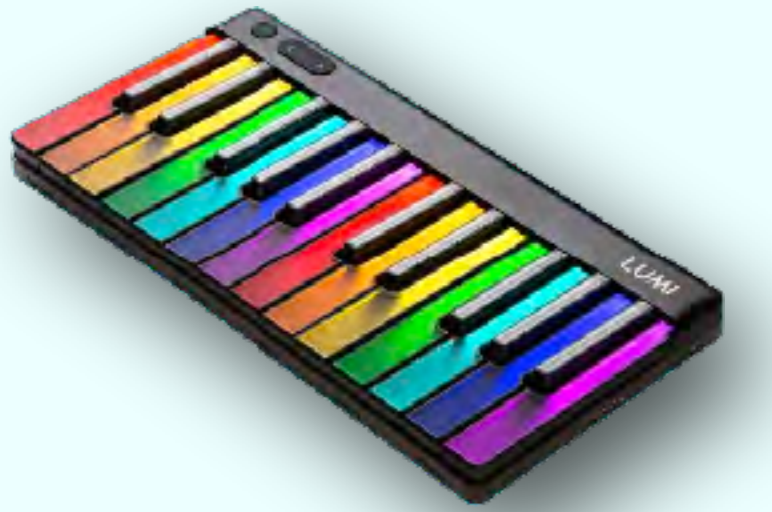
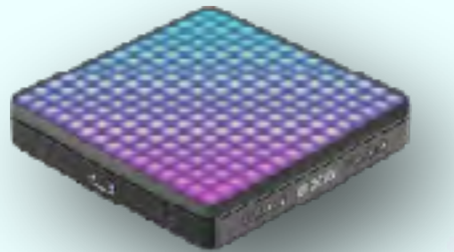
- Derivative of MIDI Modes 3/4; enabled with RPN-6/0
- Can be Channel-Per-Note (for Keyboards, like the Seaboard) or Channel-Per-Row (String) (GeoShred, LinnStrument, Guitar Controller).
- Expression Control Conventions (per Channel)
  - o KeyX – Pitch Bend (Roli calls this *Glide*)
  - o KeyY – CC-74 (Roli calls this *Slide*)
  - o KeyZ – Channel Pressure (Roli calls this *Press*)

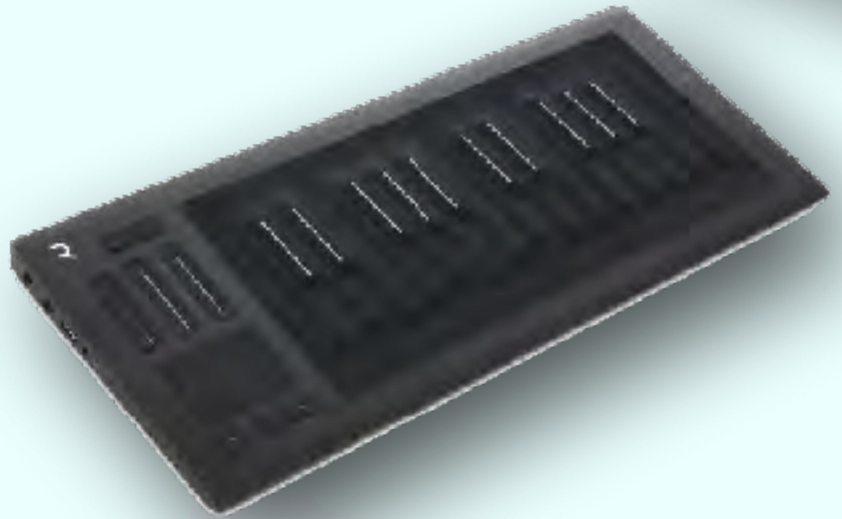
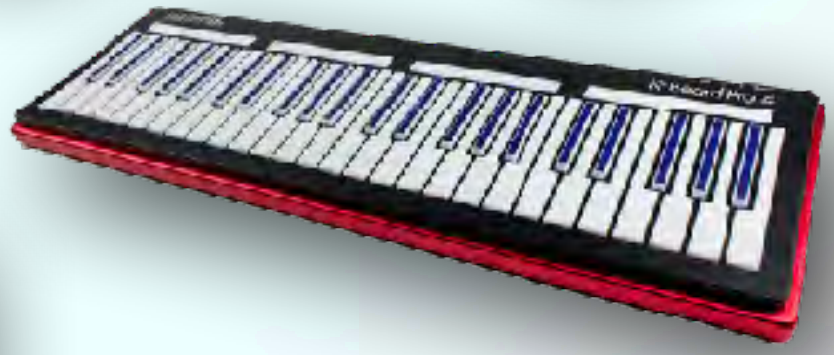


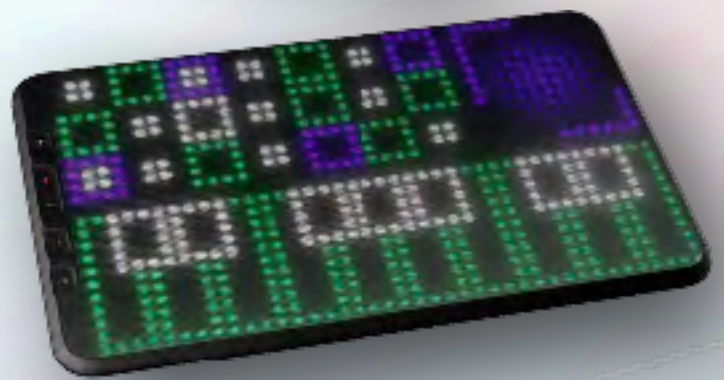
- Provides for Manager Channel (typically 1 or 16) that globally controls the MPE Member Channels (ie modWheel to all Member Channels)

Some MPE Controllers





- Haken Continuum
- Lumi Keys
- Roli Blocks
- KMI K-Board Pro 4
- Ere Touch
- Ableton Push 2
- Exquis
- Osmose
- Guitar Controllers
- Sensel Morph
- Artiphon INSTRUMENT 1
- Joué
- GeoShred
- Roli Seaboard
- LinnStrument

MPE makes a whole new generation of controllers possible. **Whatever instrument makers dream up!**





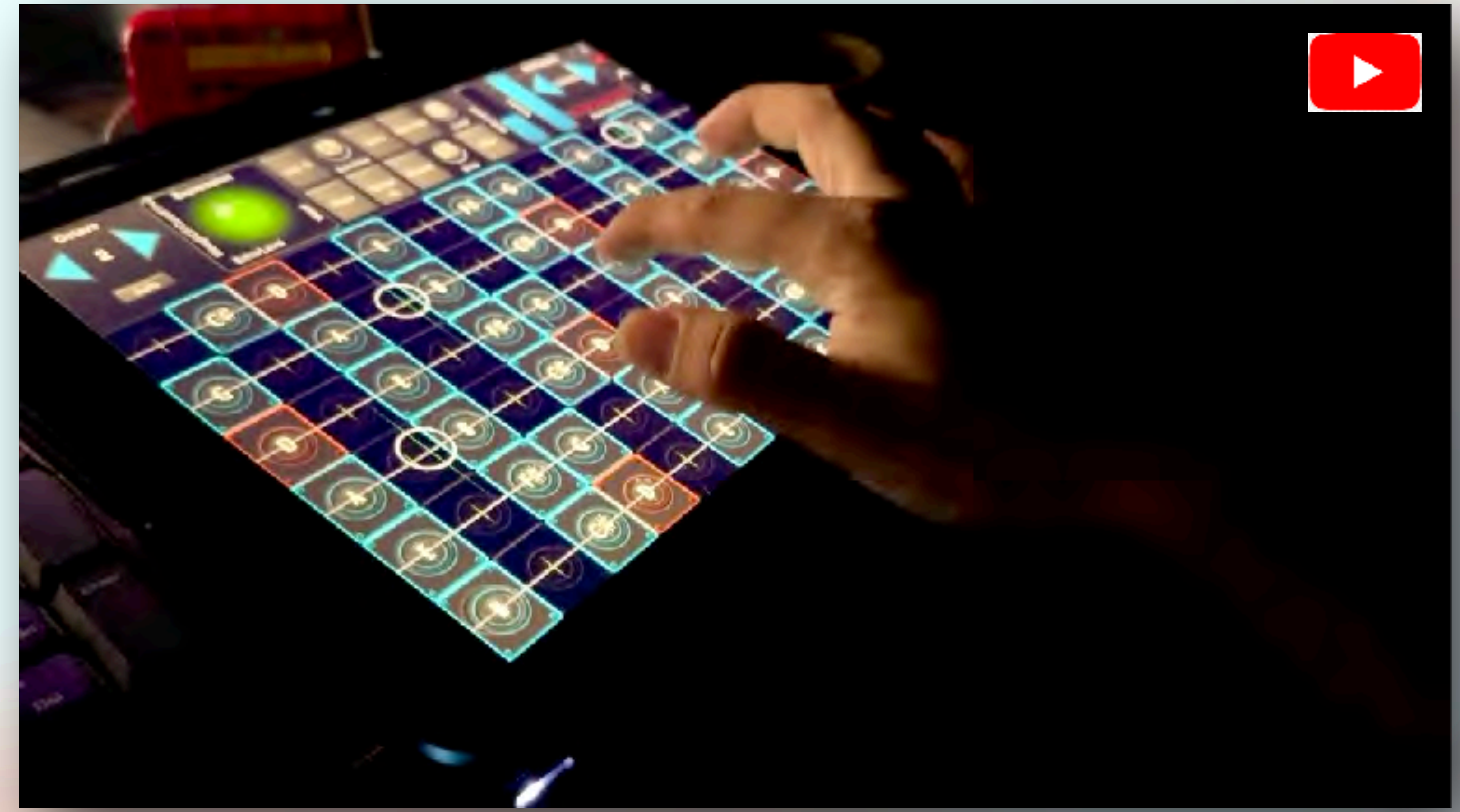


MPE and The Importance of Pitch Fluidity

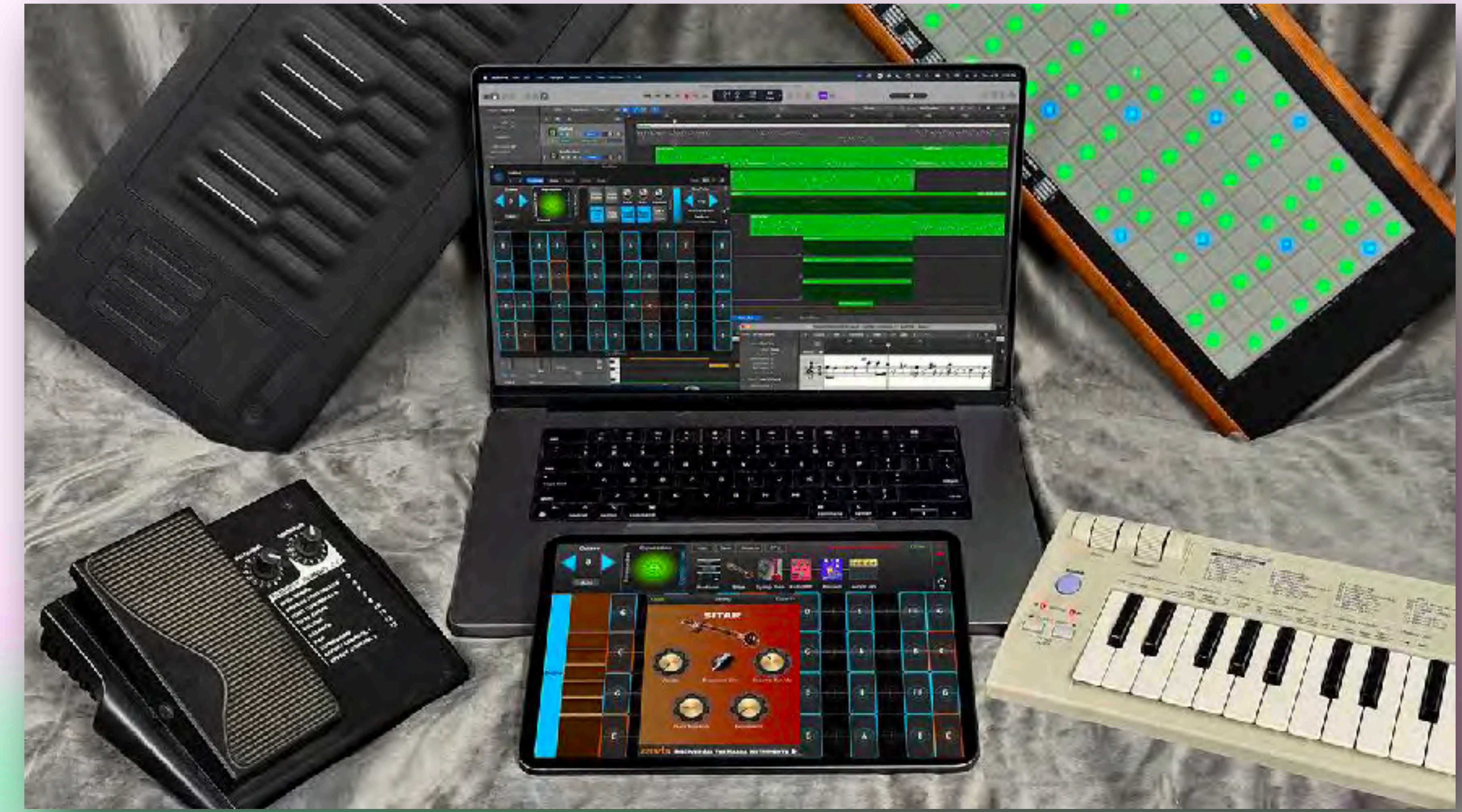
- Pitch Fluidity is an essential expressive metaphor for musical performances around the world.

- MPE directly addresses Pitch Fluidity by supporting per-note, multi octave pitch Bending.

- Though not a part of MPE, Pitch Rounding is essential to enable performers to play in-tune in any given temperament. Roli Seaboard, LinnStrument, GeoShred, et al support pitch rounding.



Full Musical Ensembles Using Physical Models



Full Musical Ensembles 2026

About SWAM

- Created and sold since 2017 by Audio Modeling, an Italian company run by Stefano Lucato, Emanuele Parravicini, and Simone Capitani.

- SWAM instruments model 33 string, woodwind, and brass instruments. They are fundamentally based on digital waveguide synthesis technology.

- Performable, real-time synthesis instruments with many MIDI-controllable parameters. Parameters can be mapped to curves that optimize their response in performance.

- Highly expressive, lifelike, and MPE-capable.

- Have MUCH smaller RAM and disk footprints than sample instruments and can load presets nearly instantly.

- Can run standalone or as plugins in DAWs, and also come in iOS versions. GeoSWAM is a family of SWAM instruments built to run in MoForte GeoShred.

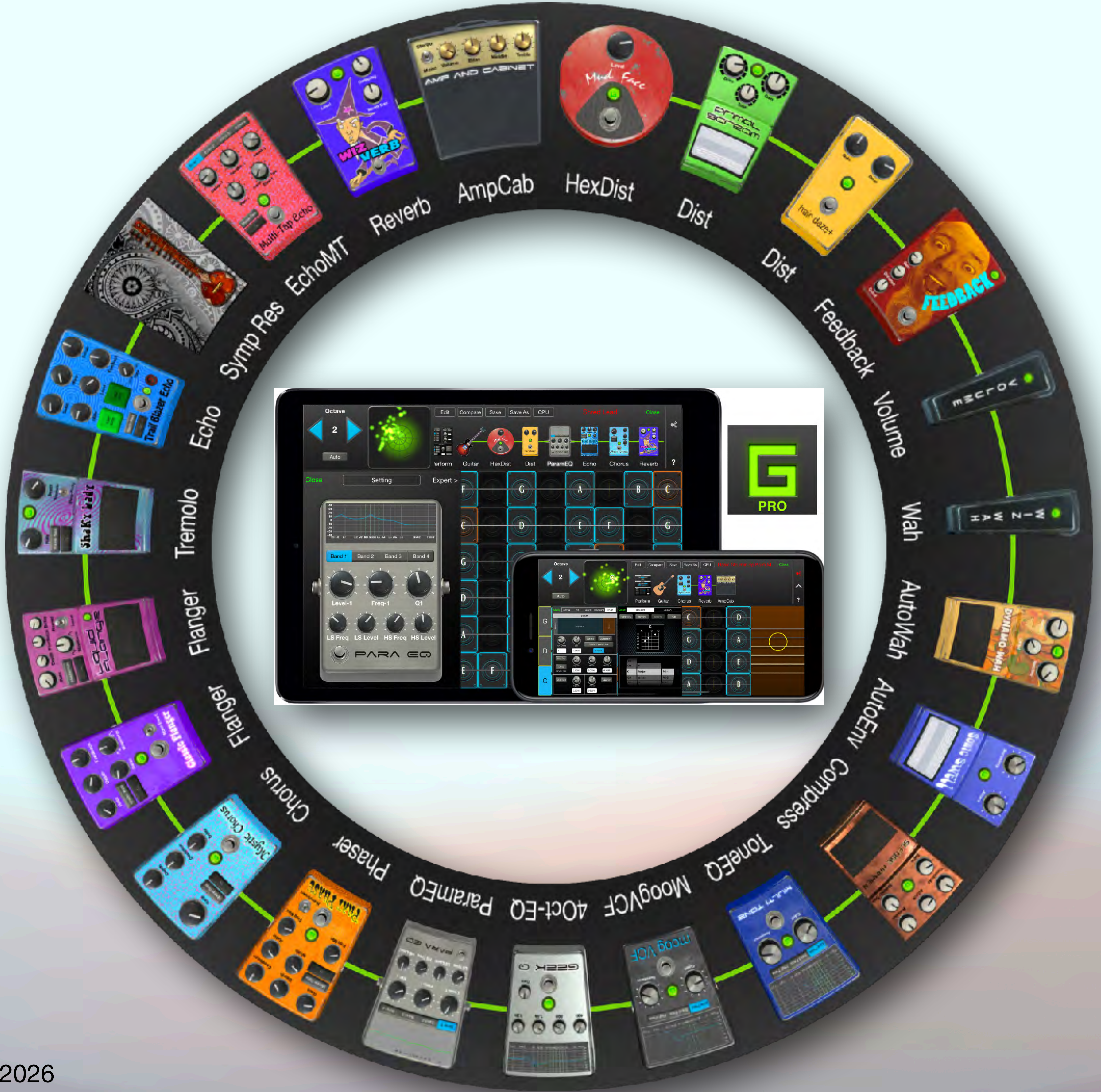
About GeoShred

A Framework for Modeled Instruments and Effects

45 Physically Modeled Instruments

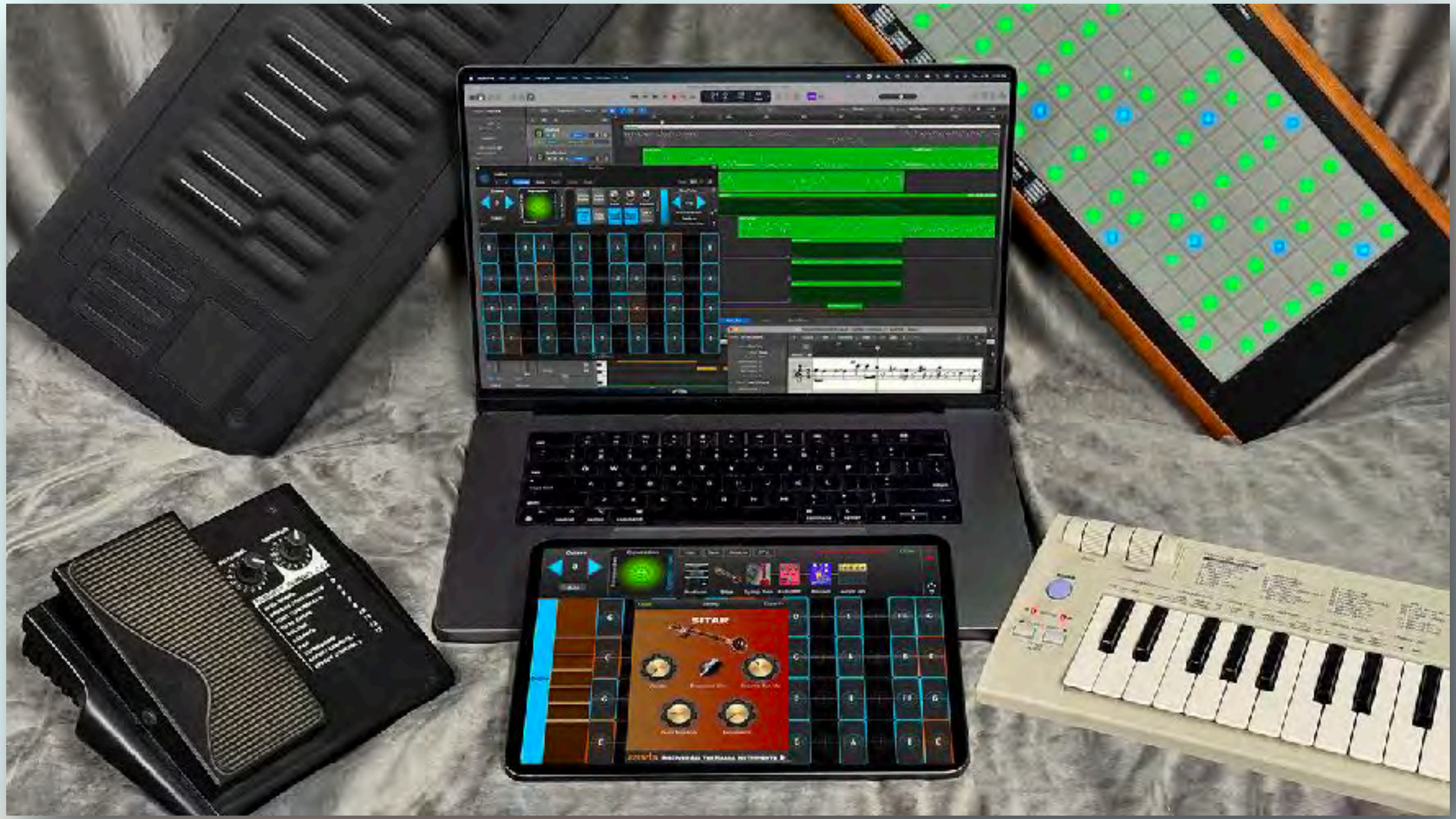
- Guitar Instruments
- 11 GeoSWAM Instruments
  Orchestral
- 33 Naada Instruments
  Asian

22 Modeled Effects - Circuit Models



GeoShred

- iOS/iPadOS and macOS. Windows 2026
- 270k users world wide in 120+ countries. 40% of users are in India
- Unique isomorphic MPE keyboard with “Almost Magic” Pitch Rounding.
- GeoShred Keyboard has XY expression on iPad and XYZ on iPhone 8,9,10.
- Supports GeoShred Keyboard, MPE Controllers, Conventional MIDI Controllers and Wind Controllers.

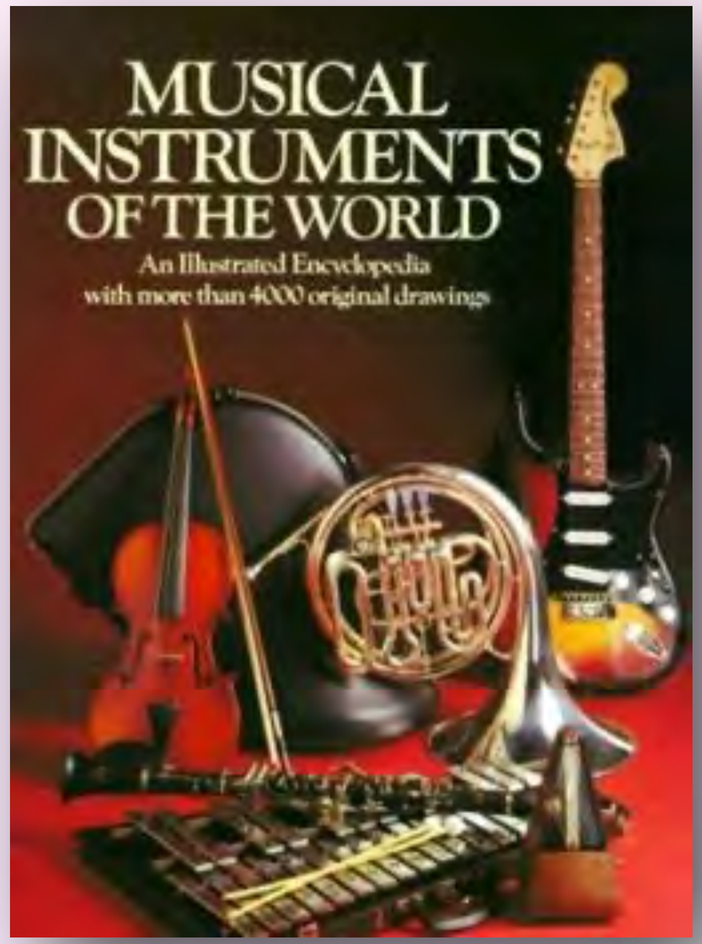


Physical Models for Orchestration

Jordan's Thoughts on Physical Models as a Pallette of Colors

Prog Orchestration

The Near Future …



- More models. AI calibrated models will replace some sample libraries and eliminate the need for key switches.
- MIDI 2
- The MIDI 2 Orchestral Articulation Profile.
- What does Orchestral Articulation mean for live performance film/game scoring and virtual performers.

MIDI-CI Profile for Note On Selection of Orchestral Articulation

Using MIDI 2.0 Protocol

Document Version 1.0

Published 2024-04-05

Developed and Published By
The MIDI Association
and
Association of Musical Electronics Industry (AMEI)

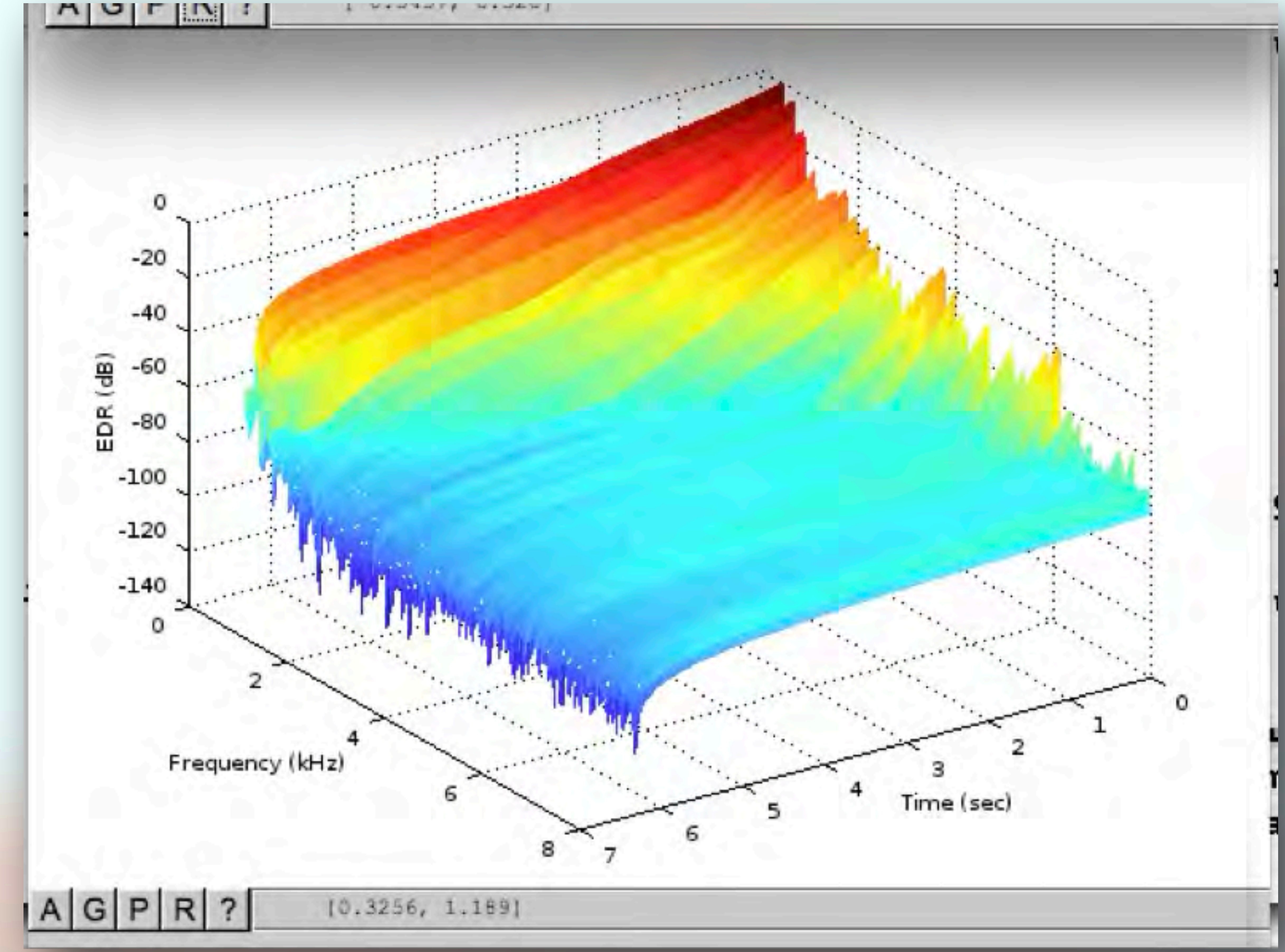
About Orchestral Sample Library Key Switches

- Key Switches are used to select alternate articulated samples
- Articulation is sometimes done in post editing
- **Note that Physical Models generally use MPE and CCs so that the articulated behavior can be performed directly**

	Spitfire	EastWest	Berlin Strings	Cinematic Studio	Vienna
C0	Long (Sustained)	Sustain	Sustain	Legato	Sustain
C#0	Legato	Legato Slur	Legato	Sustain	Legato
D0	Staccato	Legato Bow	Staccato	Staccato	Staccato
D#0	Spiccato (Bouncing Stroke)	Porato	Spiccato	Spiccato	Spiccato
E0	Pizzicato	Staccato	Pizzicato	Pizzicato	Pizzicato
F0	Col Legno (with the wood)	Spiccato	Tremolo	Conlegno	Tremoo
F#0	Tremolo	Pizzicato	Trill m2	Tremolo	Trill m2
G0	Trill m2	Tremolo	Trill M2	Trill m2	Trill M2
G#0	Trill M2		Marcato	Trill M2	Sul Ponticello
A0	Harmonics		Sul Ponticello (near the bridge)	Harmonics	Harmonics
A#0	Sul Tasto (near the fingerboard)			Con Soridino	Detaché
B0	Con Sordina (with mute)				

Models Are Calibrated With Real Recordings

- **Currently** model calibration is a filter design problem.
- For the guitar, samples are collected for every fret for every string
- Analysis done in Matlab. Goal is to design low order loop filters that match the partial decay rates in the original recordings. The process is labor intensive.
- **Moving forward, a trained AI may be better at analyzing recordings and designing model filters**. Whole sample libraries can be converted to expressive models,



AI Calibrated Models to Specific Instrument

- The problem is stated as a “Synthesis Sound Mathching Problem”.
- At a high level there is a synthesis Algorithm with “Knobs”, recordings of the target sound and some process compute the values of the “Knobs” so that the synthesis algorithm matches the target sound.
- This is currently a hot research topic, combining DSP, Optimization and Neural methods.

Neural Parameter Estimation for Musical Sound Synthesis

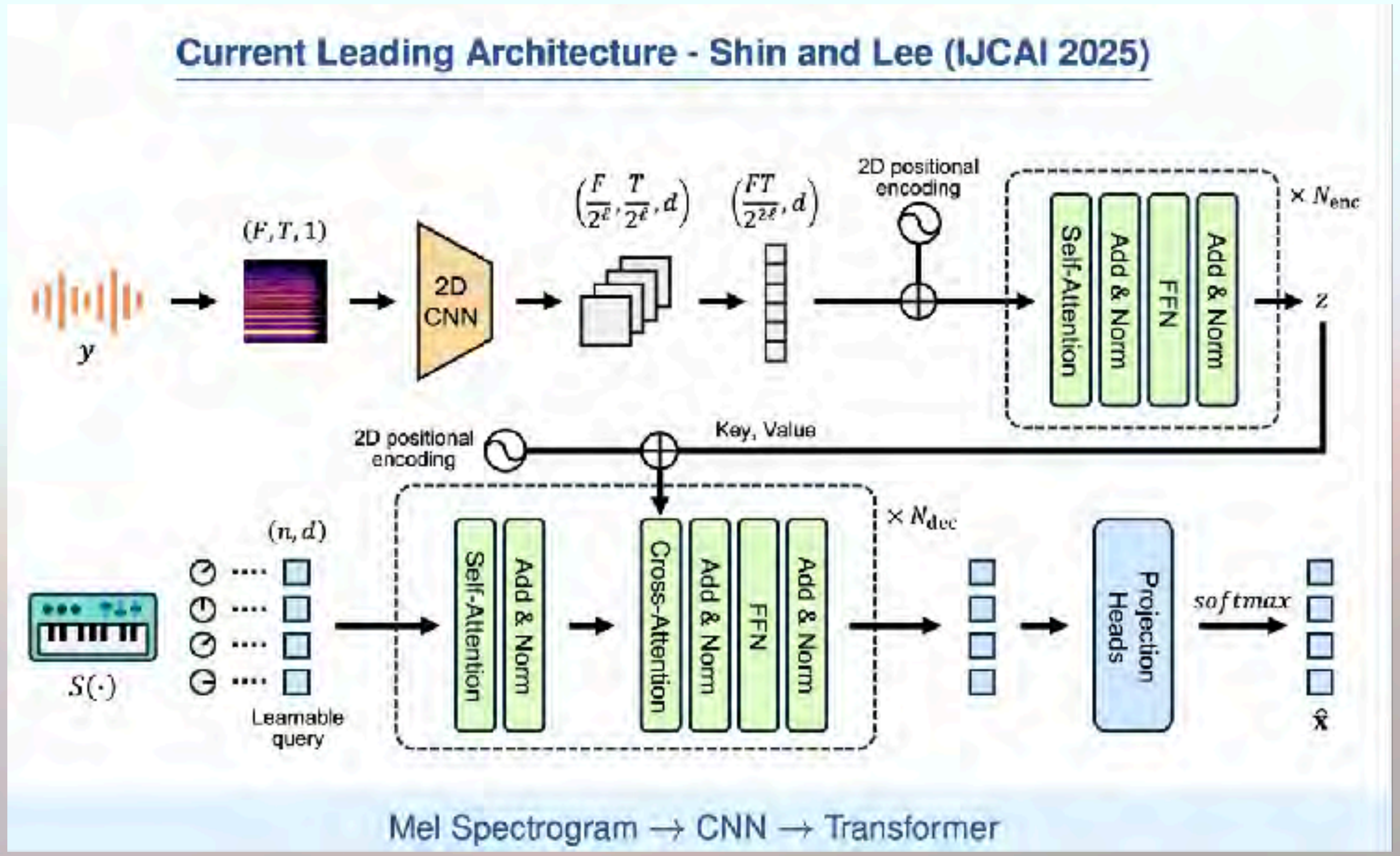
Julius O. Smith III, Gabriel Soule, Soohyun Kim, Jennifer Zheng
CCRMA, Stanford University

Fifth Meeting, Acoustical Societies of America and Japan
Session ”Music Data Science”

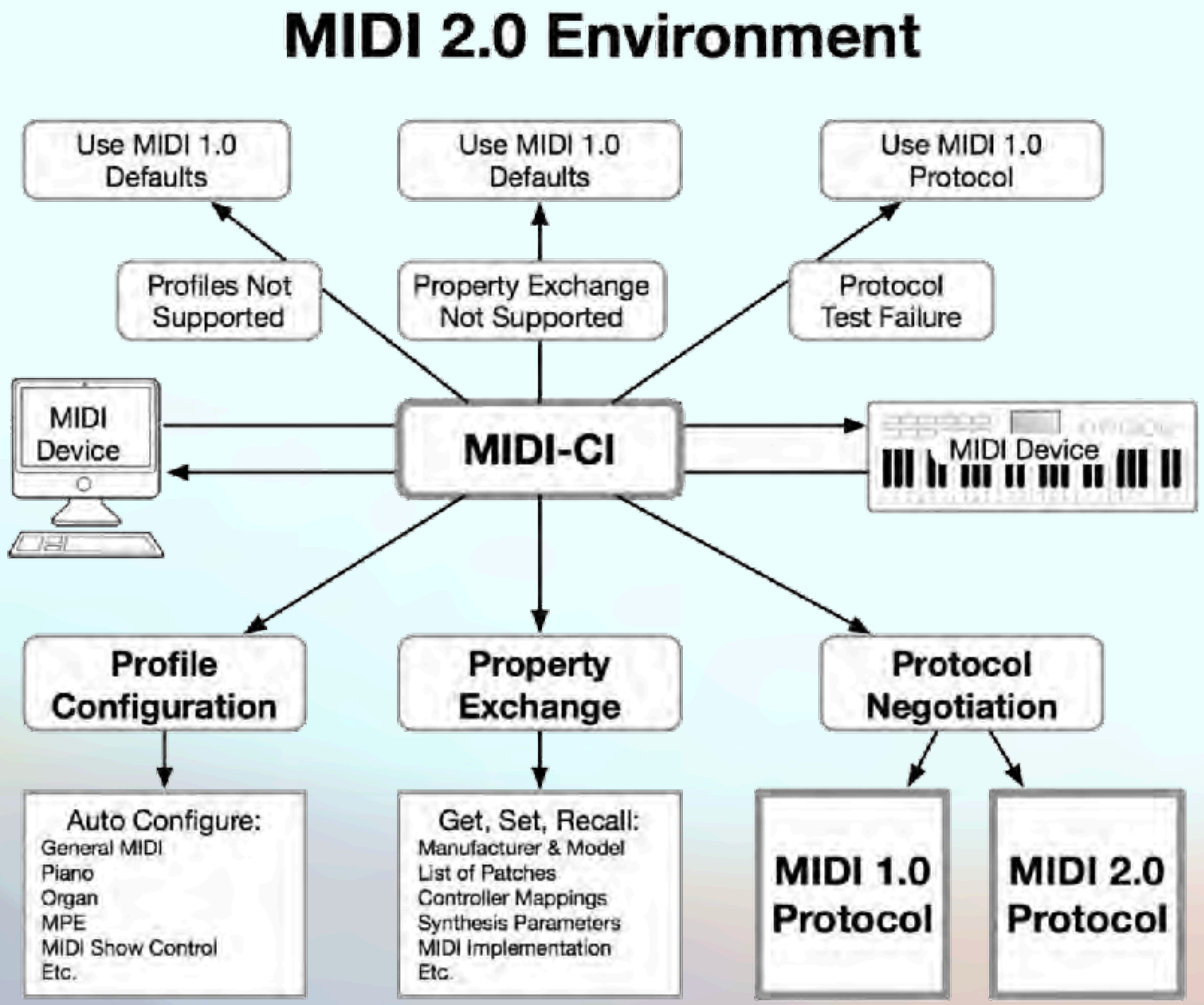
December 1-5, 2025

AI Calibrated Models to Specific Instrument

- This research applies not only to physical modeling, but other parametric synthesis methods like FM, Subtractive.
- At some point in the near future it should be possible to calibrate physical models to specific instruments, just as a sample library is recorded from specific instruments.

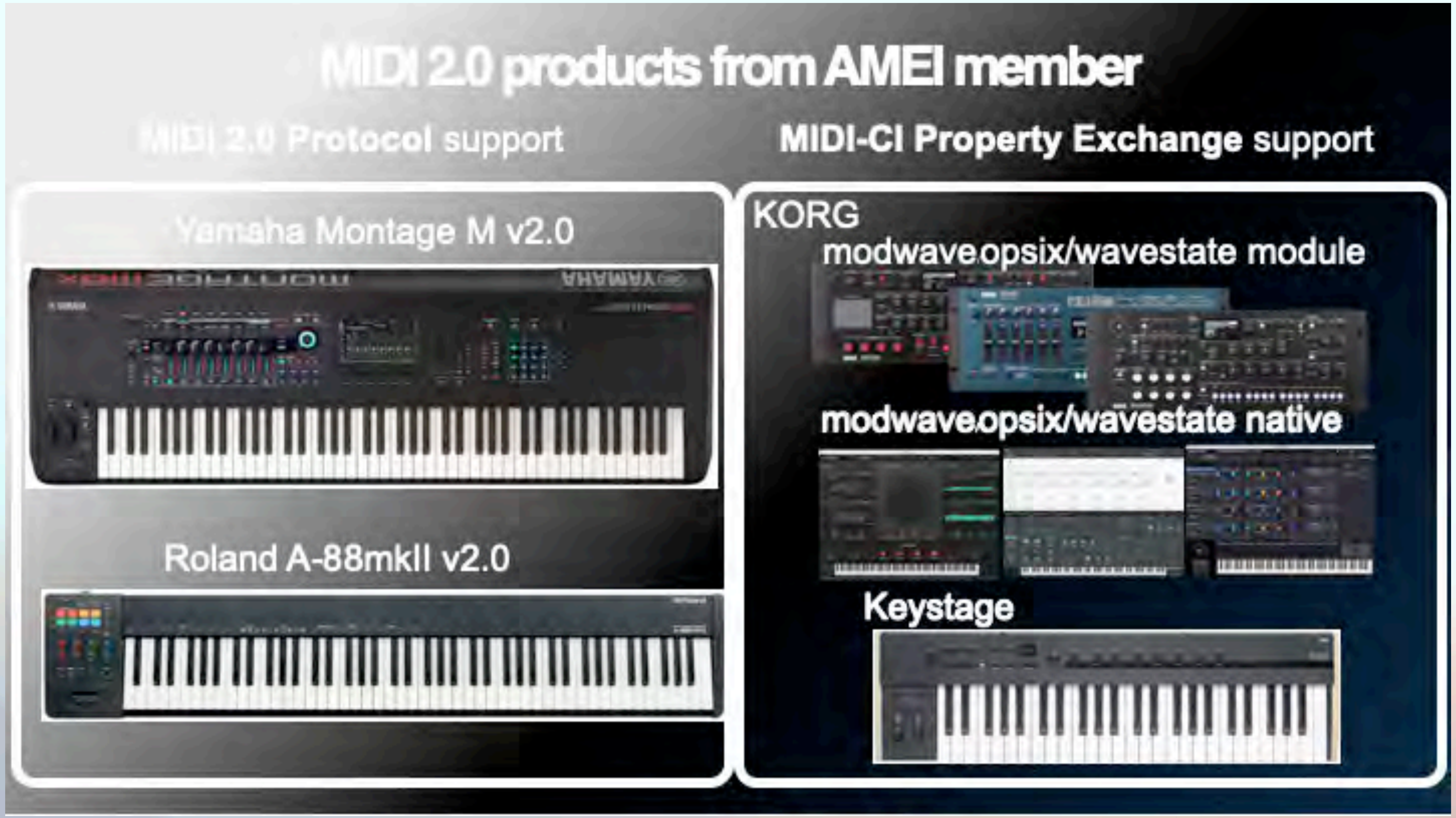


MIDI 2



- MIDI 2.0 is bidirectional. It changes MIDI from a monolog to a dialog
- MIDI 2.0 Protocol mirrors and extends the MIDI 1.0 Protocol.There is a new Universal MIDI Packet (UMP) which offers higher resolution performance controllers (32 bit), more controllers (32k) and 16 channel groups for 256 channels.
- MIDI-CI (Capability Inquiry) supports profile and property exchange which can be used to configure devices for specific applications.
- Compatible with MIDI 1.0
- Future Proofing, UMP is transport agnostic and can be implemented on USB, Ethernet, Bluetooth and future transport mechanisms.

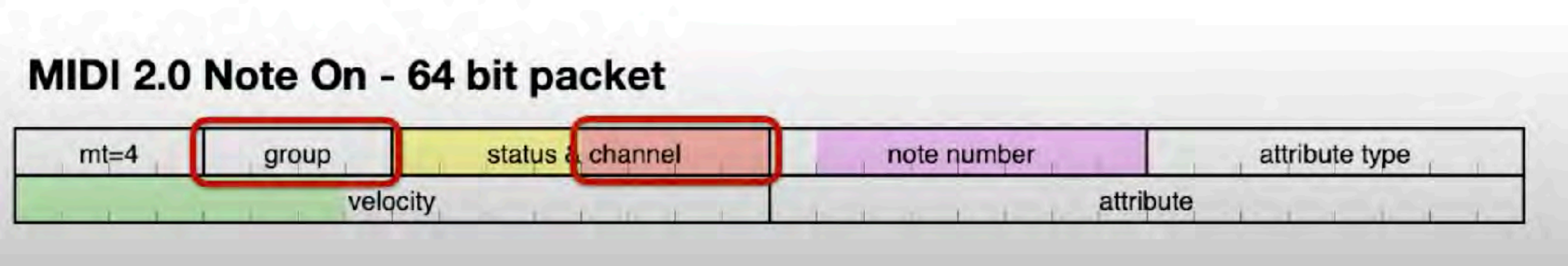
MIDI 2 Status



- Available in Linux, macOS, iOS/iPadOS, Windows, Android.
- Yamaha, Korg, Roland have MIDI 2 products available.
- DAW companies will soon support MIDI 2

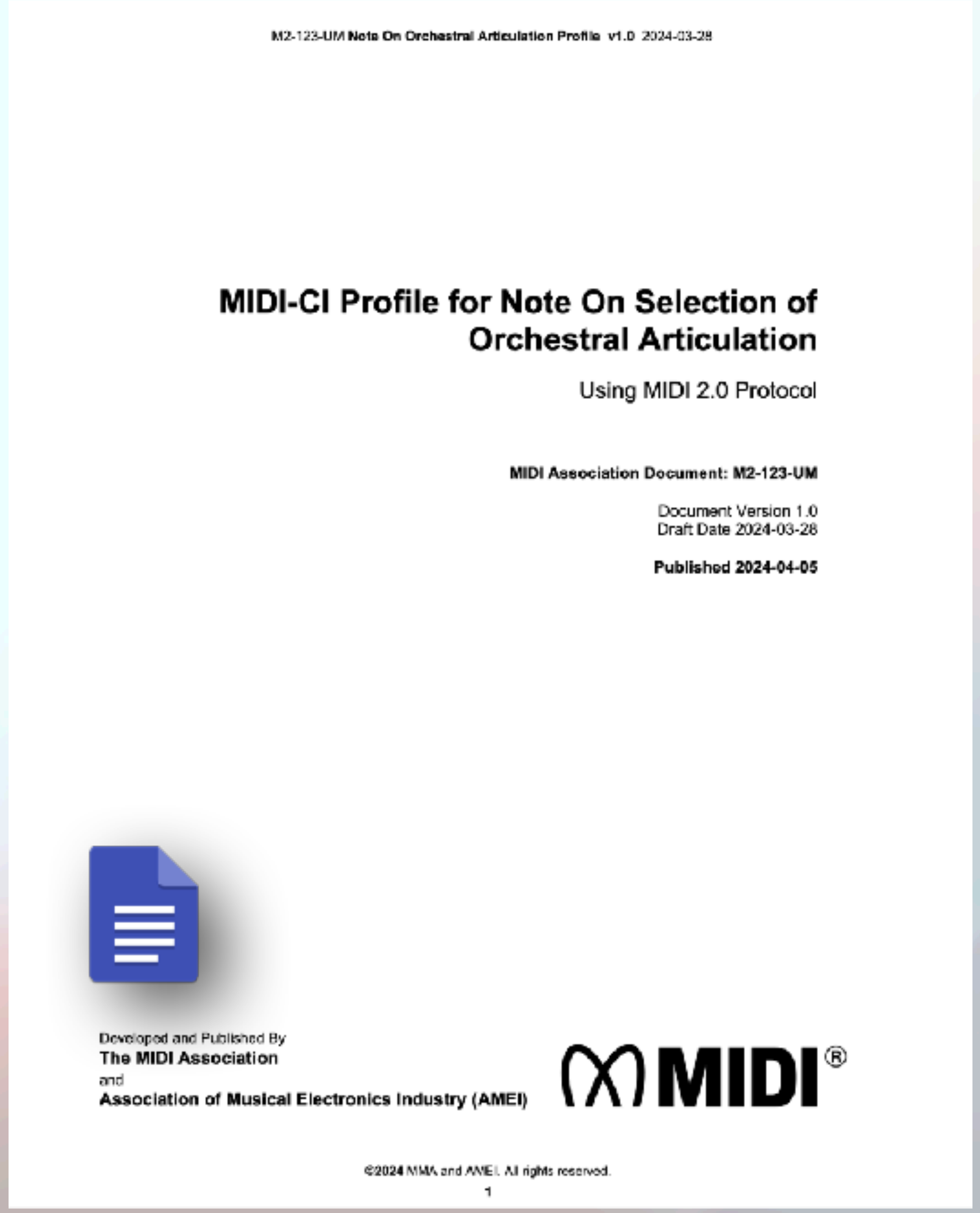
UMP - Universal MIDI Packet

- Uses existing MIDI 1.0 semantics and mechanisms
- New Universal MIDI Packet (UMP) which offers higher resolution performance controllers (32 bit), more controllers (32k) and 16 channel groups for 256 channels, better NRPNs and full 8 bit SysEx.
- Based on 32bit words. There are 32,64,96,128 bit UMPs



Profiles

- Profiles represent a semantic contract between senders and receivers.
- Piano
- MPE
- Drawbar Organ
- **Orchestral Articulation**
- …


M2-123-UM Note On Orchestral Articulation Profile v1.0 2024-03-28

MIDI-CI Profile for Note On Selection of Orchestral Articulation

Using MIDI 2.0 Protocol

MIDI Association Document: M2-123-UM

Document Version 1.0
Draft Date 2024-03-28

Published 2024-04-05

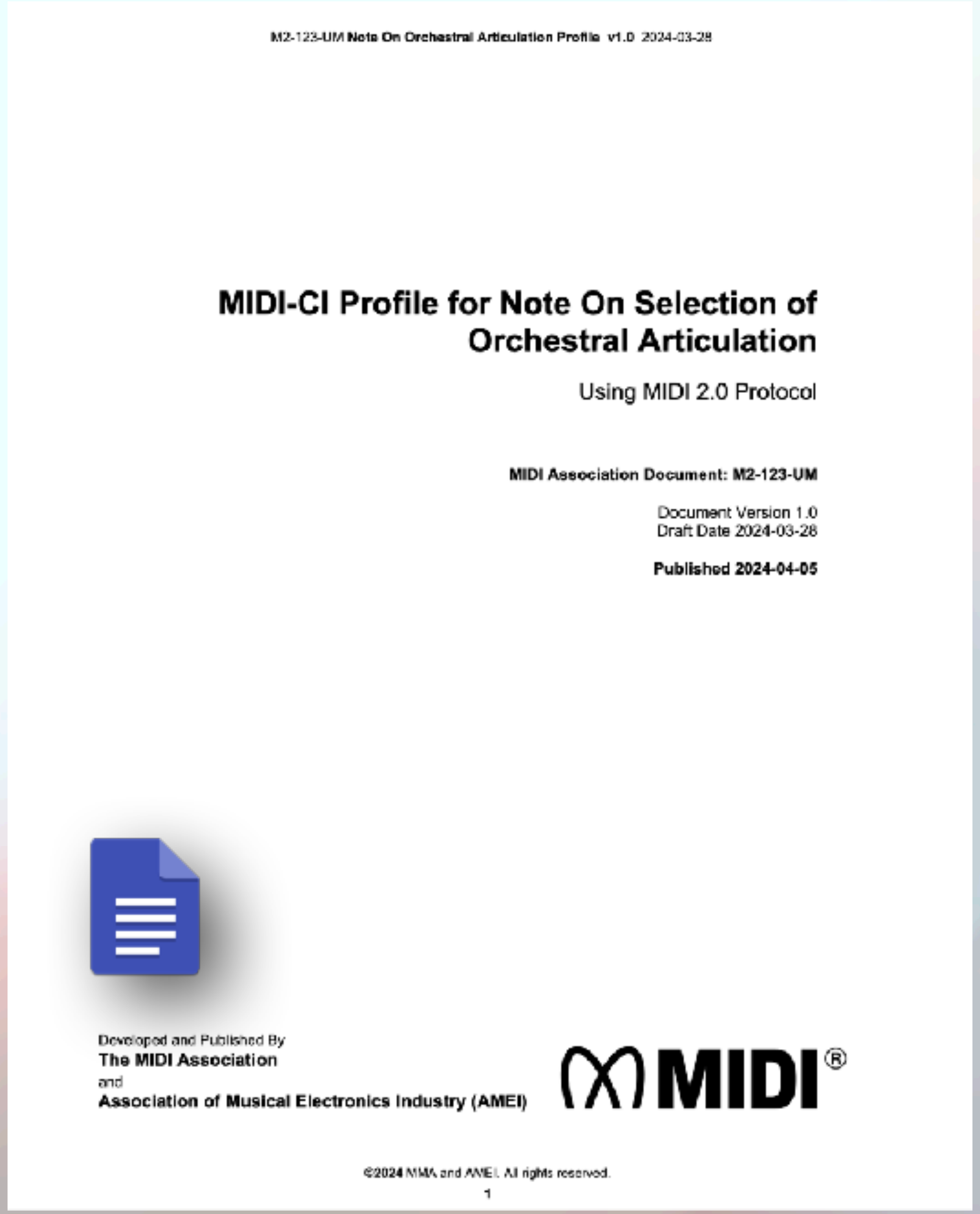
Developed and Published By
The MIDI Association
and
Association of Musical Electronics Industry (AMEI)

©2024 MMA and AMEI. All rights reserved.

1

MIDI 2 Orchestral Articulation Profile

- Released April 2024
- Provides a MIDI 2 standard for selecting articulations.
- Several hundred articulations across all families of instruments are defined
- Attribute tags in the UMP are defined for various articulations



M2-123-UM Note On Orchestral Articulation Profile v1.0 2024-03-28

MIDI-CI Profile for Note On Selection of Orchestral Articulation

Using MIDI 2.0 Protocol

MIDI Association Document: M2-123-UM

Document Version 1.0
Draft Date 2024-03-28

Published 2024-04-05

Developed and Published By
The MIDI Association
and
Association of Musical Electronics Industry (AMEI)

MIDI®

1

What Does the Orchestral Profile Mean for Live performance

- MPE controllers and Physical Models will most likely support the Orchestral Articulation Profile
- New classes of controllers to support direct manipulation of articulation parameters
- Many different sensors: touch, pressure, sliding, breath, spatial motion… all expressive opportunities

What Does the Orchestral Profile Mean for Scoring

- A standard way to articulate orchestral scores
- Much more comprehensive than current key switches. **Note PM can directly render Orchestral Articulations. NO key switches!**
- The profile supports **articulation equivalence** so that scores can be re-orchestrated with different instruments and the instruments respond with analogous articulations.

What Does the Orchestral Profile Mean for Virtual Performers

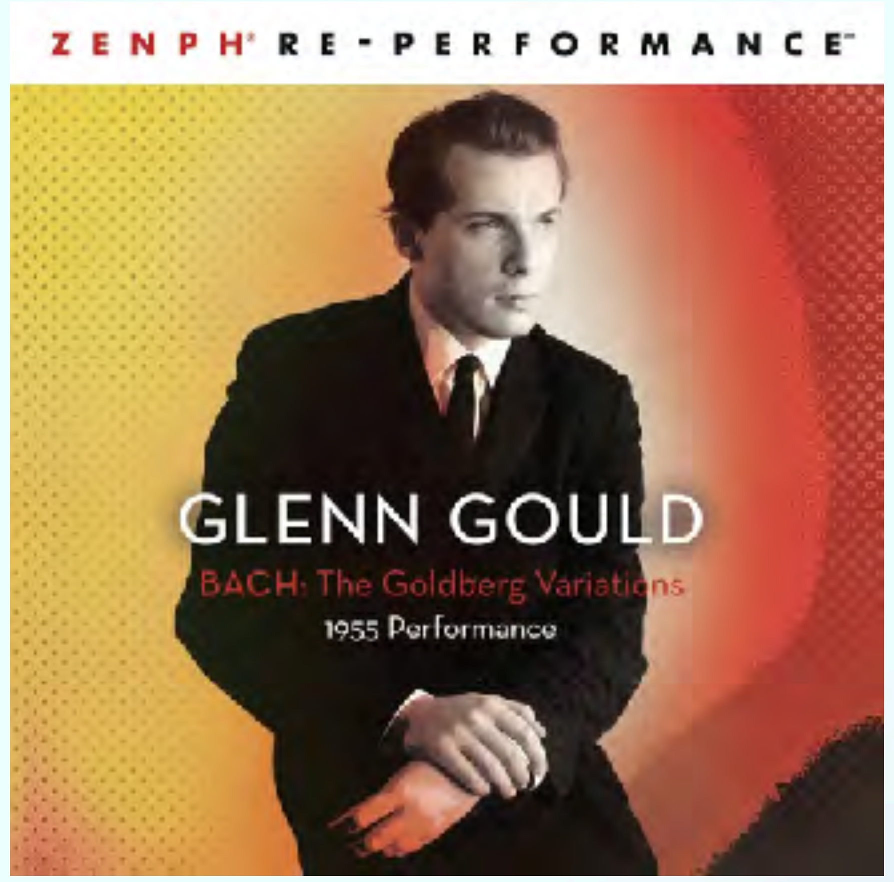
- Sometime in the moderately near future there will be a corpus of MIDI 2 articulated instrument performances.
- These performances may be used as training data for Virtual Performers.
- **Note that currently, there are unresolved ownership and ethical issues around creating AI tools using training data.**

Example Virtual Performer: JAM_BOT

- MIT Media Lab. Realtime Generative AI, trained on a collection of Jordan Rudess' MIDI performances.
- Performance with the JAM_BOT is conversational, with Jordan playing and then the JAM_BOT responding in the same style

Example, Zenph “Re-Performance”

- DSP and ML techniques to extract articulation data directly from recordings
- Velocity, timing, pedal estimate
- Stored as “10bit” MIDI
- Re-performed on a high resolution Disklavier
- Required intensive manual intervention. Zenph ceased operations in 2012.
- Now, over 10 years later these techniques could be revisited for monophonic instruments to create a corpus of articulated MIDI 2 performances based on real performances



Ethics Statement About Training Virtual Performers Using Orchestral Articulation Data

- The ownership of training data and resulting performances remains unclear
- There are issues of ownership, consent, attribution, labor displacement, authenticity.
- It’s probably inevitable that Virtual Peformers will emerge.
- The entire music eco system must start talking about these issues.
- In the IASIG there is an AIWG that is developing a crowd source wiki of best practices around music, voice over and game sounds.
- **Keep the human in the compositional and performance workflow**

Gratitude

Mary Albertson
Eric Bateman
Athan Billias
Simone Capitani
Chris Chafe
John Chowning
Perry Cook
Jon Dattorro
David Jaffe
Mike Kent
Joe Koepnick
Max Matthews (RIP)
Romain Michon
Denis Labrecque
Scott Levine
Roger Linn
Fernando Lopez-Lezcano
Keith McMillen (RIP)
Yann Orlarey
Stephane Letz
Stefano Lucato
Stanford OTL
Larry the O
Emanuele Parravicini
Danny Petkevich
Nick Porcaro
Bill Putnam
Jordan Rudess
Danielle Rudess
Kent Sandvik
Julius Smith
Tim Stilson
David Van Brink
Scott Van Duyne
Yamaha

And CCRMA

Questions?

You can reach me at
gps@moforte.com

Extra Slides for Specific FAQs

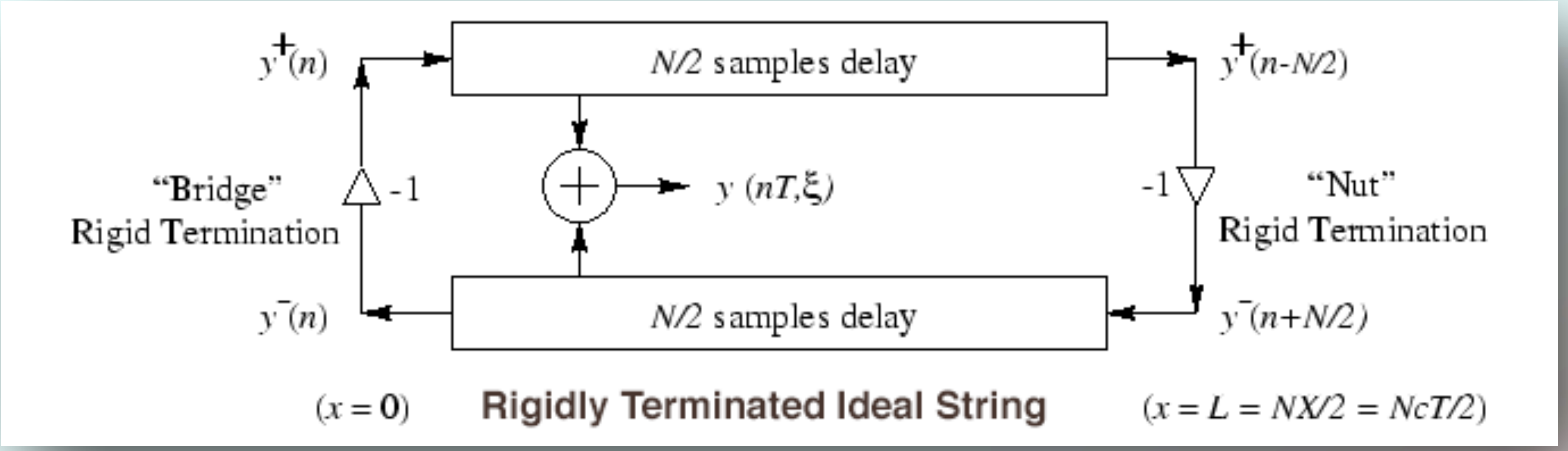
Why Android is a Challenging Platform for Audio Products Targeted for Musicians

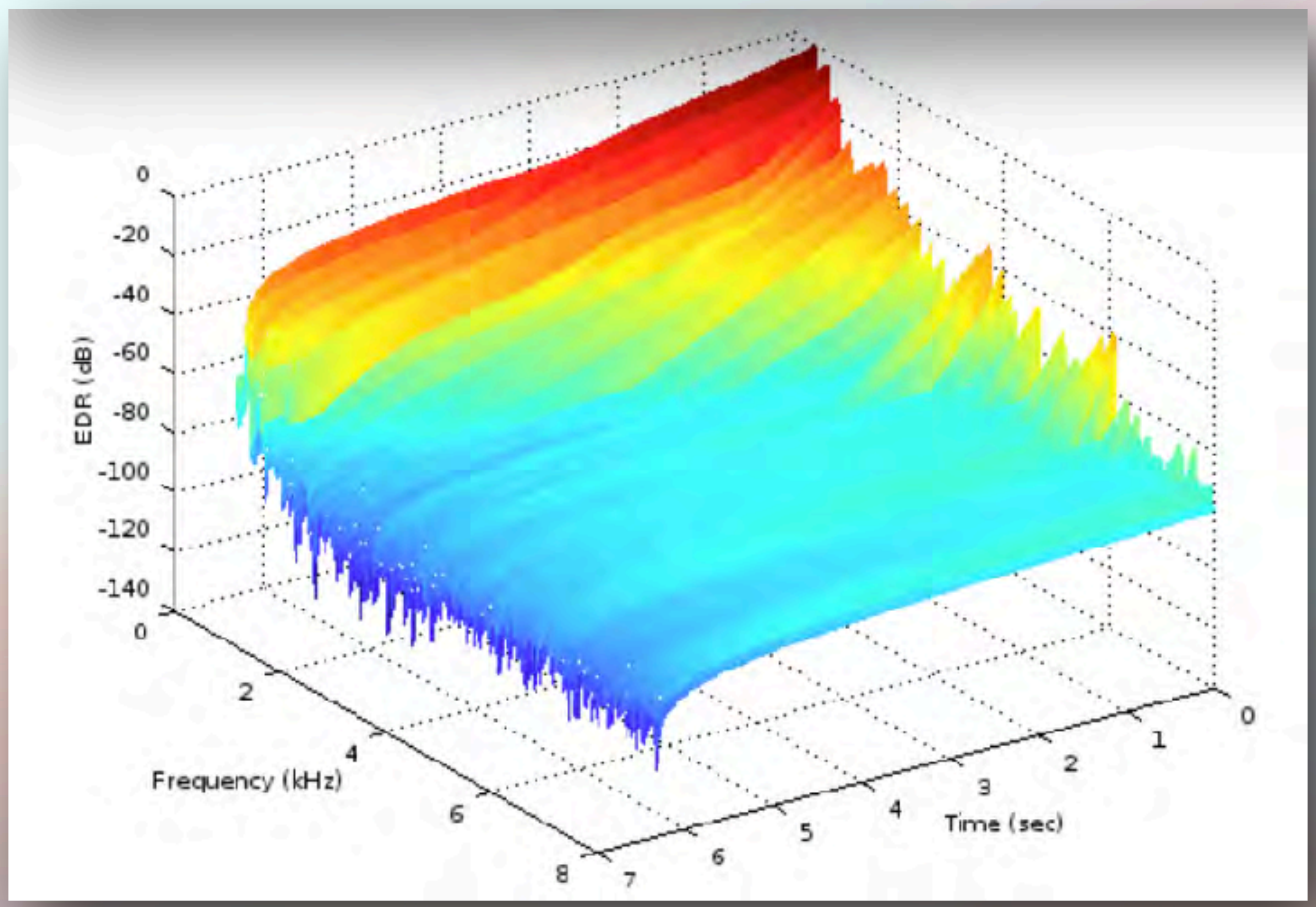
- Most Android devices have **audio** latency/jitter issues.
- Roli has measured a Mobile Audio Quality Index MAQi.
- With the exception of a few Android devices, only iOS devices have suitable audio latency/jitter properties

The Guitar Model

- A hybrid of Extended Karplus-Strong, Waveguide, Commuted Synthesis with extensions:
  - Harmonics and pinch harmonics
  - Pre-computed pickup excitations
  - Collisions for fret excitation
  - Sitar Bridge model
  - Body Model
  - Hexaphonic split
  - Doubling of courses
  - Statistical variations

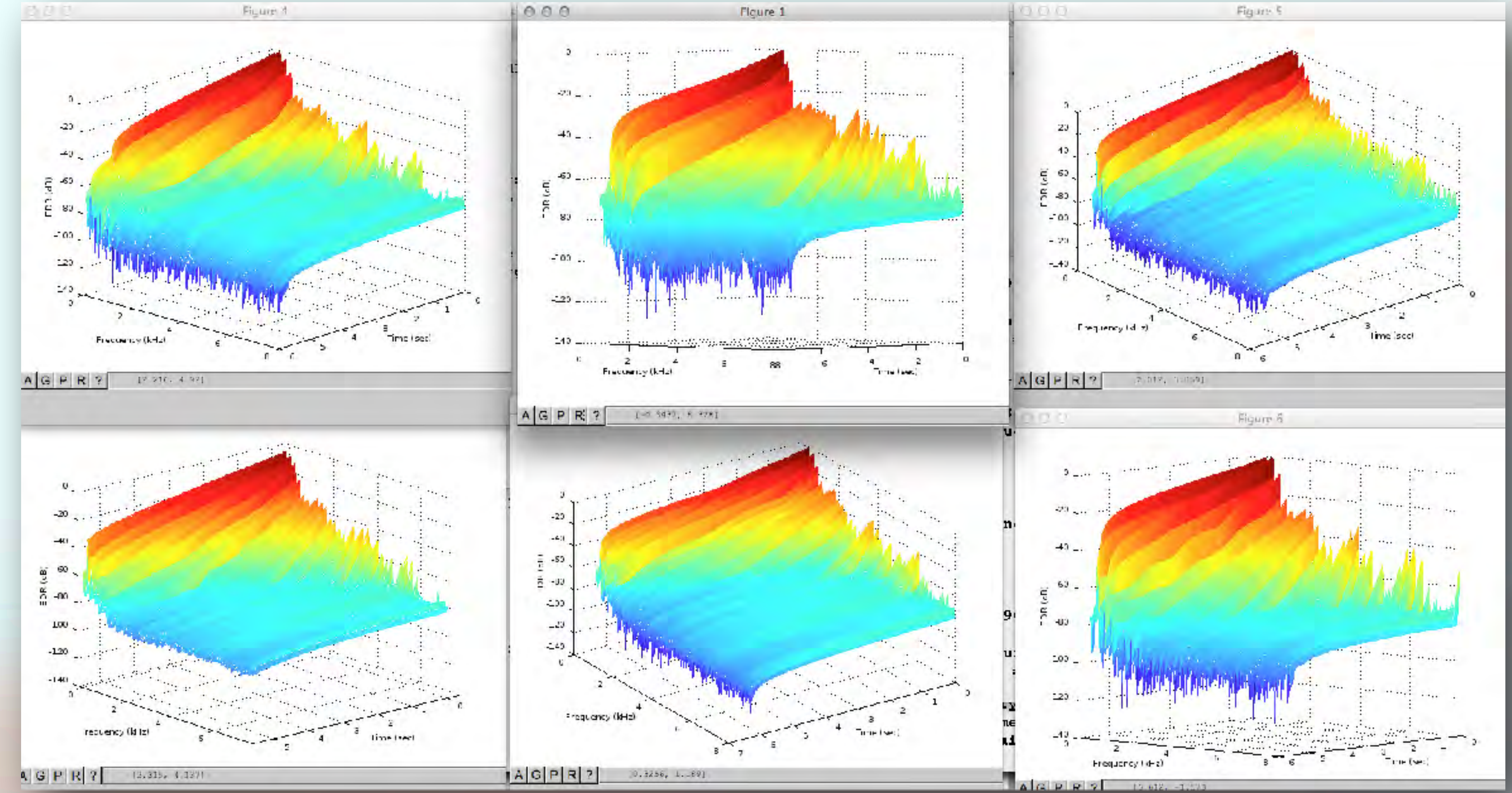
- Calibrated from real recordings





Model Calibration

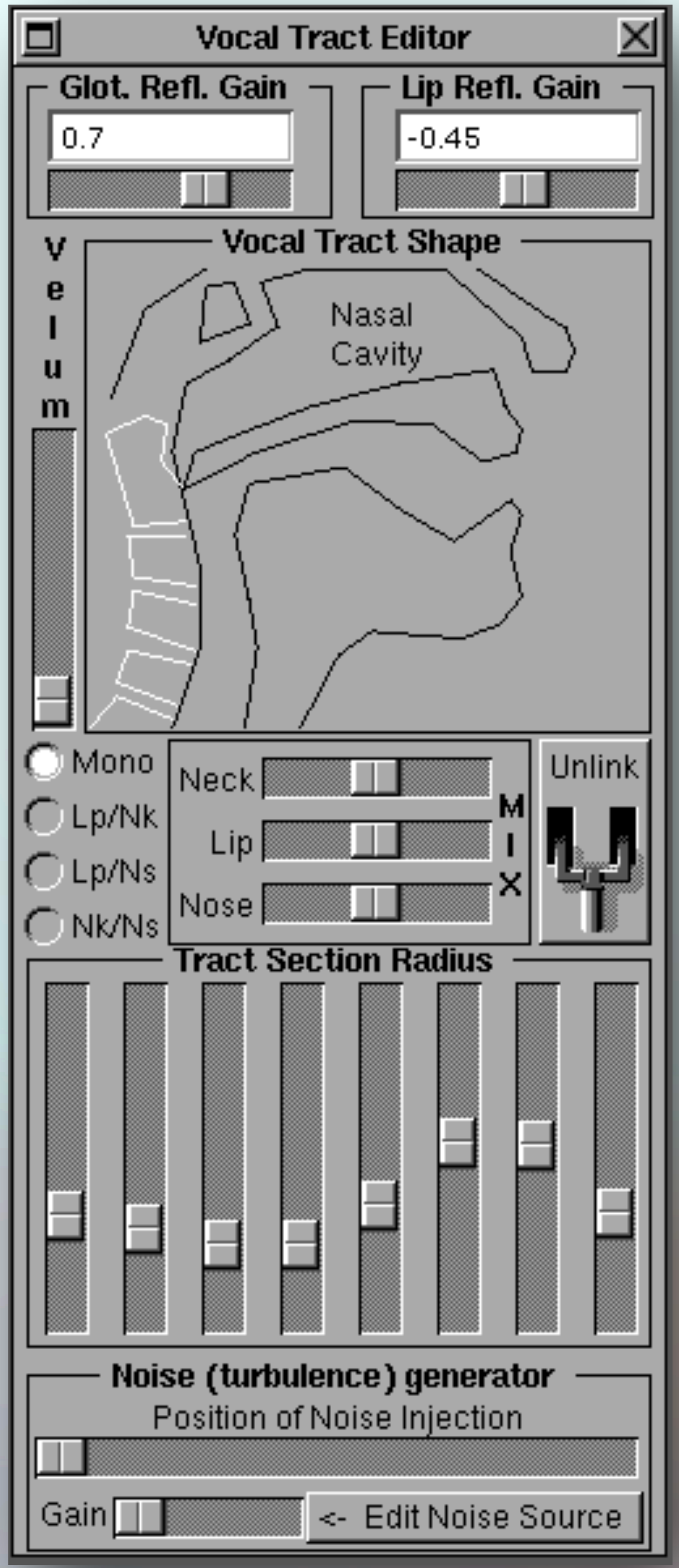
- Samples are collected for every string for every fret
- Analysis done in Matlab
- Goal is to design low order loop filters that match the partial decay rates in the original recordings.



The NeXT MusicKit (1988)

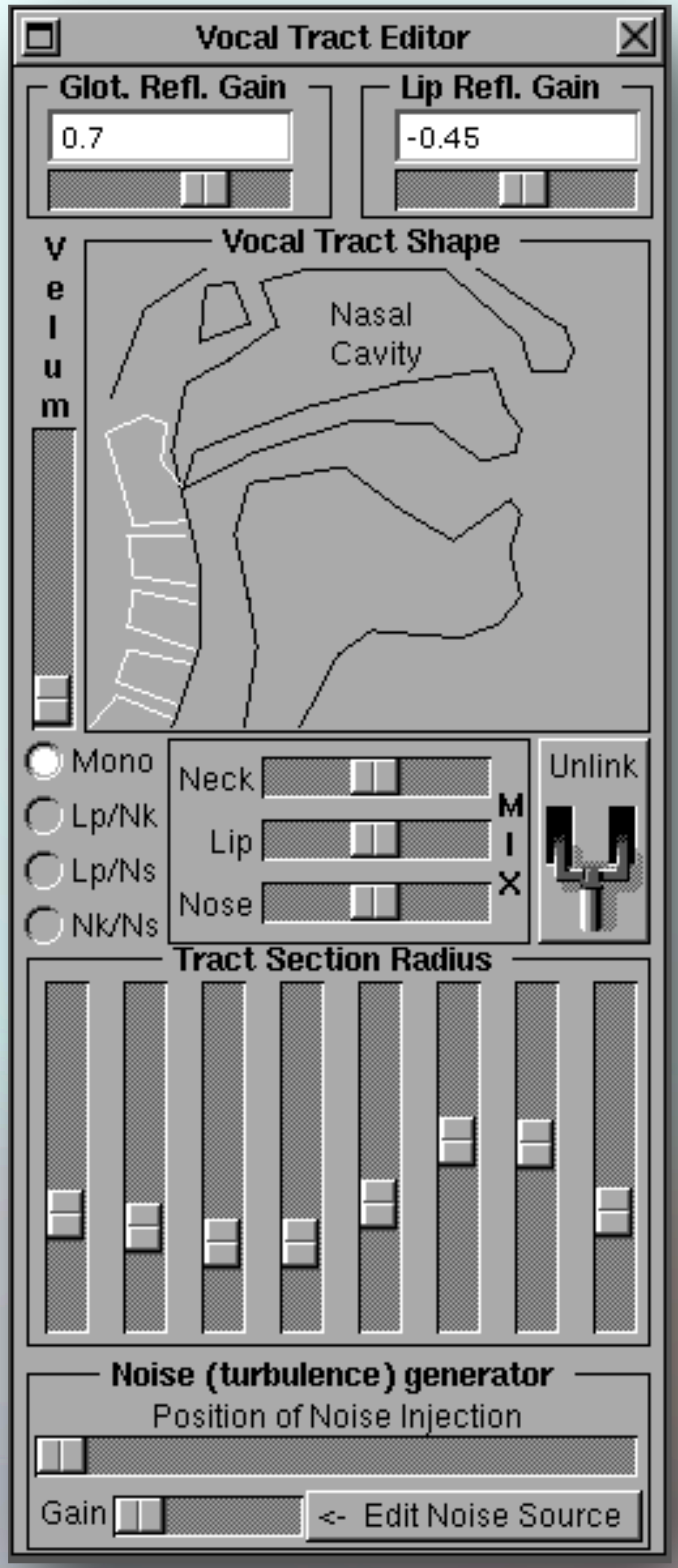
- The NeXT MusicKit unified MIDI (Control) and Music V (Unit Generators) Paradigms. (Jaffe, Smith, et al.)
- The launch of the NeXT Machine in 1988 included a performance of a 6 string physical model along with Dan Kolbialka playing Violin.
- In 1989 Mike Minnick created SynthEdit using the MusicKit and the NeXT Draw Program
- In 1992 CCRMA took over supporting the NeXT MusicKit.
- in 1993 Eric Jordan and David Jaffe created GraSP using the MusicKit and the NeXT Draw Program

Sheila Vocal Track Modeling (Cook 1990)



Perry Cook's SPASM **"Singing Physical Articulatory Synthesis Model"**

Sheila Vocal Track Modeling (Cook 1990)



Perry Cook's SPASM **"Singing Physical Articulatory Synthesis Model"**

What about Latency?

- The largest source of latency (for ios) appears to between screen interaction and the guitar model. Note that the audio buffer latency is about 5ms.
- We started at 180ms screen to audio out.
- We brought this down to 25-35ms by replacing Apple's gesture handlers with a custom gesture handler. This makes sense. Gesture handling requires analysis of a moderate amount of state to initiate an action.
- MIDI to Audio Latency is about 20-30ms.
- PowerStomp which is audio-in/effects chain/audio out is around 18ms.
- Latency to the internal speakers on iOS devices seems to have gotten a bit poorer over time. Probably due to DSP processing for the head phone jack.