A Custom Integrated Circuit Based Audio-to-CV and Audio-to-MIDI Solution

Brian Kaczynski

Second Sound, LLC, Miami, FL, USA

Presented at the 146th AES Convention in Dublin, Ireland on March 22, 2019

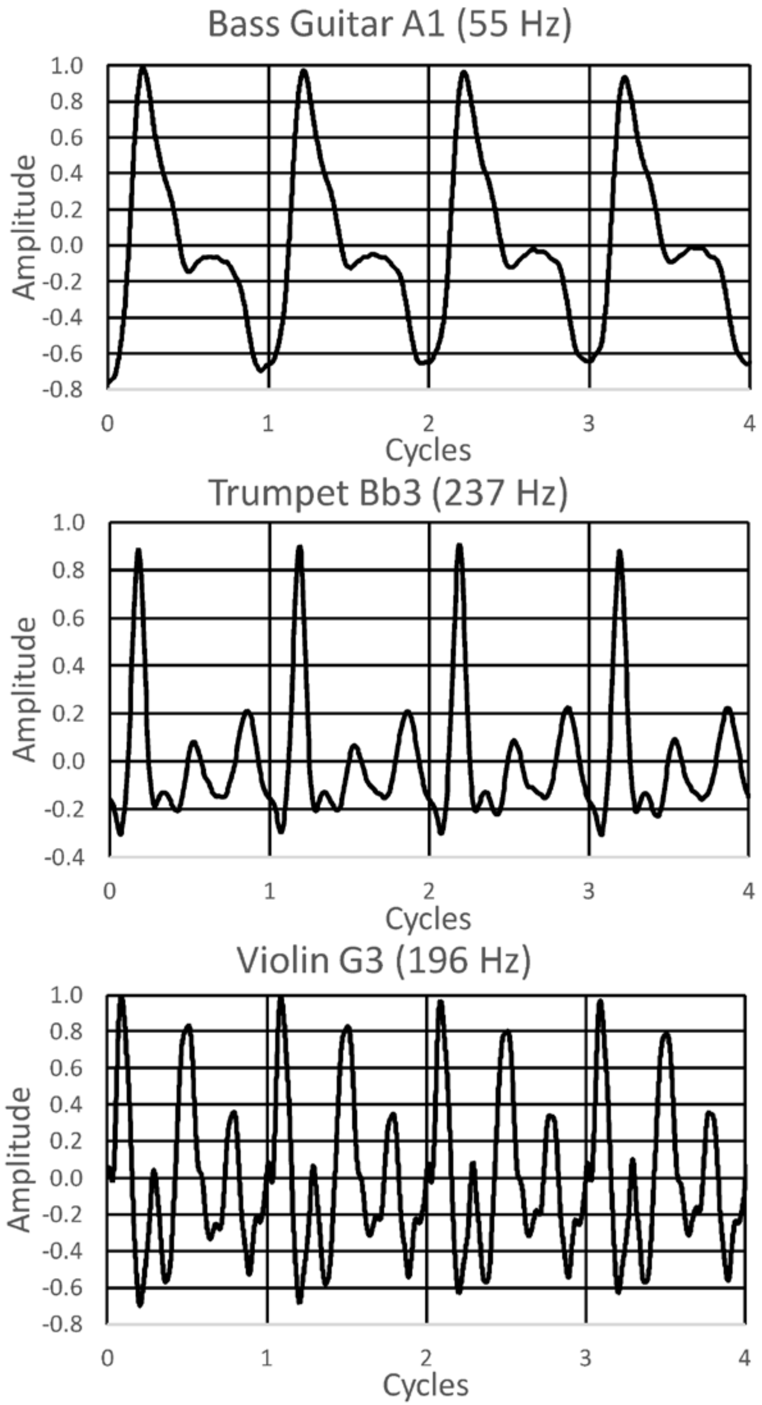
A Very Brief History of Pitch-Tracking Tech

- Pitch-tracking techniques originally motivated by early 20th C. speech processing applications like telephony – LOTS of literature on this!
- If we restrict ourselves to music applications, Korg (MS-20 in 1978) Roland (SPV-355 and GR-300 in 1980) and were the first.
- Otherwise innovative products, but their pitch tracking circuits were based on techniques like the Threshold Analysis Basic Extractor or TABE [1] (Korg) or fundamental frequency detectors with tunable filters that apply a closed-loop control system (Roland) from the mid-50s [2].

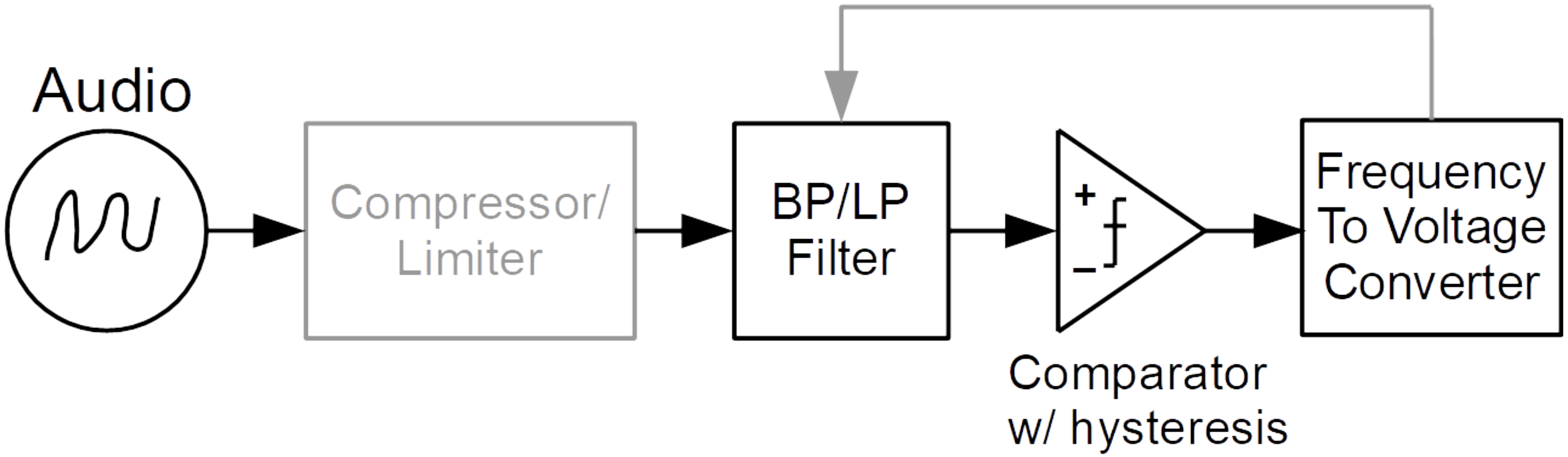
[1] McKinney, N.P. “Laryngeal frequency analysis for linguistic research” (1965).
[2] Edson, J.O. and Feldman, C.B.H. “Derivation of vocoder pitch signals,” *U.S. Patent No. 2,906,955* (filed 1956).

Difficulties of Pitch-Tracking

- For some instruments, no amount of filtering can produce clean zero-crossings at fundamental period!
- 5-string bass can play as low as 31Hz (low B0); transient response of filter tuned so low causes noticeable latency.
- Certain combinations of notes can cause tracking circuits to get “stuck” on a harmonic.

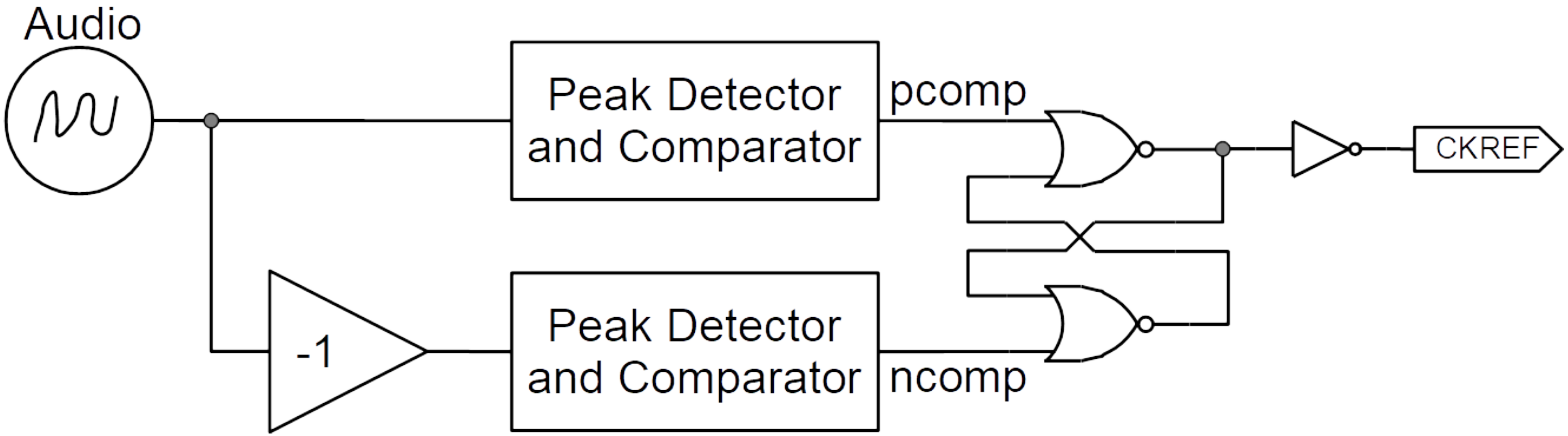


Commercial Pitch-Tracking Before 2000



- Compressor and feedback to BP/LP filter optional
- Roland SPV-355 adds envelope follower and servo loop to control BP filter cutoff frequency
- Detection similar to “filter + zero crossing” and subject to errors
- Tunable BPF adds latency especially for bass instruments

Peak-Detector-Based Pitch Tracking

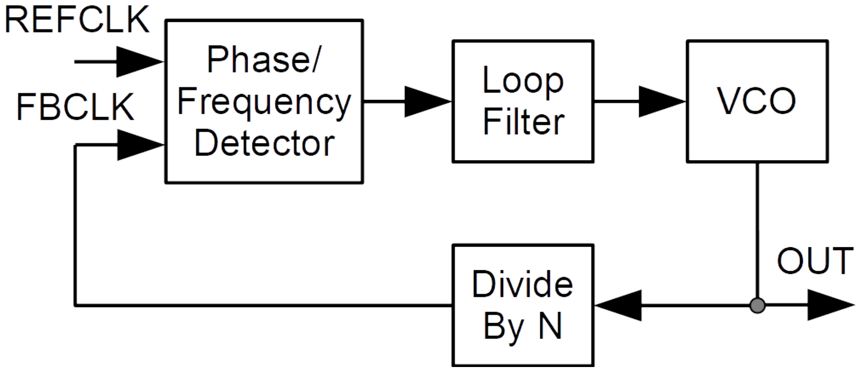


- Described by Angus, Garner and Howard in 1998 [3].
- Traditional peak detector with fixed decay time works only over narrow range of frequencies, but used effectively in EBS OctaBass (2002).

[3] Angus, J., Garner, P. and Howard, D. "Fundamental frequency estimation for use with a singing pitching development system for primary school children," *Proceedings of the Institute of Acoustics* (1998).

Peak Detectors with Sliding Time Constant

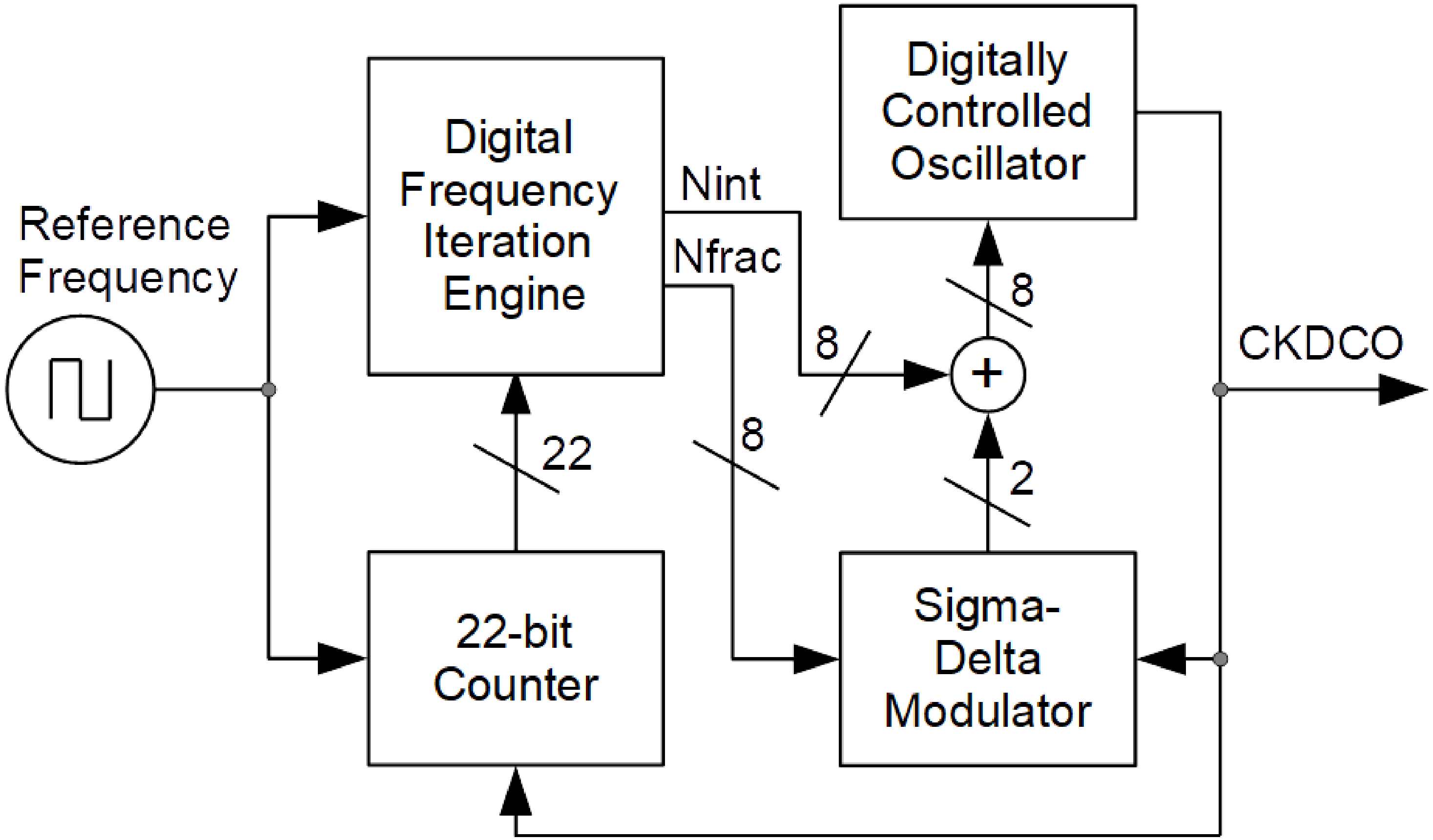
- Switched-capacitor circuits can give us time constants which scale with sample rate, but we need a frequency multiplier for sampling.
- Traditional PLL has poor transient response if we want wide tuning range [4].



[4] Gardner, F.M., “Charge-Pump Phase-lock Loops,” *IEEE Transactions on Communications* (1980).

Digitally-Controlled FLL [5]

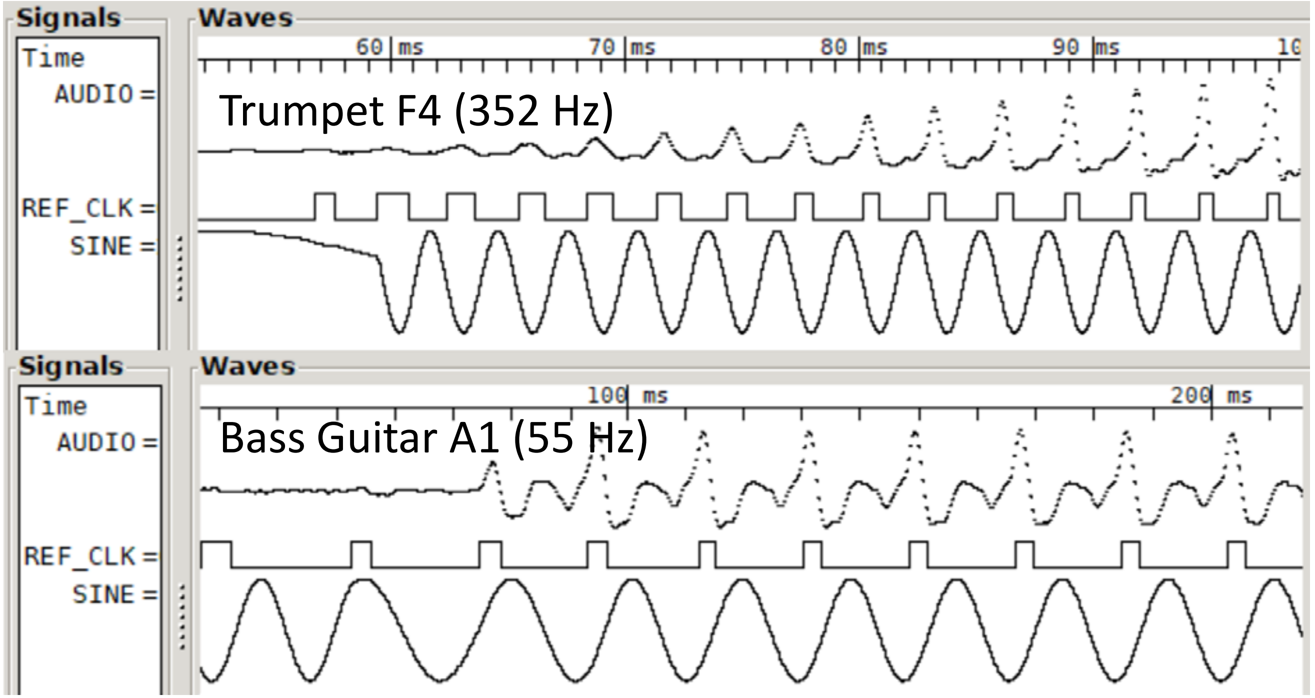
- FLL locks in one cycle
- Tracks from 25Hz to >5kHz
- F_{DCO}/F_{Ref} = 8,192
- 16-bit (integer + fractional) frequency word used for pitch CV and MIDI commands



[5] Kaczynski, B., "Fast-locking frequency synthesizer," *U.S. Patent No. 9,685,964* (filed 2015).

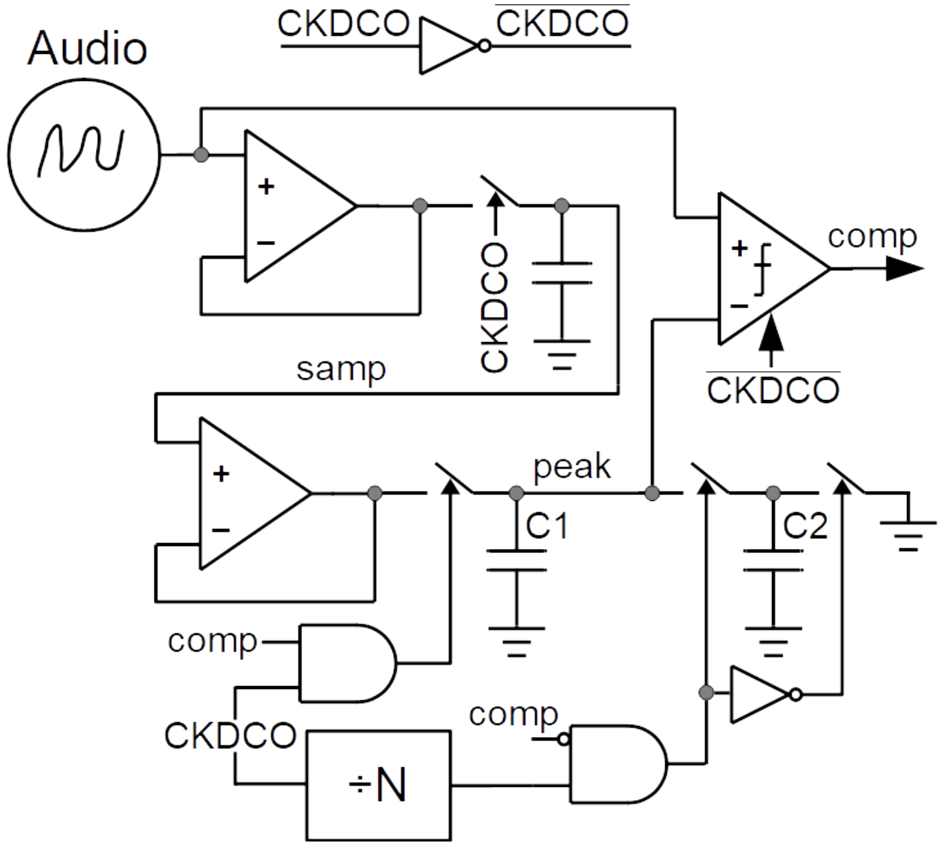
FLL Simulation Waveforms

- System simulated with Icarus Verilog
- Event-driven simulator mandatory due to variable sample rate
- Peak detector and FLL tightly interrelated…



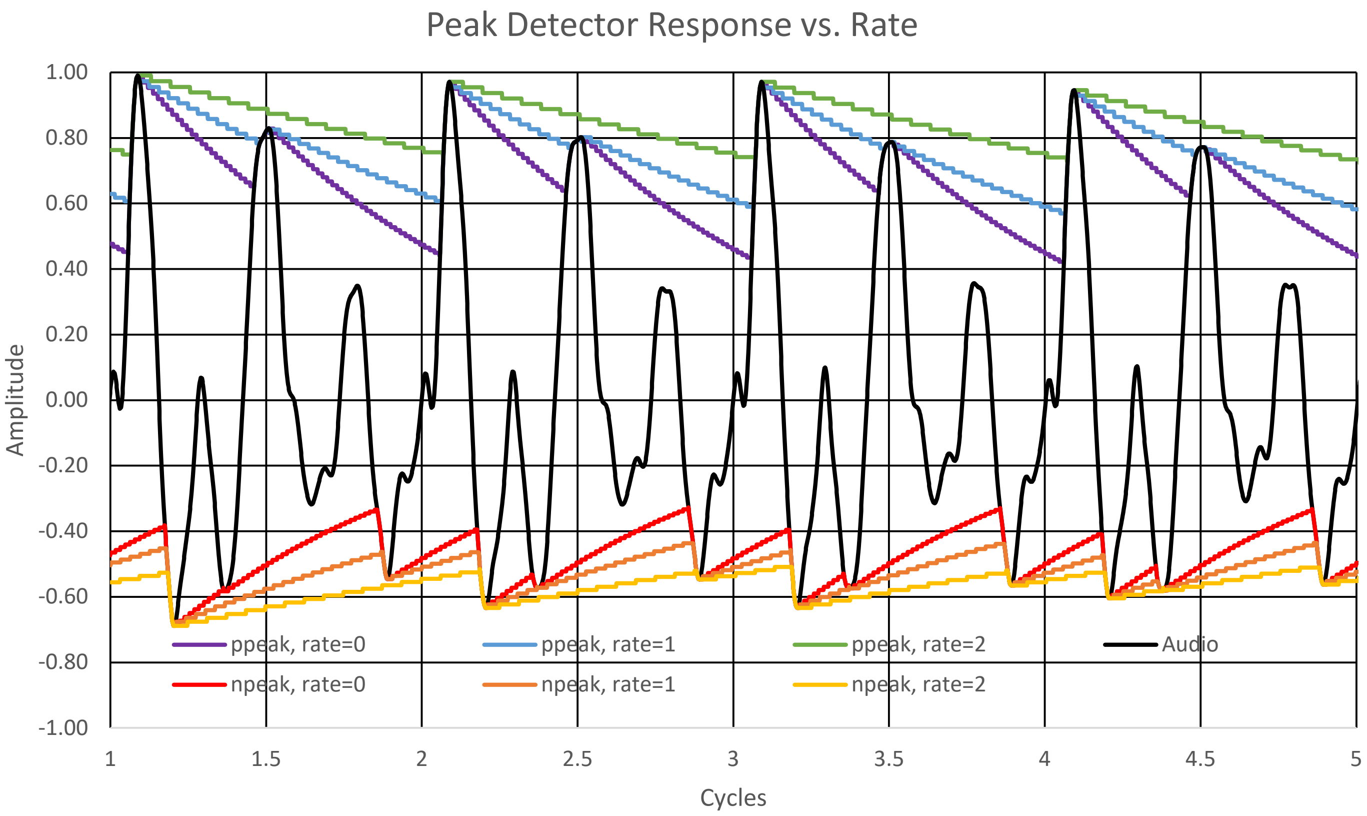
Switched-Capacitor Peak Detector

- First sample-and-hold (S/H) amp always sampling input
- Input sample transferred to “peak” via second S/H if input is greater than current peak
- Peak decays by periodically shorting discharged C2 across C1 every N DCO cycle, provided comp is low
- Decay/cycle = $\left(\frac{C1}{C1+C2}\right)^{8192/N}$



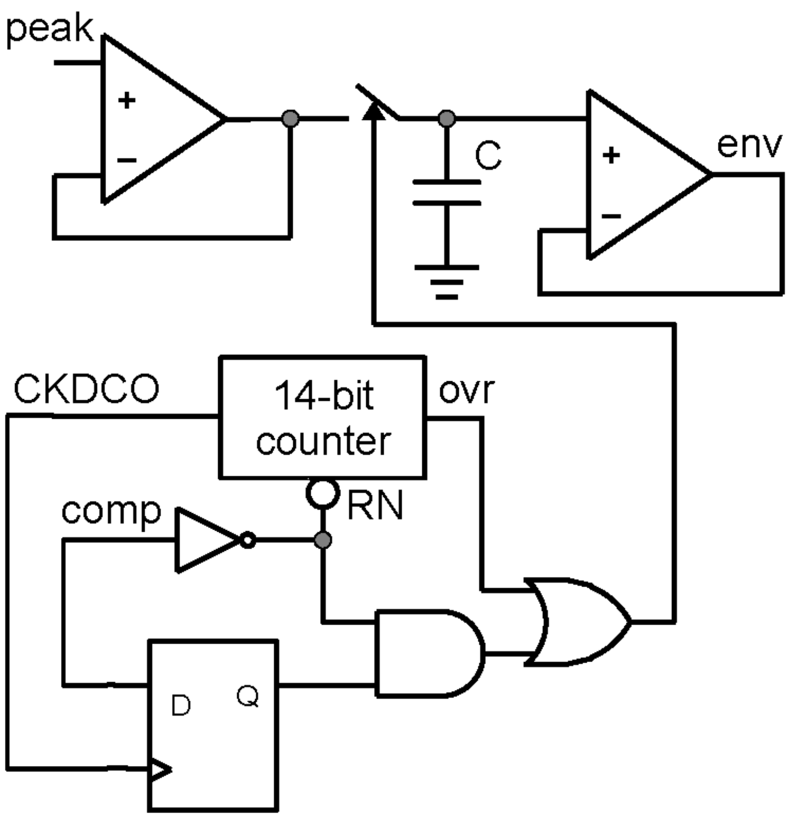
Peak Detector Decay Time Tuning

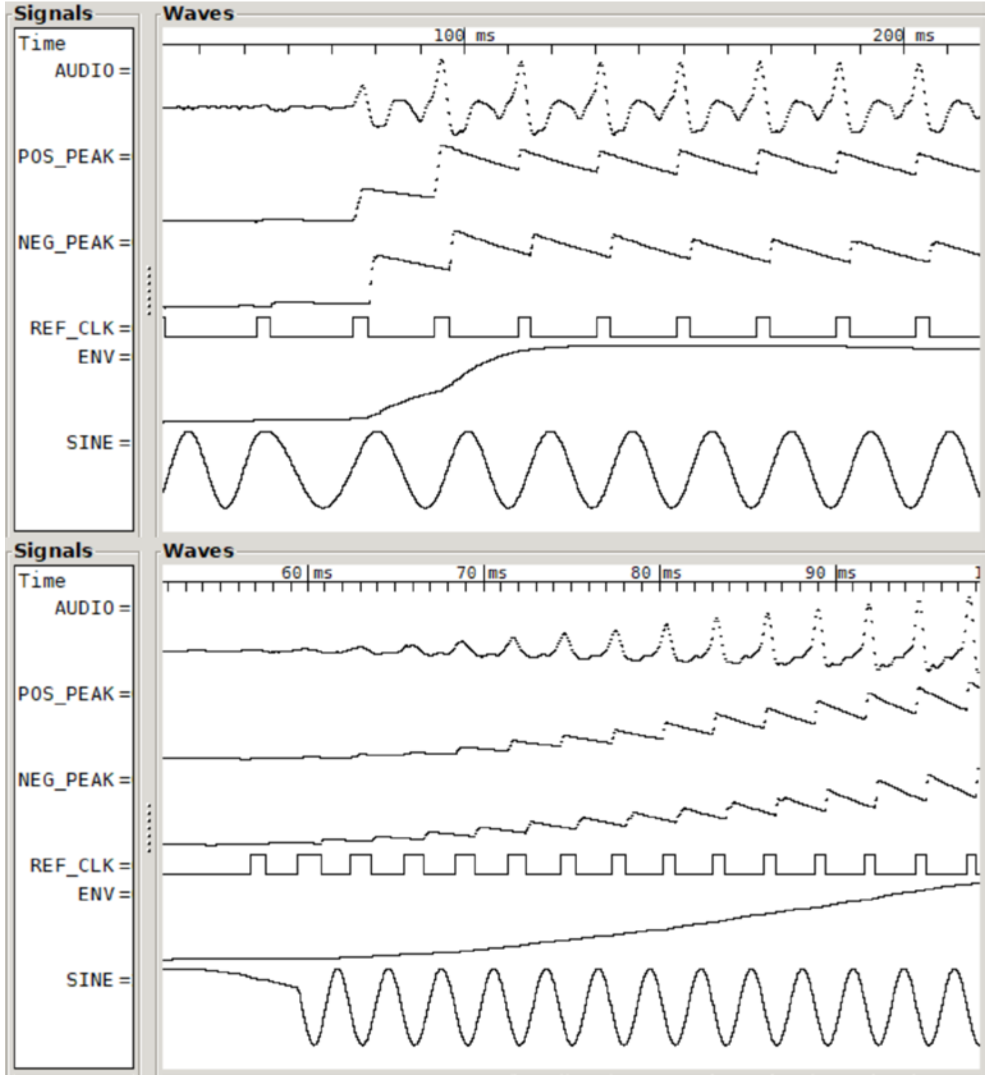
- Count N tunable to 128, 256, 512, 1024
- Longer decay (1024) rejects more harmonic energy but can skip cycles when note decays
- Shorter decay good for fast attack/decay instruments like bass



Zero-Ripple Envelope Follower

- "peak" sampled to "env" on comp negative edge OR counter overflow





Audio to MIDI Performance

- Adaptive polarity selection chooses negative peaks
- 16-bit frequency sent to uC via 3-wire SPI interface clocked at 1,024x audio
- MIDI note on takes 640us with running status
- Optional pitch bends sent each cycle
- Note off when gate falls uses velocity 0x00

Conclusions

- Real-time hardware audio to CV/MIDI opens up new worlds of creative possibilities for musicians, sound designers, composers…
- Breathes new life into vintage analog synths which were not so easily “playable” before and adds a human element to digital synths that some might argue was lacking.
- Applications not discussed: automatic scoring, toys, speech processing (back to where it all started?)…
- High-level integration means Audio-to-MIDI could be more ubiquitous in devices: Sound cards? Mobile phones? What else can you imagine?