Back To The Future

Emulating Classic Audio with 21st Century Technology

Emulating Classic Audio With 21st Century Technology

Why do we care about “old stuff?”

Historical Reference

- Certain sounds and instrumental techniques have become “gold standards.” E.g. Stradivarius violins.
- Focus groups repeatedly chose familiar sounds over unfamiliar ones in preference tests. (They also prefer familiar material.)
- Student musicians and audio engineers learn their craft from teachers who developed their skills on classic instruments and equipment.

Classic Instruments & Equipment

- The most enduring instruments designed in the 20th century are electro-mechanical. Instruments that generate sound purely by electronic means fall out of fashion rapidly.
- Most electro-mechanical instruments have quirks (defects?) that performers have exploited for their unique expressive capabilities.
- Legacy studio equipment often had unique distortion characteristics that added “warmth” or “color” to the sound.
- Certain purely electronic instruments are still regarded as iconic, but rarely used in contemporary settings.

Classic Gear

- Why not just get the old stuff?
  - Expensive
  - Rare
  - In poor condition
  - Spare parts & consumables unavailable
  - Big and heavy
  - Requires specialized maintenance and storage

State of the Music Business

The good news is that streaming is growing at a wonderful pace.

And that's about it for the good news.

– Donald Passman, *All You Need To Know About The Music Business*. (9th edition.)

Music Business Summary (2015)

- Industry earnings are only 50% of what they were in 2003.
- CD sales represent less than 25% of the business, and falling. Download sales have declined at about 15% per year for the past several years.
- Musical instrument manufacturers face challenges from financial (e.g. Steinway bankruptcy) to shortages of skilled labor for assembly. Most “classic” instruments and equipment are from defunct companies!

Music Technology Summary 2017

- Equipment becomes more affordable everyday. Educational opportunities in audio engineering abound using contemporary gear.
- Advances in audio processing have diminished the need for expensive treated acoustic studio spaces. (E.g convolution reverbs.)
- Software emulations of legacy instruments/equipment are widely available at various price/quality levels.
- Session musicians can work remotely and transfer recordings via ftp or cloud services. The talent pool is worldwide.

Emulation Approaches

- Subtle complexities of electro-mechanical instruments are not easily emulated via synthetic sound generation alone.
- The most common approaches are digital sampling, physical modeling, or a combination of both.
- Synthesis techniques like FM and waveform filtering do not adequately capture the nuances of physical instruments.
- Approaches to emulating purely electronic studio equipment vary from TSAR (“That Sounds About Right”) to modeling circuits in SPICE or other tools.

Evaluating Emulations

So how good is this stuff?

Audible Perception

- Perception is highly subjective and cannot be easily described, let alone quantified.
- Historically, listeners claimed “perfect” fidelity on media we could consider inferior, e.g. Thomas Edison’s wax cylinders.
- Most recorded pop music cannot be compared to a live performance since studio productions cannot be performed live.
- Recorded and live music are really two different art forms.

Artistic Perception

- Physical instruments often allow performance techniques that are difficult or impossible on emulated interfaces.
- Performers generally feel more comfortable with physical instruments and often yield a better performance on them.
- Even accomplished musicians generally find it *impossible to recognize real instruments vs emulations* on recordings.

Stages of Music Production

- Composing/Arranging/Organizing
- Recording
  - Tracking
  - Overdubbing
- Mixing
- Mastering
- Production & Distribution

Composing & Scoring

If you don't have it in your part, leave it out because there is enough missing already.

– Eugene Ormandy, Philadelphia orchestra conductor

Composing & Scoring

- Some sort of musical "roadmap" is needed for musicians to perform together. Otherwise chaos reigns (and costs increase!)
- Many musicians use tablets to display music and "turn pages" using footswitches connected via Bluetooth or USB.
- Although music notation hasn't changed significantly in over 200 years, software notation programs struggle with many details.
- Areas of active R&D in music notation:
  - Scanning (i.e. "OCR"). Current products exhibit poor performance.
  - Automated transcription of music notation from audio.

Recording - Instruments

- Microphones (for acoustic instruments and vocals)
- Drums & Percussion
- Guitars
- Keyboard instruments
- Strings/Brass/Woodwinds

Microphone Technologies

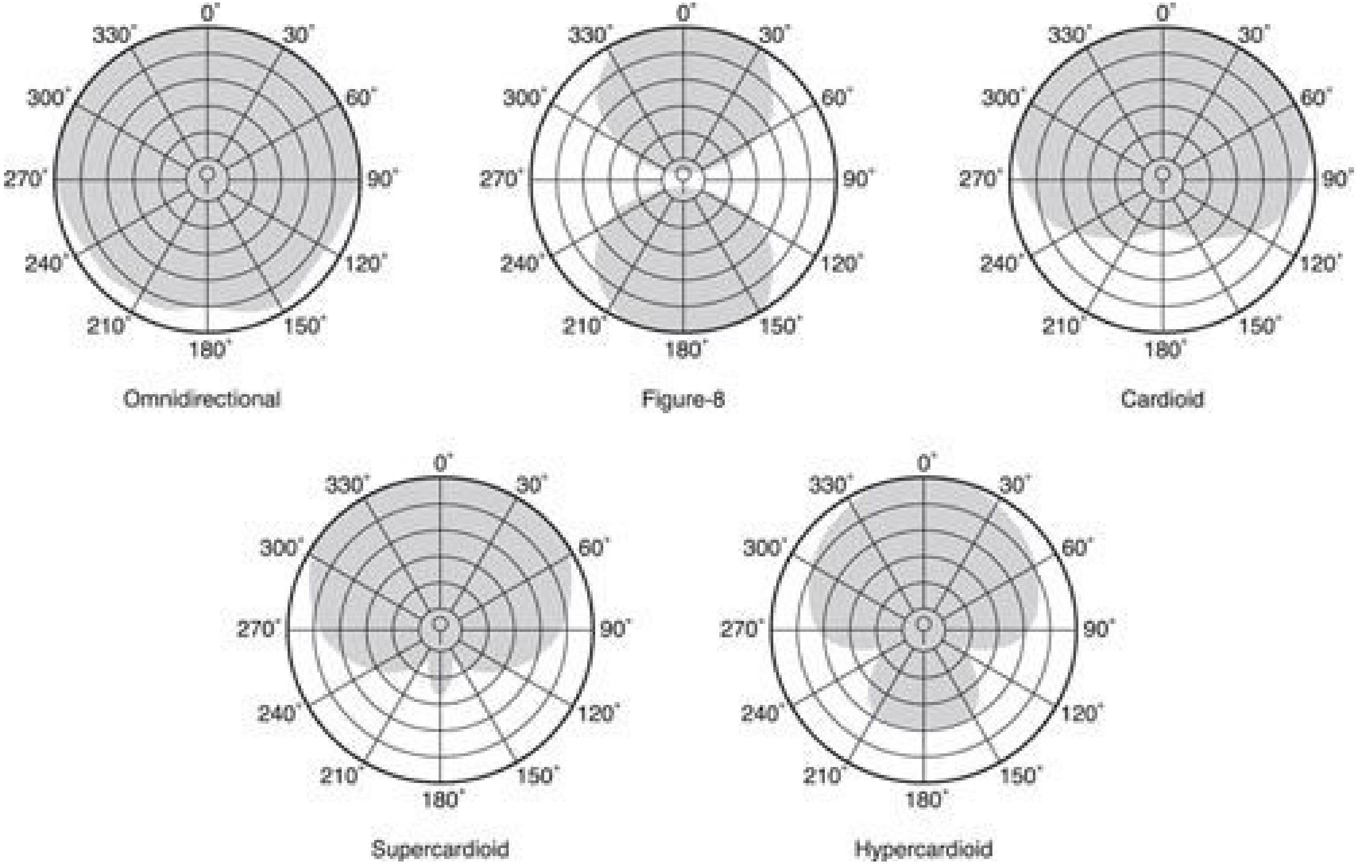
- Ribbon
- Dynamic
- Condenser

- Most studios will have several of each type.

Microphone Characteristics

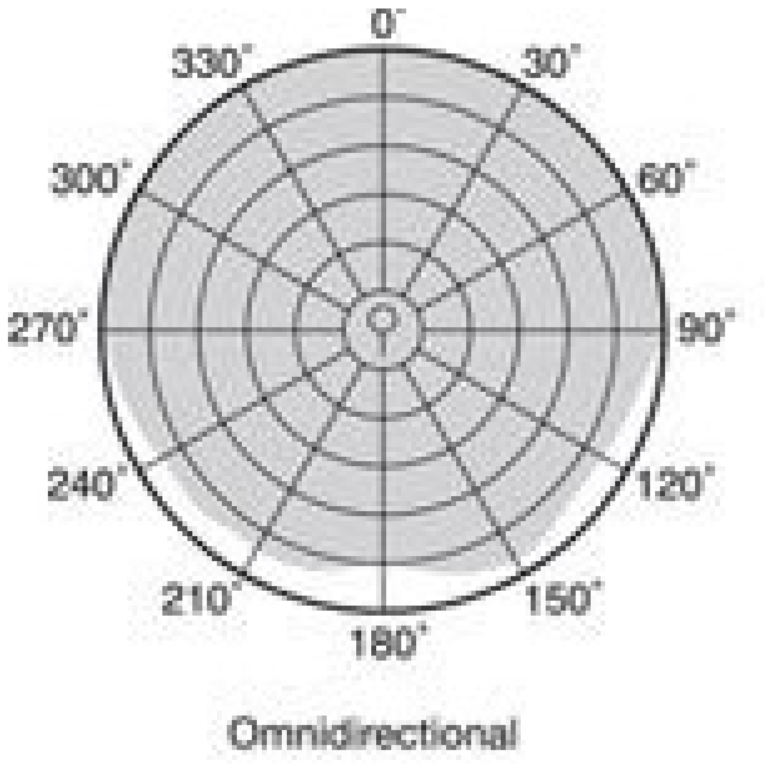
	Ribbon	Dynamic	Condensor
Pattern	Cardioid	Figure 8	Any
Frequency Response	Fairly flat, lacking high end	Uneven, lacking high end	Extended range, variable response
Cost	Medium - High	Low - Medium	Low- High
Active?	Some	No	Yes
Output Impedance	Very high	Low to High	Low
Output Level	Very Low	Medium to High	Low to Medium
Safe SPL	Very Low	High	Moderate
Suitability for Mods	High	Low	Varies
Ruggedness	Very Low	Very High	Low

Microphone Pickup Patterns



Omnidirectional Pattern

- Note typical pattern is never completely omnidirectional due to interference of mechanical mounting systems.
- Requires quality acoustic space, which is rarely the case in low-budget studios.



Cardioid Pattern

- Attempt to make pattern more directional.
- Note good rejection from rear, but relatively little rejection to sides (< 3dB).
- Side effect: bass increases as source distance decreases due to "proximity effect."

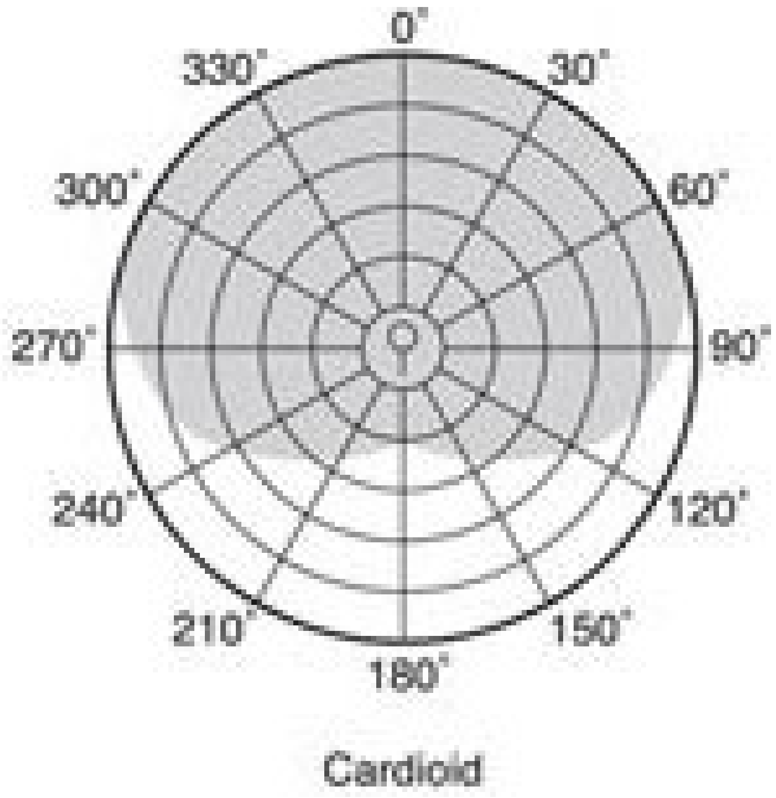


Figure 8 Pattern

- Note excellent side rejection, but almost symmetrical front/rear response.
- All ribbon mics inherently have figure 8 pickup patterns.

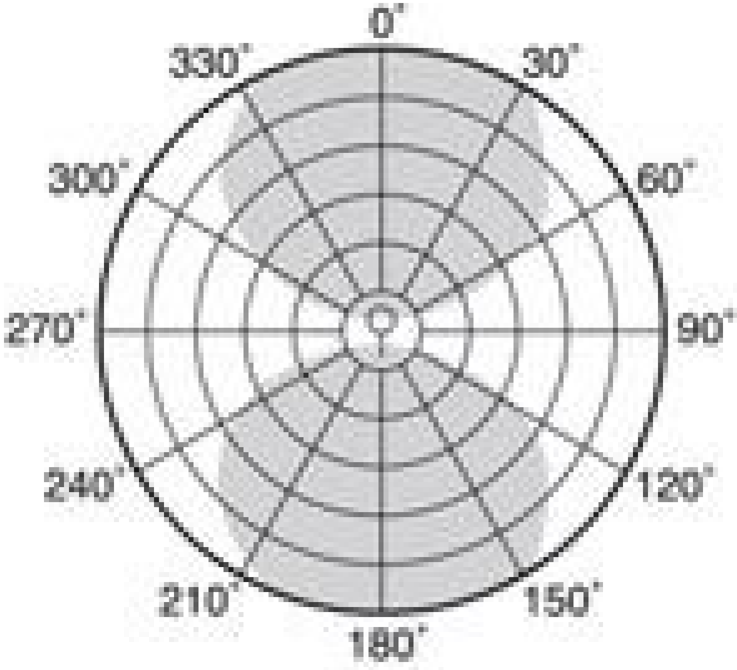
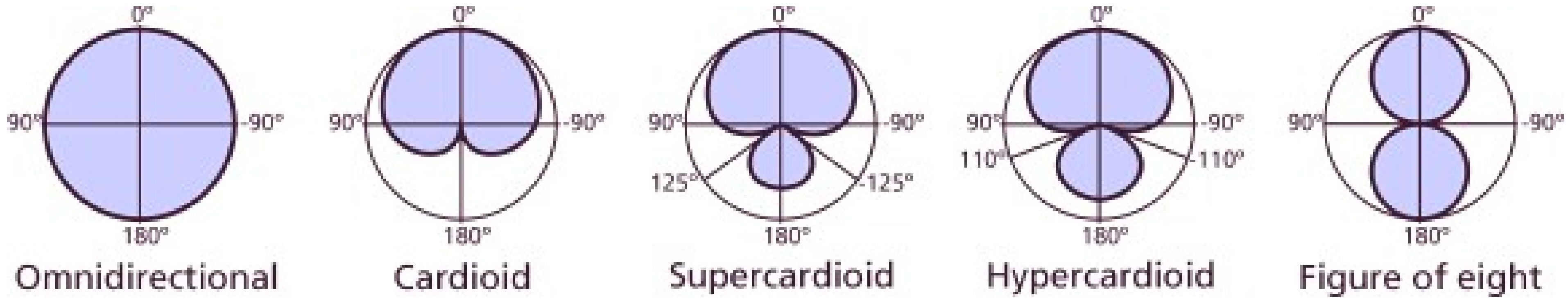


Figure-8

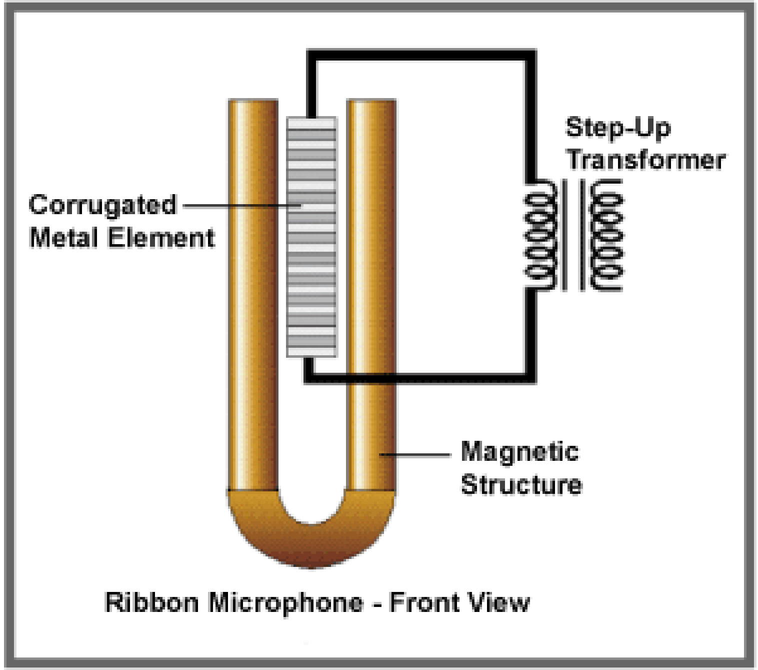
Intermediate Patterns

- As we reduce side pickup, we increase rear pickup.
- So a Cardioid eventually morphs into a Figure 8.

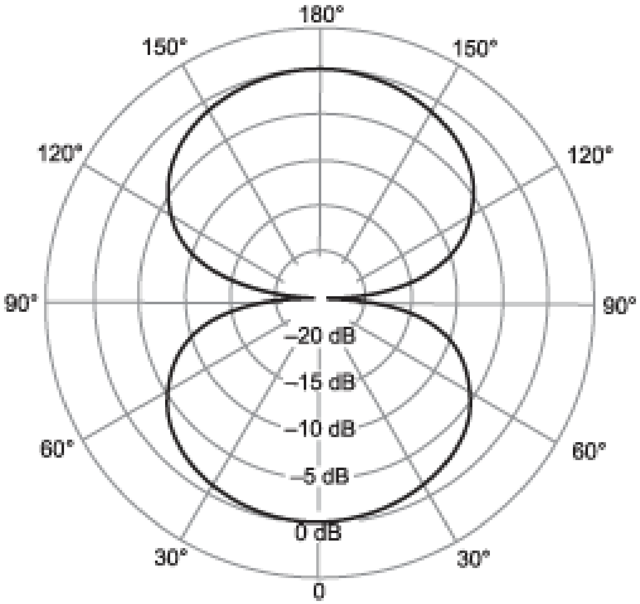


Ribbon Microphones

Ribbon mics are *very* simple:



How a ribbon microphone works



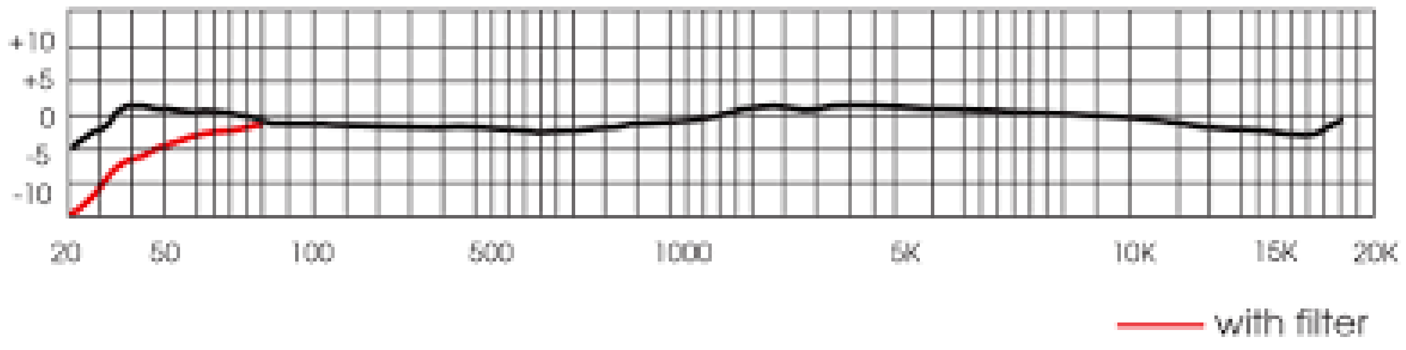
Bi-directional polar or pickup pattern

RCA 44B & 44BX

- One of the earliest ribbon mics commonly used by the broadcast and recording industries.
- Produced 1932 – 1955.
- Average current street price $2,500.
- Fragile and high-maintenance.

Royer R-122

- Contemporary (active) ribbon mic.
- Cost ~$1,900.
- Good frequency response.
- Somewhat less fragile than classic ribbon mics.

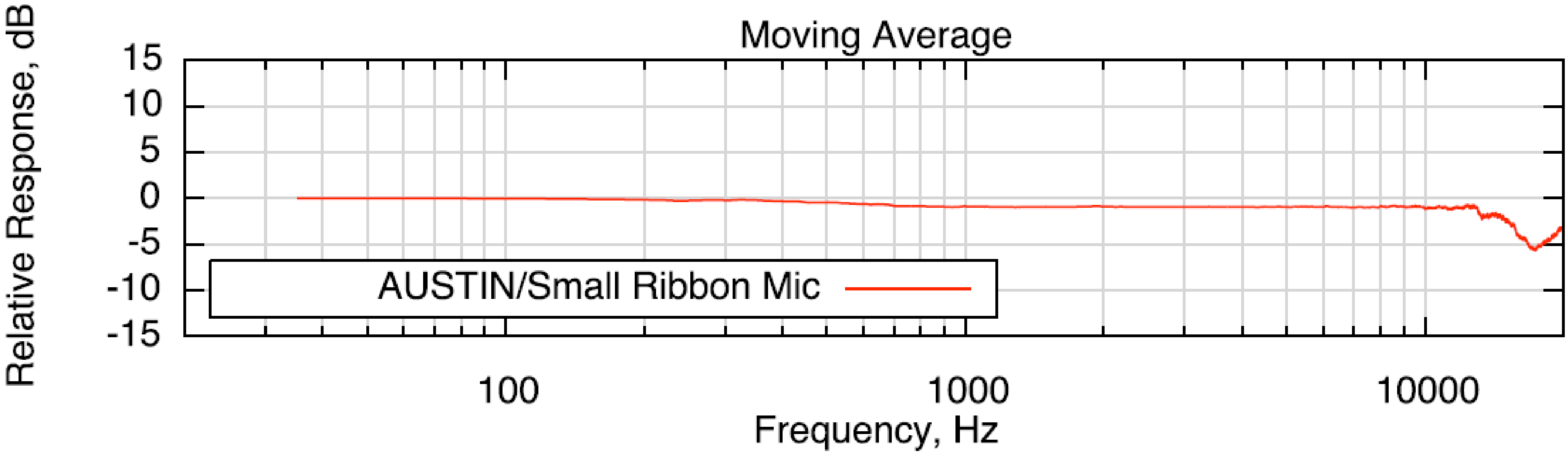


Budget DIY Ribbon Microphone

- Austin DIY Kit $250 - $380 (depending on transformer).
- Ribbons can be repaired ($330) and transformers can be replaced.

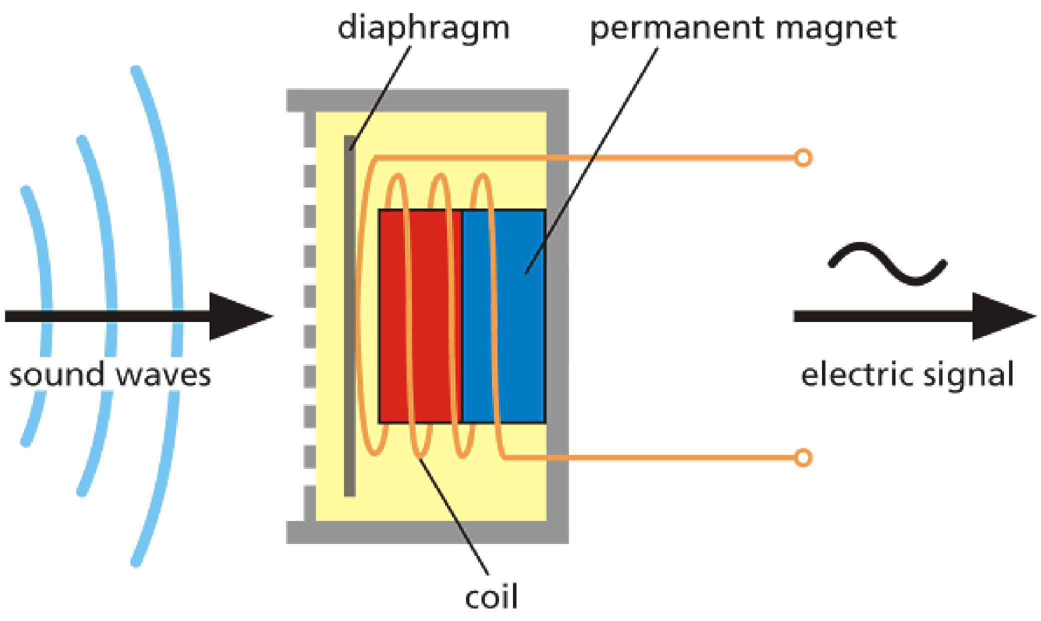
DIY Ribbon Microphone Response

- 0.8 dB max variation from 25 Hz to 16000 Hz:



Dynamic Microphones

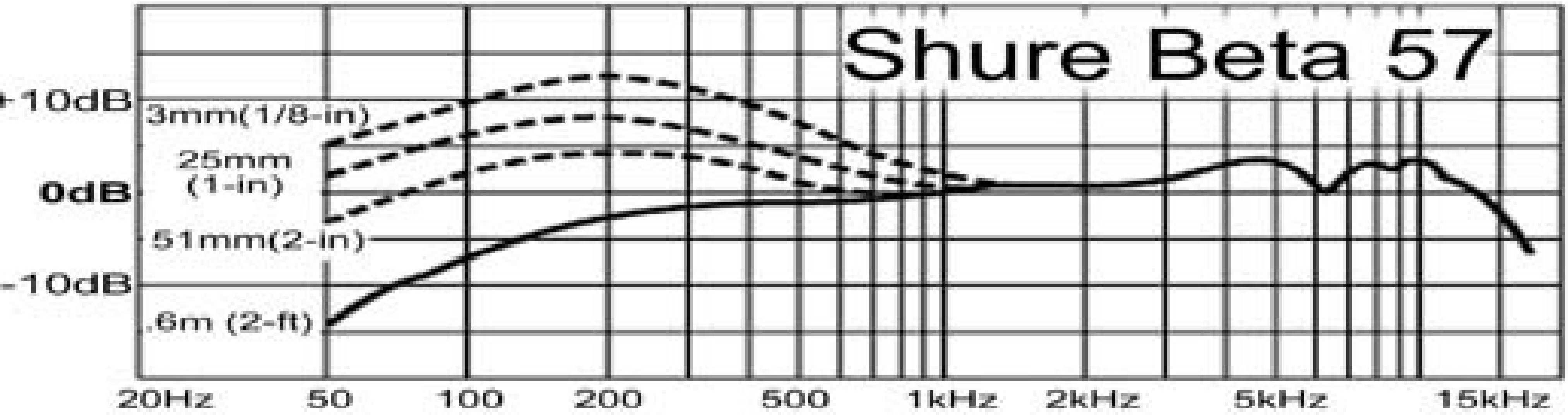
- Coil is attached to a diaphragm which moves over a magnet to generate current.
- More mass than a ribbon mic so less response, but much more rugged.
- Most common mic for live performance and “loud” sources (e.g. drums, vocals.)



Shure SM57 & SM58

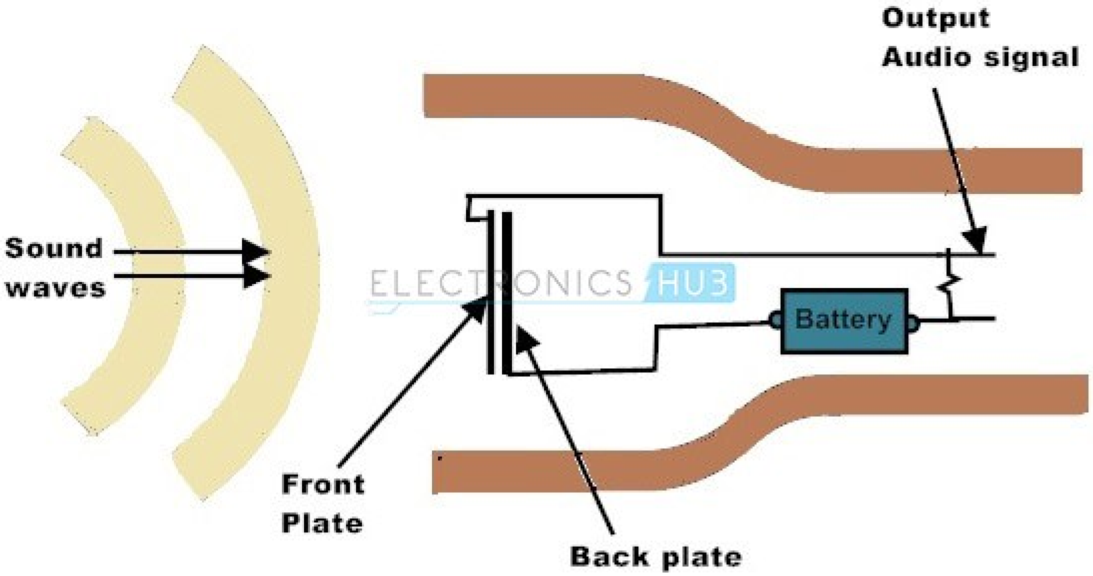
- Produced 1965 - present.
- Street price $70 - $100.
- The most-used mics in the world.
- Can easily handle high SPLs from vocals, drums, guitar amps, etc.
- Little point in modifying these (but people do.)
- SM58 has “ball” windscreen to minimize close proximity bass boost.

Directional Mic Proximity Effect



Condenser Microphones

- Condenser mics act like a capacitor where sound waves alter the distance between the plates.
- Most condensers require external power.
- Typically the most sensitive mics. Plate thickness typically 3-6 microns.



AKG C12 / Telefunken Ela M 250/251

- Produced 1953 – present.
- First patented condenser mic design.
- Has 9 pickup patterns (cardioid, omni, figure 8, and various "in between.")
- Contains vacuum tube amplifier and custom transformer.
- Street price of original $10,000 - $16,000.
- Street price of current models $6,700 - $8,500.
- A staple of Abbey Road Studios & BBC recordings.

Neumann U47/67/87

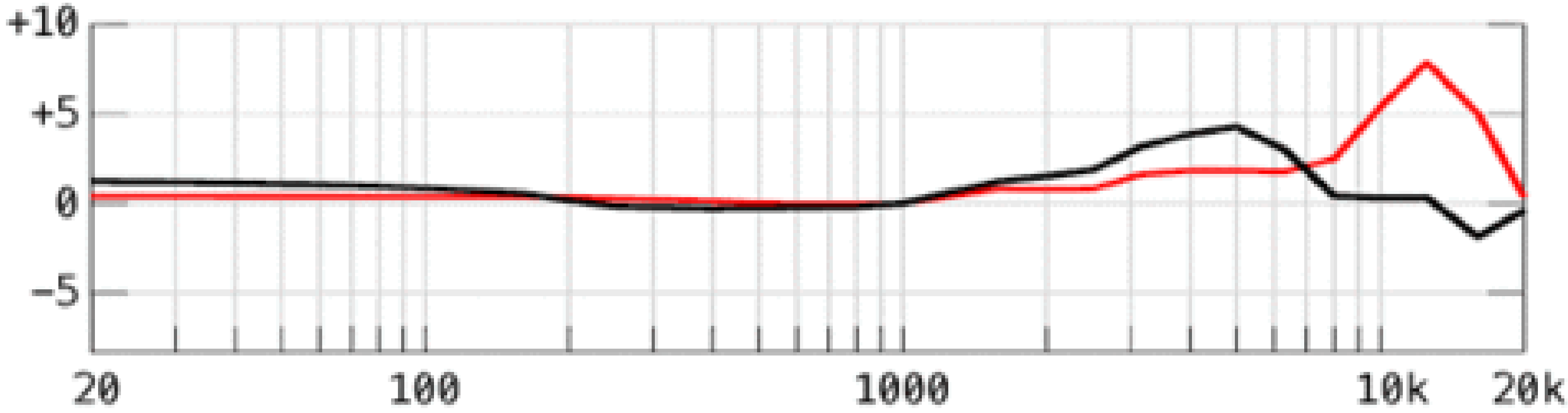
- Produced 1960 – present.
- Street price of current models about $5,000.
- Contains vacuum tube amplifier and custom transformer, with unique pre/deemphasis feature for noise reduction.
- Most widely copied microphone series in history.

Budget Condenser: MXL 990

- Produced 2005 – present.
- Street price new $50-$100, used $30-$50.
- Cardioid pattern only.
- Typical Chinese clone of Neumann U67, using pre-emphasis capsule *without* corresponding de-emphasis circuitry, resulting in overly harsh high-end.
- Favorite mic for DIY mods.

Typical Stock MXL990 Response

- Red line is stock mic: pre-emphasized capsule, but no corresponding de-emphasis circuitry in electronics.
- Black line shows a flat-response capsule installed, still with stock electronics (no de-emphasis.) High-end boost is reduced but not eliminated.



Typical MXL 990 DIY Modifications

- Replace the capsule ($160.)
- Replace the electronics with (improved) de-emphasis circuitry ($129.)
- Remove excess headbasket screen mesh ($0) or buy new headbasket ($34.)

… OR …

- Hire all upgrades done ($399.)

Typical Modification Results

- Red line is modified MXL 990 ($350-$400.)
- Black line is stock Neumann U87 ($5000.)

Another Approach - Modeling

- Several manufacturers now offer special mic/preamp combinations that model classic microphones in software.
- With this approach the modeled microphone can be changed *after recording!*
- Cost ~$1,000, plus extra mic model software modules.

Using Microphones Effectively

- An acoustically-treated recording space is important. Most small studios lack treatments due to expense. Costs are often $250K and up- if they're possible at all.
- The biggest problem with home studios is standing wave issues primarily caused by low ceiling height. Ceilings should be 10' high at minimum. The bigger all room dimensions, the better.
- The most pragmatic approach is to deaden the space as much as possible, then add room ambiance artificially via convolution reverb software later.

Electronic Drums

- Make minimal acoustic sound. Cost from $300 to $3,000+.
- Piezoelectric sensors in heads connect to a trigger-to-MIDI box, which connects to a computer to playback samples. Multiple sensors per pad/cymbal.
- Digital sampling works well because there is little interaction between the various drums and cymbals. Drum sounds are fairly short.

eDrum Sound Software Example

- “Superior Drummer” allows configuration of trigger pads (or keyboard notes) to drum samples.
- Customized pad sensitivity.
- Virtual mic placement and bleed (leakage) can be configured.

Edrums video

- http://edrum.for.free.fr/static/videos/PFozzDrumming.mpg
- This is about a two-minute video showing someone playing a real drum set modified with piezo triggers, into a computer running the “Superior Drummer” sample set.

Edrums video II

- http://edrum.for.free.fr/static/videos/CVHatsDemo.mpg

- Another video by the same guy, showing different sounds from a HiHat cymbal rigged up with separate piezo triggers on the bell, bow, and edge.

eDrum Sample Package Bloat

- Superior v1 (DFHS) – about 6 GB.
- Superior v2 – 19 GB.
- Superior v3 – 235+ GB. SSD recommended for best perforamance. Recorded in “11.1 surround” (with mics placed high above the drum set for “3D ambiance.”) 64 bit software package only.

Some Iconic Guitars

1959 Martin D28
$7,000-$15,000

1953 Gibson Super 400
$8,000-$12,000

1968 Coral Electric Sitar
$2,000-$5,000

Fender Stratocaster
$60-$30,000

Modeled Guitars – Line6 Variax

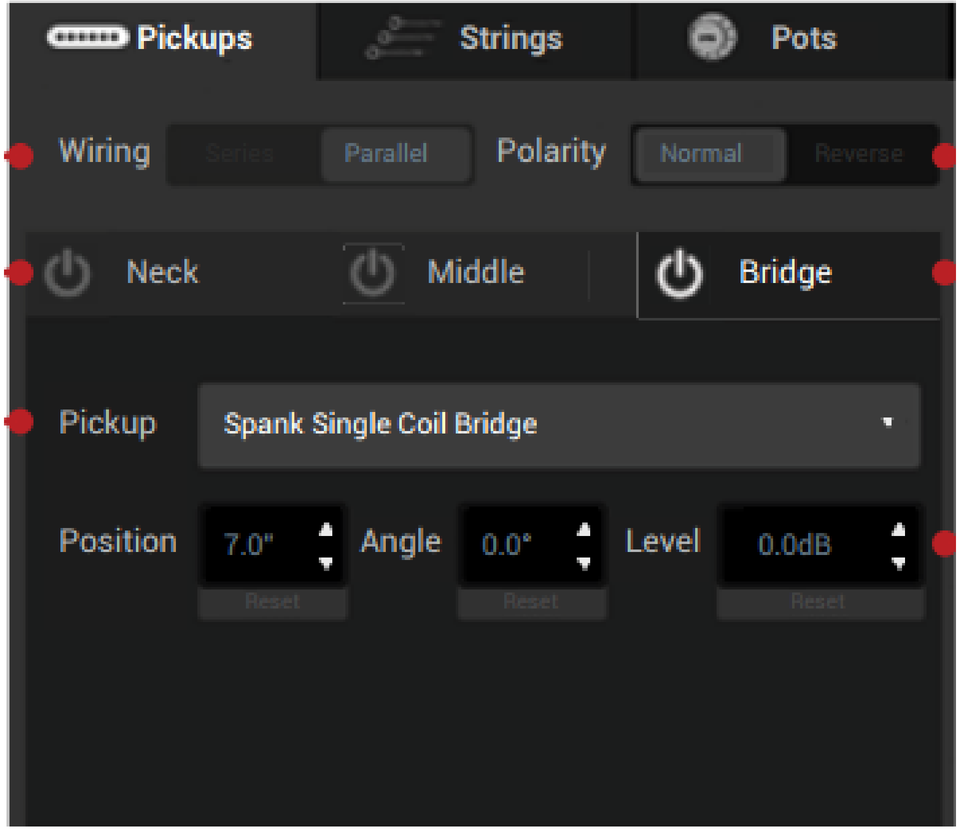
- Piezo pickups in bridge connect to internal electronics.
- Controls are software-definable for any purpose, e.g. alternate tunings.
- Has both digital and analog outputs.
- Cost $300-$700.

Customization software

- Many instrument parameters can be tweaked to customize instruments, or create ones that never existed.
- Acoustic guitar models allow selection of different body sizes & virtual mic placement.
- Electric guitar models allow unusual pickups & locations, even resistance of control pots.

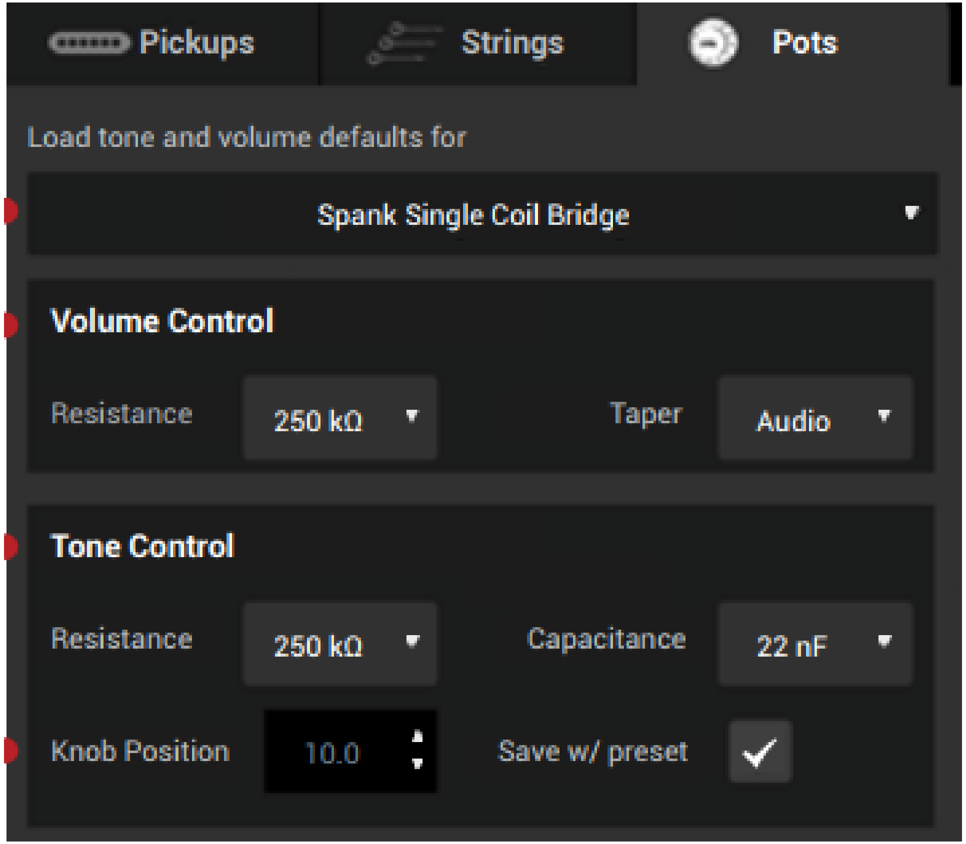
Example Pickup Customization

- Options include pickup wiring, angle, distance from the bridge and each other, angle (bass vs treble side,) individual string balance, etc.
- Some versions allow different virtual pickups on different strings, or ranges of strings.



Other Customization Options

- Pot options include resistance, taper, and capacitance.
- Functions of control knobs can be configured.
- String tuning and gauges can be selected, even those that would be impossible on a real guitar.



Keyboard Instruments

- Grand Pianos
- Electric Pianos
  - Rhodes Piano
  - Wurlitzer Piano
- Hammond Tonewheel Organ
- Hohner Clavinet

The Grand Piano

- Although not a legacy instrument, grand pianos are becoming an endangered species.
- Concert (i.e. ~9') grand pianos from Steinway or Bosendorfer cost ~$100,000.
- Exotic pianos cost even more (e.g. Fazioli, Ravenscroft $250,000).
- Most acoustic piano companies "authorize" software emulations of their products now (especially Ravenscroft!)

A Piano Is A Piano, Right?

- No. Size matters!
- The longer the string length, the less inharmonicity, hence better sound.
- Grand pianos have significantly different (better) action mechanisms than upright pianos.
- Emulations address all these qualities and allow things impossible on physical pianos.

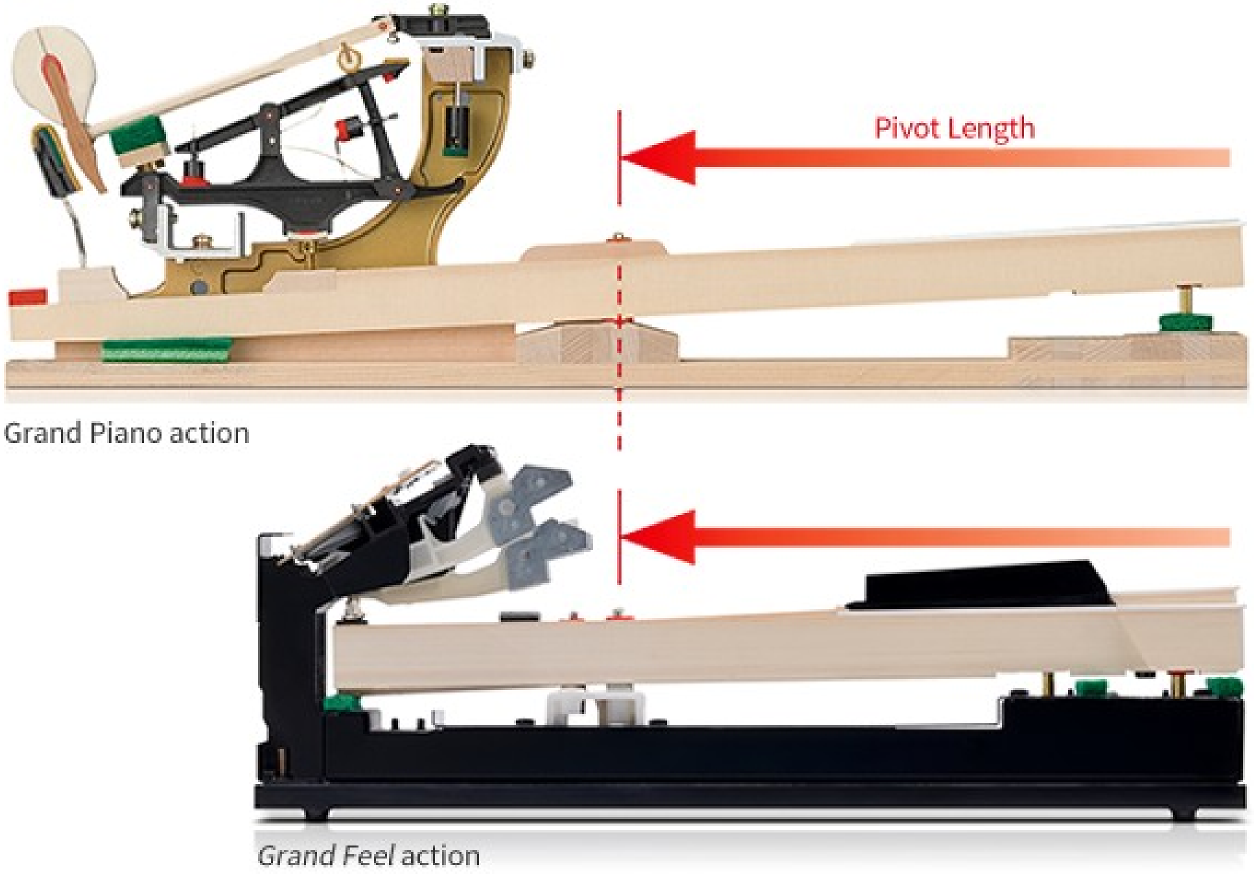
Piano Characteristics

- 88 notes (or more.) Upper notes have 3 strings per note, midrange notes have 2 strings per note, bass strings only one. Pianos contain up to 12,000 individual parts.
- Pianos can be “voiced” by conditioning the hammers for brighter or darker sound.
- Damper pedal releases dampers on all strings allowing sympathetic vibration and smoother legato playing.
- Only grand pianos have:
  - *Una corda* (soft) pedal moves hammers sideways to contact only one string.
  - *Sostenuto* pedal lifts dampers only on notes already held down.
  - “Double escapement” action retracts hammers in two steps to facilitate rapid repeated notes.

Controller Example: Kawai VPC1 (~$1,800)

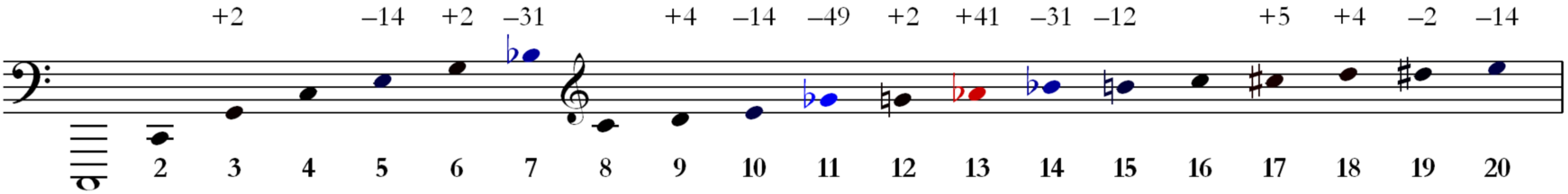
Piano Actions – Under The Hood

- Upper is real grand piano, lower is VPC1. Note same pivot point on both.
- VPC1 has clever “reversed” hammer with triple optical position sensors and counterweights.



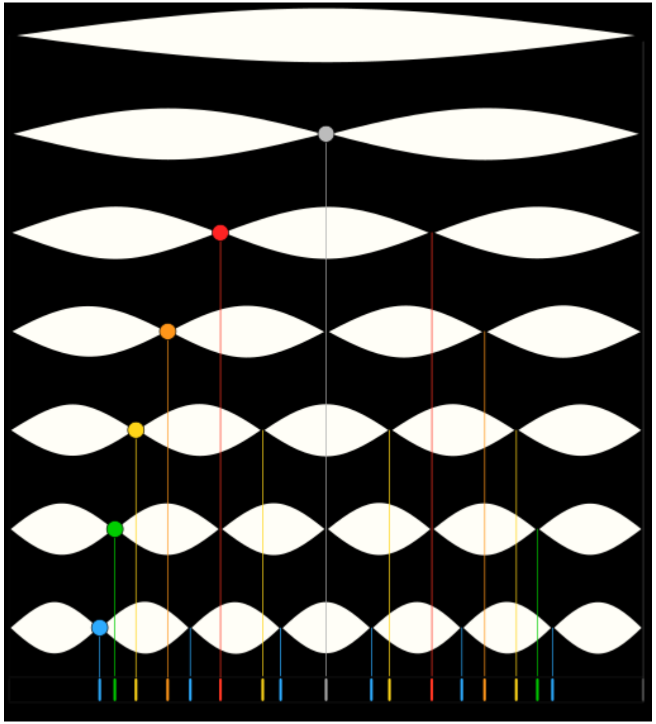
Piano String Impact Location

- Piano hammers contact the strings at 1/7th of their length (node of the 7th harmonic, considered the first “bad” harmonic, typically 31 cents flat from ideal.) But the location varies for each string!
- Of no concern to electronic emulations of course.

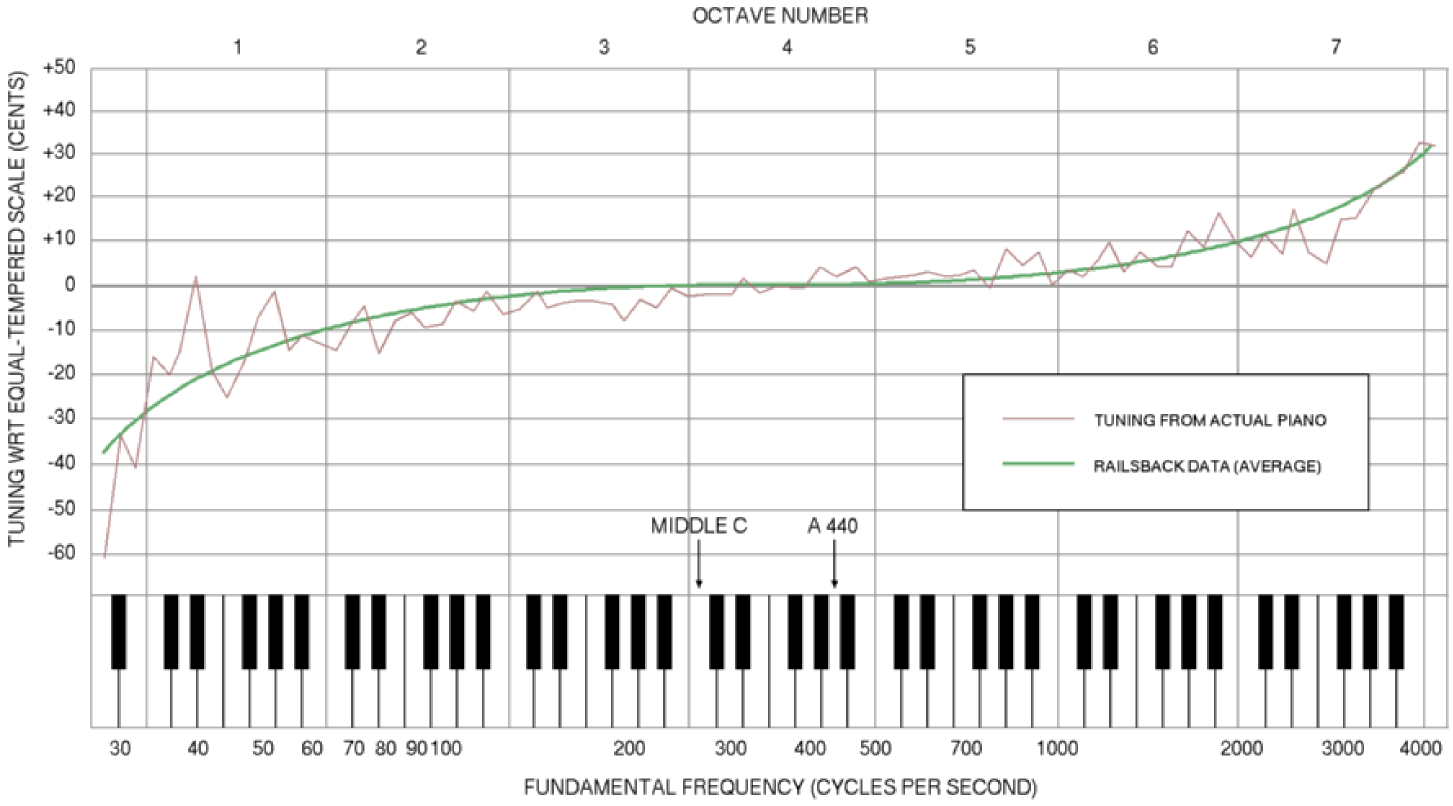


Piano Strings & Inharmonicity

- String vibration nodes are only at perfect harmonic intervals if a string is perfectly flexible. None are of course.
- The shorter the string in proportion to its thickness, the stiffer it is. Which means it vibrates faster. So …
- The higher the harmonic, the sharper it is! Which is why pianos with longer (i.e. less stiff) strings sound “better.”



Piano “Stretch” Tuning



Piano Simulations

- String inharmonicity doesn’t exist. But it can be emulated in software if desired.
- Unconventional tunings are possible, and can be switched on the fly while playing.
- Intentional detuning, hammer wear, and mechanical noises can all be emulated by switching samples or via computations.

Physical Modeling vs Sampling

- Purely sampled piano libraries have a large storage footprint (e.g. ~122 GB+) but samples are still acquired one note at a time, in isolation.
- Samples cannot deal with harmonic interactions of notes being played together, or partial depression of pedals.
- Steinway recently pulled their “authorized” status from a sampled piano library and gave it to a physically modeled emulation.

Pianoteq Modeling Parameters

Pianoteq Virtual Mics & Listener

- Up to 5 virtual mics can be positions in X/Y/Z space.
- Lid can be closed or open to varying degrees, or off completely.
- The speed of sound to each mic can be individually adjusted (!)

Fender-Rhodes Electric Piano

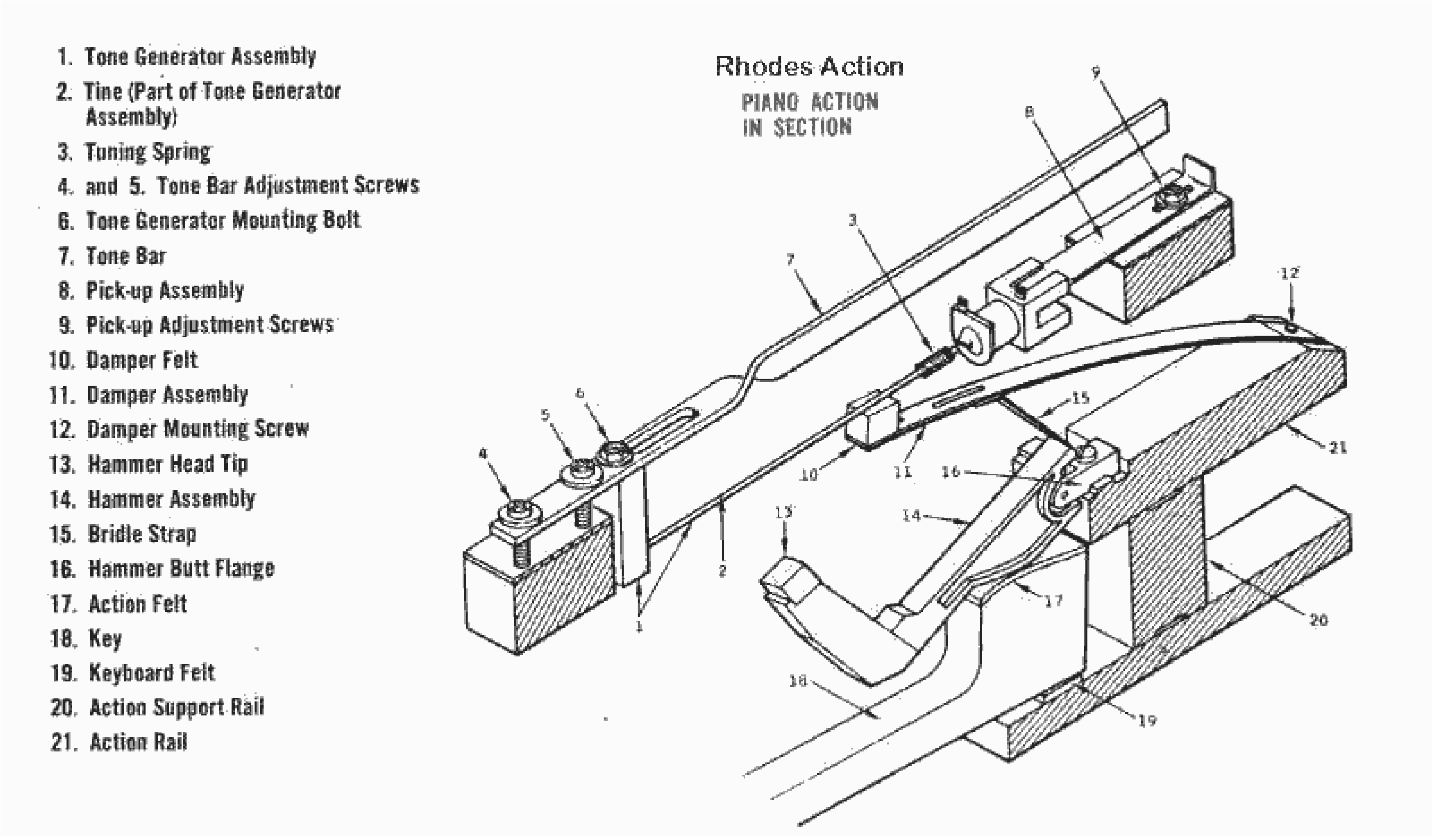
- Invented by Harold Rhodes in WWII to provide entertainment for convalescing GIs.
- Licensed to Fender and mass produced 1965-1984.
- “Stage” (shown) and “Suitcase” versions (with integral amp) produced in both 73- and 88-key versions. Weight 75-100 lbs.
- Significant resurgence of popularity this century.
- Restored versions cost ~$3,000-$5,000.

Rhodes Under The Hood

- Mechanical hammers strike a “tine,” essentially one arm of an asymmetrical tuning fork.
- Notes are tuned by moving a coil spring along the length of a tine.
- A magnetic pickup is positioned behind each vibrating tine.

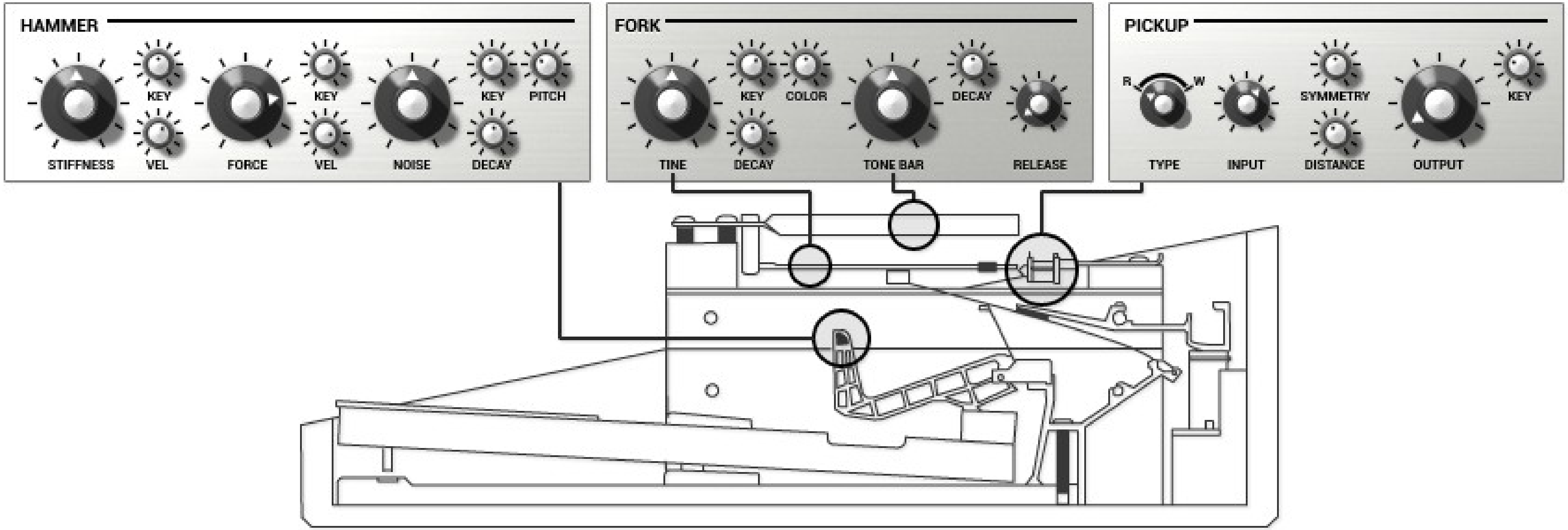
Rhodes Action

- Pickups could be adjusted back and forth and at angles to the tine for various tones.
- Significant mechanical noises contribute to a Rhodes' distinctive sound.



Applied Acoustics Lounge Lizard

Allows tweaks of essentially all mechanical adjustments available on a physical Rhodes piano.



Wurlitzer Electric Piano

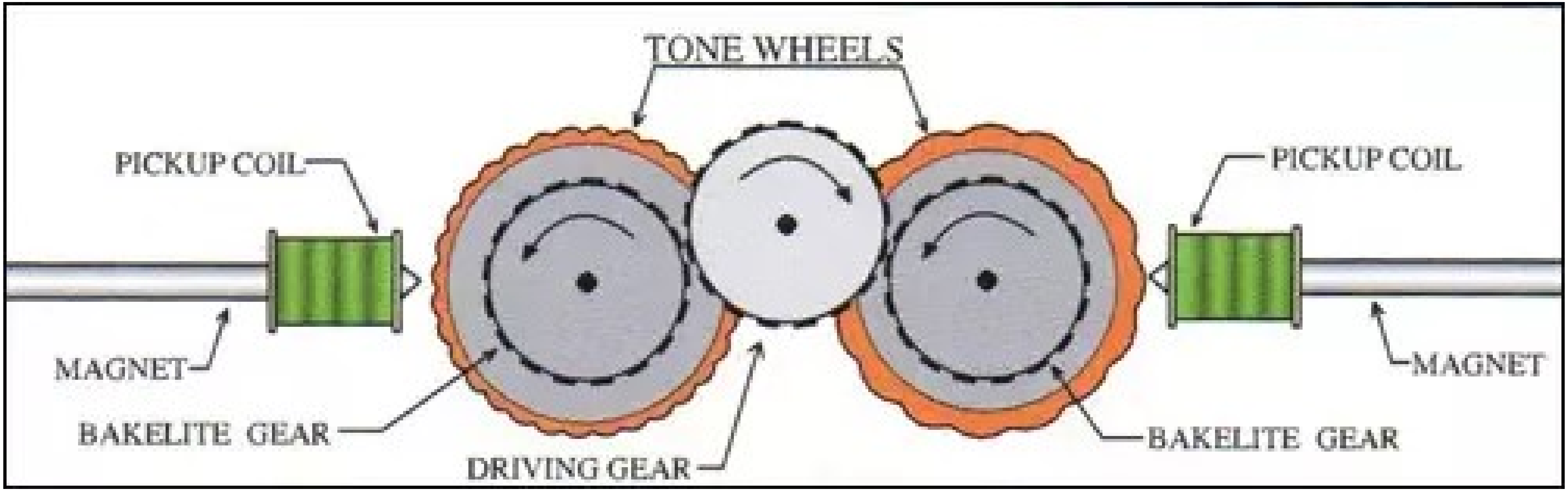
- Various models produced 1954-1980. Only 64 keys. 56 lbs. Small internal speakers.
- Street price $2,500-$5,000 restored.
- Internally similar to a Rhodes piano, but with different tines and a distinctive softer sound.
- Tuned by soldering (and filing off) globs of solder on the tines. A tedious job.
- Many software emulations.

Hammond Tonewheel Organ

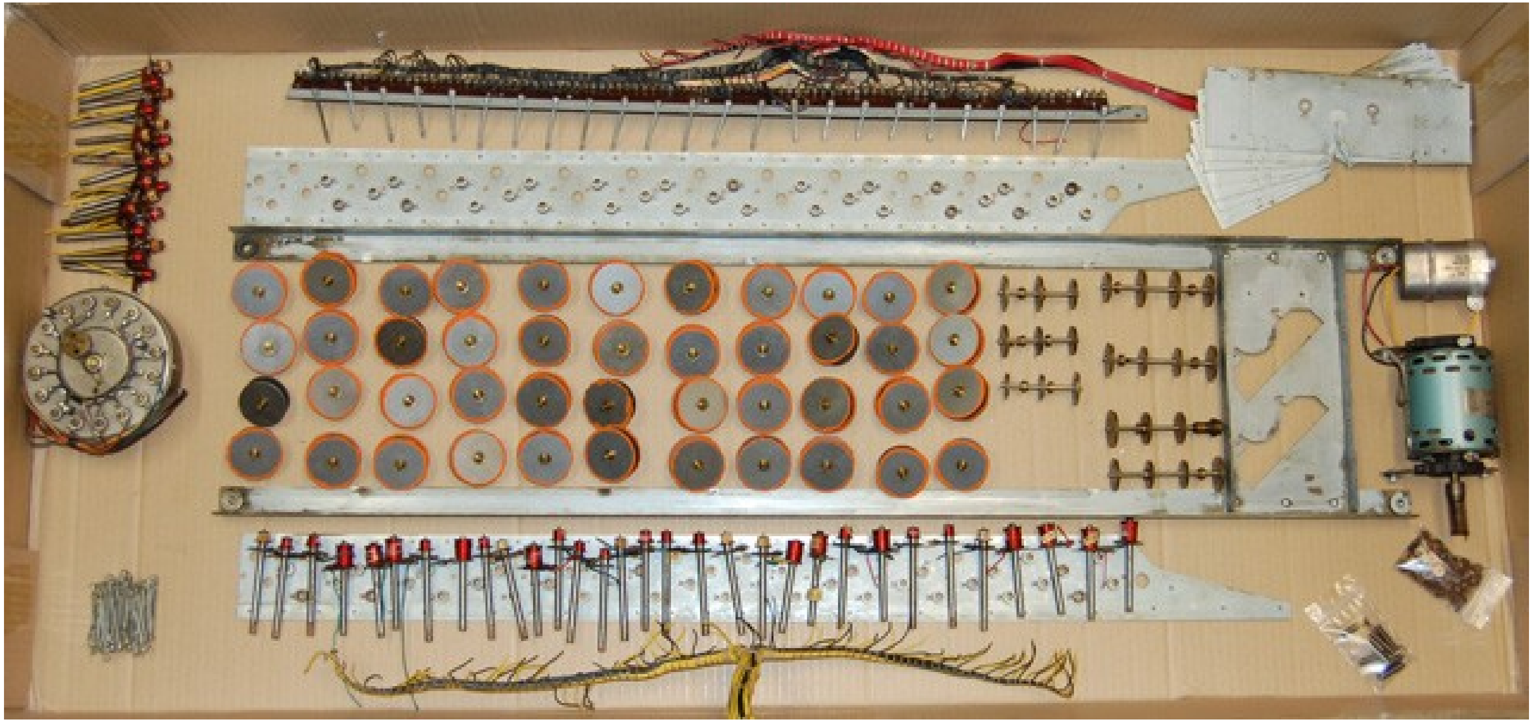
- Street price $3,000 – $13,000. Average $5,000 for B3 model (most popular).
- Produced 1935-1975. Electronic variants produced 1976 – present.
- Weight of B3 model is 425 lbs. without Leslie speaker (another 149 lbs.)
- Originally intended as church organs, but adopted by many iconic jazz, rock, and pop artists.

Hammond Tonewheel Operation

- Tonewheels are motor-driven “gears” whose numbers of teeth are determined by musical intervals.

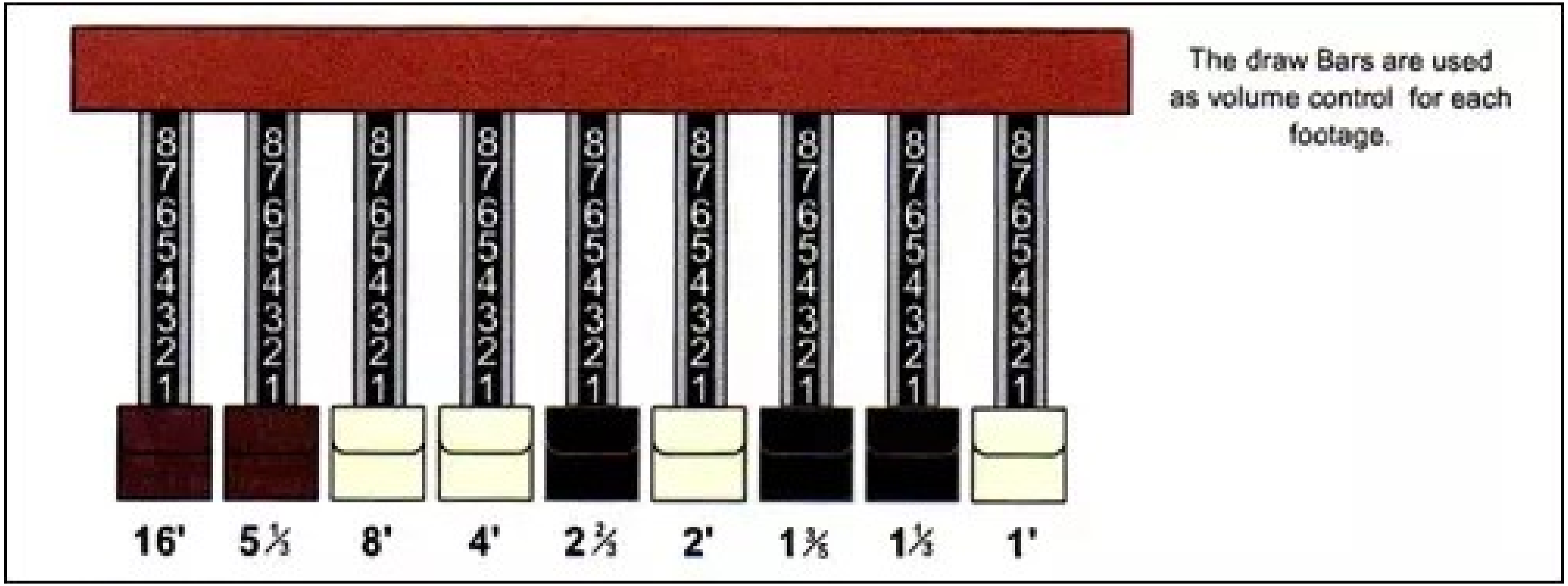


Hammond Tonewheel Generator Disassembled



Hammond Tonewheel Controls

- Sets of “Drawbars” control the relative amplitude of the pickups for each tonewheel.
- Note pipe organ “foot” terminology.



Leslie 122/147 Speaker

- Note dual rotating treble horn near top, and single woofer firing down into a rotating “elbow” baffle below.
- Horn and bass baffle rotate in opposite directions.
- Performers could select “fast” or “slow” speeds (but not “off”). Horn/baffle change speeds at different rates.

Hammond Quirk #1 – Key Click

- A design defect caused arcing of the key contacts when a player initially pressed a key, resulting in a “click.” Subsequent key presses caused no clicks as long as other keys were held down.
- Players capitalized on this quirk to add variety to note attacks as they played.
- Hammond eventually eliminated this defect but restored it when musicians complained.
- All emulations include this “defect.”

Hammond Quirk #2 – Crosstalk

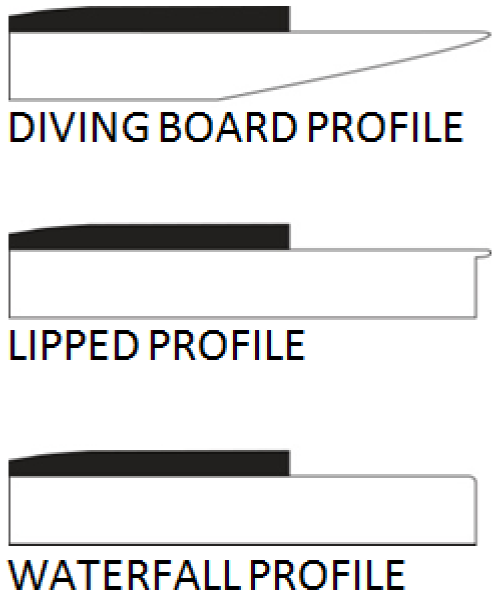
- Pickups can sometimes sense neighboring tone wheels, resulting in impure tones.
- Most common on real instruments than need maintenance.
- Considered part of the sound, most emulations allow for adjustable degrees of crosstalk.

Hammond Quirk #3 – Pitch Bend

- When the organ is switched off while a note is playing, the motor spins down before power dies in the tube electronics. This results in a momentary downward "pitch bend" (until power dies completely or the player switches power back on.)
- This causes significant wear and tear on the motor of a real organ and violates the recommended power up/down procedure! Players do it anyway.
- Most emulations allow pitch bend (in both directions) via keyboard controllers.

Hammond Quirk #4 – Palm Gliss

- As a cost saving measure, Hammond dispensed with the “lip” on the front key edge that was traditional with ivory piano keys.
- This allowed players to play a multi-note glissando with their palm on the front key edge, instead of a single-note glissando with their fingernail.



Hammond Emulator – NI B4

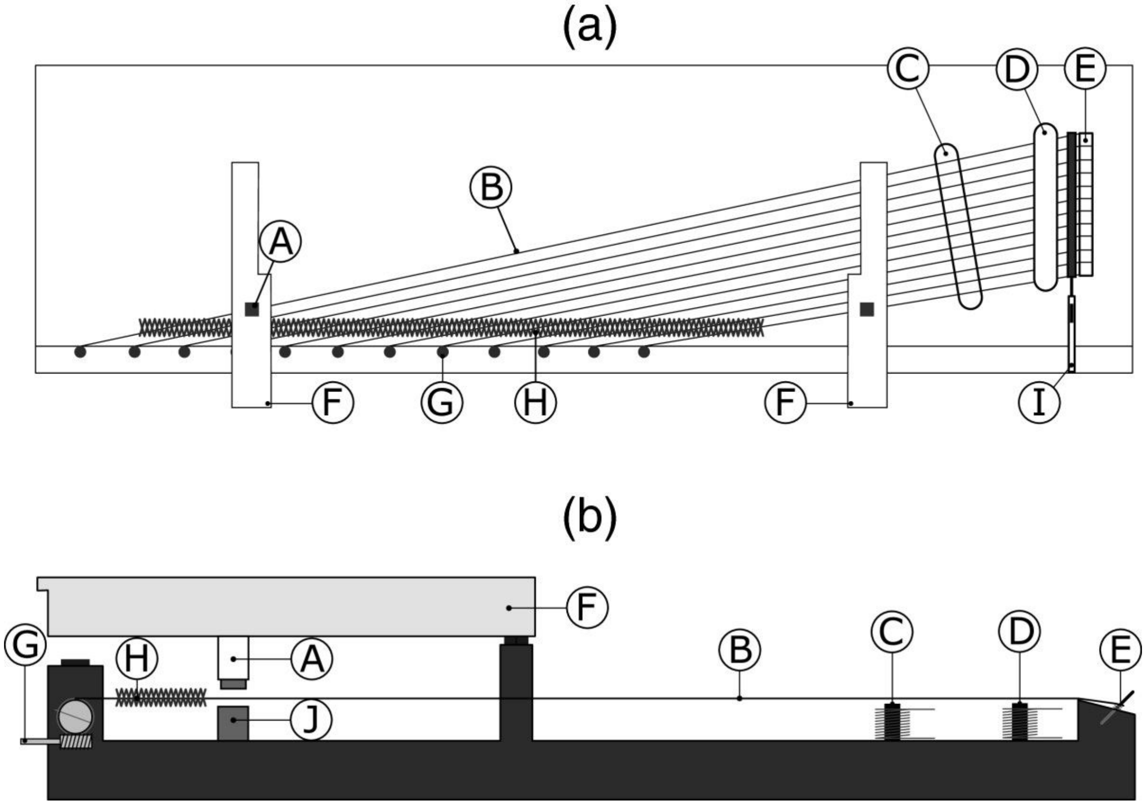
- Early Hammond Organ emulation allows all “quirks” previously mentioned.
- Can map single keyboard controller into zones for upper and lower organ keys.
- Models 16 levels of deterioration & crosstalk. Adjustment of keyclick tone & amount. Virtual mic/listener location for Leslie speaker.
- Extends capabilities to include velocity sensitive keyboard response, piano sustain pedal emulation, distortion, Leslie speaker acceleration tweaks.

Hohner Clavinet

- Inspired by the clavichord, but internally quite different.
- Distinctive “funky” sound still popular today.
- Several models produced 1964-1982. “D6” most popular.
- Restored versions $2,000-$3,000.

Hohner Clavinet Action (Top & Side Views)

- When the key (F) is pressed, the tangent (A) slams the string (B) into the anvil (J), causing string vibrations for pickups (C) and (D).
- Upon release, the note is dampened by felt (H).
- Strings are banjo strings tuned with a screwdriver at (G).
- Tangent (A) has a replaceable rubber hammer similar to a pencil eraser.



Clavinet Problems

- Extremely noisy electronics. No internal shielding.
- Strings eventually wear a groove in the rubber hammers, causing the strings to get “stuck” when pulled up on release, then make an unwanted “click” sound when they break free. (This is so common it is part of most software emulations!)
- “Vibanet” (contemporary Clavinet semi-clone) addresses most mechanical problems. Cost $5,000+.

Clavinet Emulations

- Many, including freeware.
- Most mimic the traditional white “D6” rocker switch controls and mute slider seen here.
- Few (none?) allow virtual adjustments to pickups, etc.

Strings/Brass/Woodwinds

- Sophisticated sample libraries of acoustic instruments abound.
- Used extensively by film and television composers. (Only video games have sufficient budgets to hire real orchestras.)
- Most libraries offer comprehensive articulation control, like the bowed string “keyswitches” shown.

Does It Work? Ask Mark Isham

- Possibly the most successful Hollywood film composer you never heard of. 46 Awards (Emmy/Grammy/Golden Globe/etc.) Scored 336 movies and television shows, resulting in $3 billion in revenue.
- Rarely works with real musicians. Plays all parts himself on a keyboard (although he’s a trumpet player!)

Mark's Studio: What Clients Expect To See

Note giant mixing board, big video monitor, outboard rack mount signal processors, and boutique reference speakers.

Mark’s Studio: Another View

Where The Real Work Gets Done

- 12-core Mac Pro work computer.
- 8-core Mac Pro “print” computer for delivery.
- 2@ Mac Pro “sampler” computers each running Vienna Sound Libraries ($13,000 each).
- Cheap $100 keyboard controller.
- Modest monitor speakers.
- Back wall is museum gear.

Recording

Twenty years of learning my craft, hours of practice each day and NOW they call me gifted!!

– Unknown classical musician

Recording Developments Through the 1980s – Mostly Hardware

- 1949 – Les Paul aligns an additional playback head with the record head on his multi-track tape recorder, allowing him to record new material in sync with previous material.
- 1983 – The *Musical Instrument Digital Interface* (MIDI) is adopted by most manufacturers, allowing diverse pieces of electronic music gear to control each other. Sequencer software products appear on personal computers that record MIDI (but not audio) data.
- 1989-1992 – The first consumer *Digital Audio Workstations* (DAWs) appear that record multitrack audio to computer hard drives instead of tape, offering non-linear access/editing. Proprietary hardware is necessary to provide adequate throughput.

Recording Developments Through the 1990s – Mostly Software

- 1996 – The first DAW software plug-in audio architectures are formalized, allowing integration of DSP software modules for audio processing (e.g. reverb and compression) that previously required external hardware.
- 1999 – Plug-in architectures are expanded to include modules that accept MIDI input as well as audio, allowing for instrument emulations as well. A large market develops for (mostly) electronic synthesizer emulations.

21st Century Developments

- 2000 – The first hard drive streaming sampler products appear, allowing performers to stream unlimited-length audio from hard drives on demand with negligible latency. Meticulously-sampled commercial sound libraries follow.
- 2002 – *De facto* standards for inter-process audio routing emerge (e.g. ReWire, Core Audio, JACK) to route audio between separate applications, not just within tightly-coupled host/plugin environments.
- 2006 – Dante, AVB, and AES67 technologies leverage network standards like QoS and PTP to achieve sub-millisecond audio latencies over Ethernet networks, allowing distributed audio processing.

Tracking (Recording) Today

- Rarely does pop music get recorded by an entire ensemble.
- Typical workflow is to start with a “rough” (overall recording,) then re-record each individual part until everything is as good as possible.
- Recording the “rough” is generally not demanding on equipment- just like recording into an old-school tape recorder. “Draft” quality is totally acceptable.

Tracking (Overdubbing) Today

- Recording along with pre-existing material is where the problems start to come in.
- Synchronization of the incoming audio is *not* critical (it can be corrected later), but synchronization of cue audio to the performer *is* critical.
- Commercial products rarely talk about the issues involved. But it all boils down to one word ...

The Elephant In The Room

LATENCY

Latency: Critical Timings

- Speed of Sound = ~1 foot / ms.
- “Loss of Simultaneity” = ~20 ms. So …
- In the real world, musicians > 20 feet apart cannot play in sync with each other.

- Speed of MIDI data transmission = 31250 baud.
- Time to play an 8-note chord over MIDI = ~ 5.5 ms
- Time to release that 8-note chord over MIDI = another ~5.5 ms. So ...
- Virtual instruments have about 10 ms of processing time before things get sloppy.

Different Latency Scenarios

Virtual Instruments

- MIDI transmission lag.
- Audio processing.
- Output buffering.
- D/A conversion.

Analog recording

- A/D conversion.
- Input buffering.
- Audio processing.
- Output buffering.
- D/A conversion.

Parameters Determining Latency

- Driver Buffer Size.
- Sample Rate.
- Hardware I/O Efficiency.
- Software Efficiency of Drivers.
- Software Efficiency of Applications.
- General CPU processing capability, for audio and other tasks.

Buffer Size

- Number of samples to accumulate before the audio packet is transferred.
- Usually a power of two.
- The smaller the buffer size the quicker the response (as the CPU spends more time in an interrupt handler), but more CPU power is required.
- At some point the CPU runs out of time to process and produces pops/clicks/gaps in the audio stream.

Common Sample Rates

- The faster the sample rate, the less time between samples, but more CPU power is required.
- 44100 samples/second for CD audio.
- 48000 or 96000 for almost everything else (streaming audio, film, digital video, etc.)
- Optimally set to the sample rate of the delivery medium to avoid inherent loss of precision with sample rate conversion. But ...
- Common practice is to use 48K or 96K and downsample for CD targets, since they are no longer the dominant delivery medium and sample rate conversion routines are quite good now.

Even Higher Sample Rates

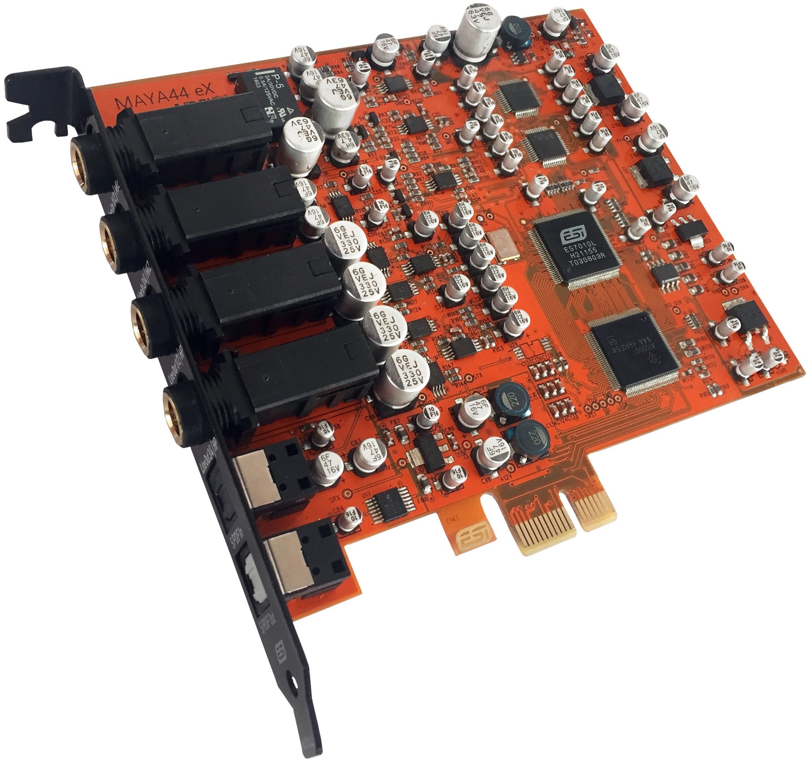
- Although no current production media requires such rates, many current interfaces can run at 192 KHz, 384 KHz, or even faster.
- The human ear cannot hear sounds above 20 KHz, but listeners often claim they can hear an audible difference in these high rates. How can that be?
- These differences can be attributed to distortion and artifacts in the converters and supporting circuitry. Each hardware system will be different.
- Audio hardware experts claim ideal sample rates around 66 KHz for minimal audible artifacts. But this is nowhere close to any standard sample rate.

Audio Interface I/O Efficiency

- Interface types (from best to worst)
  - PCI Express (Bus card, Thunderbolt 2 & 3)
  - PCI bus card
  - Firewire (IEEE 1394 A and B)
  - Ethernet (Dante, AVB, AES67)
  - USB (1, 2, 3)

PCI Express card

- ESI Maya 44EX ($180)
  - Not commonly available
  - Only 4 channels in and out
  - Only 1 card per system
  - Cannot be used with laptops



Thunderbolt

- Many interfaces available $300 - $3,000. (Focusrite Clarett 4 shown $550.)
- Thunderbolt is essentially PCI Express on a cable.
- Adapter cables allow use of Firewire interfaces.
- Incompatibilities between Thunderbolt 2 (current audio interfaces) and Thunderbolt 3 (on recent Macs and PCs.) Especially problematic on Windows.

PCI bus

- PCI bus is almost extinct even on desktop systems.
- No current pro PCI audio interfaces exist that I know of.
- Vendors typically do not update drivers for legacy devices. They want to sell you new hardware!

Firewire

- Invented by Apple but unavailable on current Mac or Windows laptops. (Thunderbolt can connect to Firewire however.)
- Developed for streaming data (video primarily) so it is well suited for audio.
- Unlike most USB implementations, Firewire can reserve a specified amount of bandwidth to allow glitch-free streaming.
- Silicon implementations have been poor. Only TI chipsets work reliably for audio. Windows drivers support has been poor and further deteriorated under Windows 10.

Ethernet

- Main technologies on the market:
  - Dante (proprietary, layer 3, fastest audio, no video support)
  - AVB (open source, layer 2, slower audio but video support)
- AES67 (yet another standard) aims to allow interoperability various Ethernet audio protocols.
- Uses IEEE-1588 Precision Time Protocol (PTP) to obtain microsecond resolution and latencies in the 1-4 ms range.
- Relatively expensive ($600-$3,000.) AVB requires qualified network switches.

USB

- The Universal Serial Bus is truly universal. The most common audio interface type found today. $80 on up.
- USB data transfers incur overhead from bus acquisition and relinquish cycles. Poorly suited to streaming applications unless a custom USB stack is utilized.
- Hardware support is spotty. Only NEC chipsets work reliably for streaming audio from a host.
- Many laptops lack any Cardbus or Express card slots to add a suitable USB interface if the on-board interface is unreliable.

Beware Vendor Latency Claims

- Focusrite Scarlett 18i20 ($500) USB 2 interface claims 2.74 ms roundtrip latency using:
  - Buffer size = 32 samples
  - Sample rate = 96 KHz
- But at more typical settings:
  - Buffer size = 128 samples
  - Sample rate = 48 Khz

  Latency is 21.92 ms!
- The best interfaces can achieve roundtrip latencies of ~6 ms at these settings, but they replace the USB software stack in the host.

Minimizing Recording Latency

- External hardware.
- “Direct Monitoring.”
- Reduced cue mixes.
- “Freeze” tracks to audio (less demanding on playback), then mute the original track. “Unfreeze” when done.
- Avoid high-latency plug-ins (e.g. compressors, pitch shifters.)

Plug-in Latency

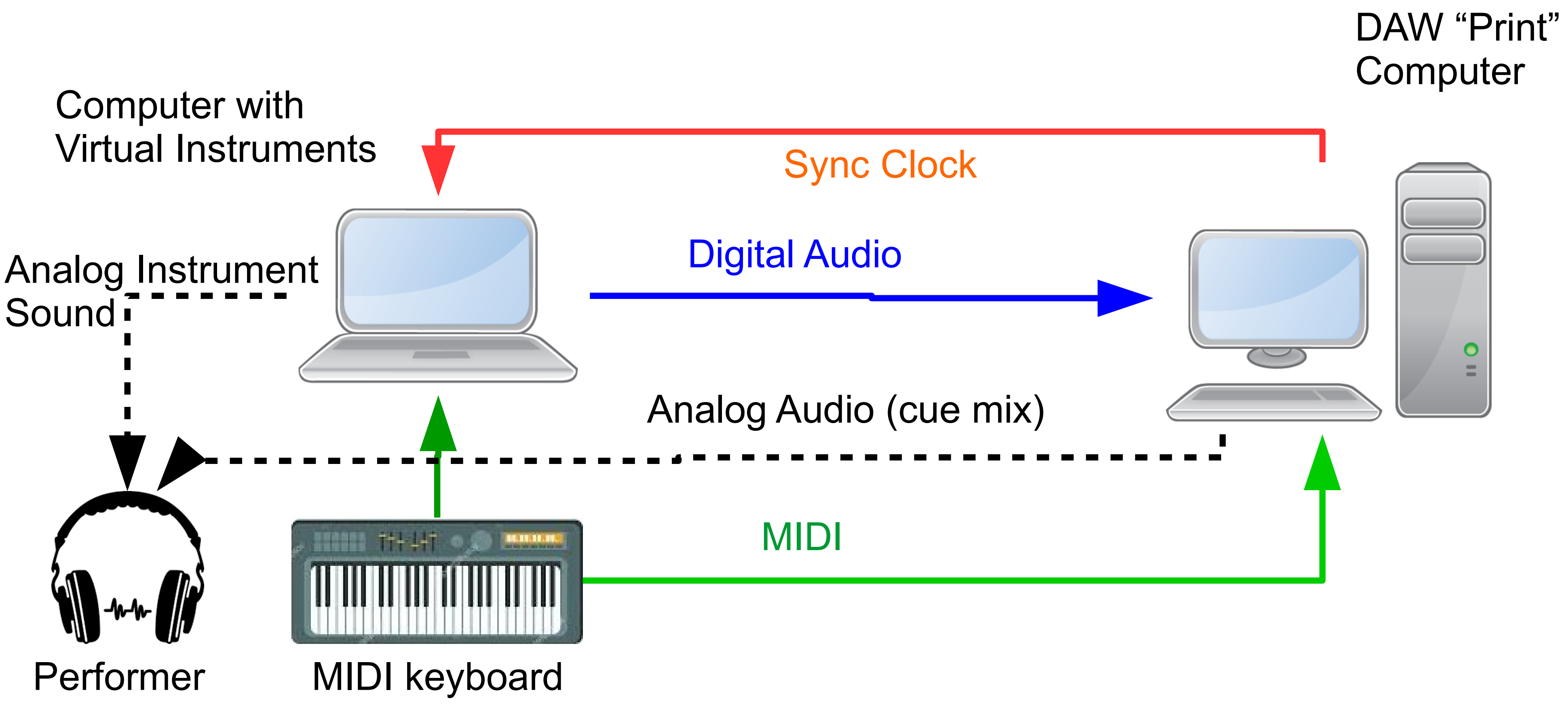
- Latencies are highly variable. Most vendors do not publish latency specs. Waves specs latency in samples, not ms.
- Plug-ins that need to "look into the future" generally have highest latencies.
- "Delay compensation" in hosts syncs all tracks to the highest latency active plug-in. Guaranteeing worst case latency!

Waves Plug-in	Purpose	48K Latency (in samples)
IR1	Convolution Reverb	0
Morphoder	Vocoder	639
C4	Multiband compressor	65
Waves Tune	Pitch correction	3072
Waves Tune RT	Pitch correction	0
DeBreath	Noise removal	35248
Grand Rhapsody	Virtual Grand Piano	64

Minimizing Virtual Instrument Latency

- Second “Instrument” computer feeds primary “print” DAW.
- Can record both (digital) audio and MIDI data allowing re-orchestration later.
- No latency contribution from A/D conversion or input buffering.
- Timing of recorded audio and MIDI may be significantly skewed on DAW, but can be corrected by aligning tracks later.
- “Freeze” tracks on DAW if/as necessary.

Virtual Instrument Recording



Mixing

Audio processing is the art of balancing subjective enhancement against objective degradation.

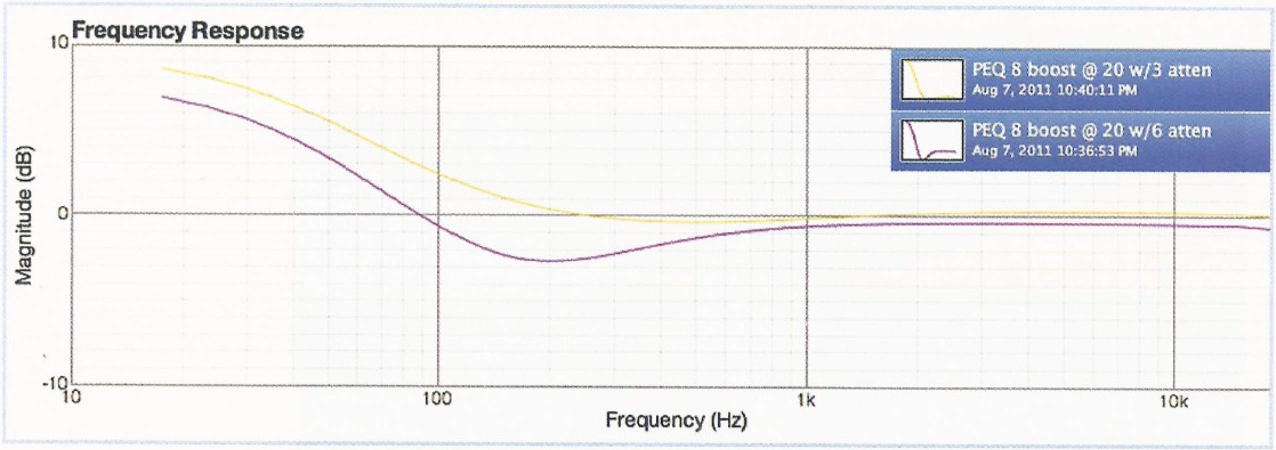
– Bob Olhsson

Mixing

- The mixing engineer's job is to modify the raw recorded multi-track audio to meet the producer's goals.
- Processing typically includes equalization, compression, and spacial location (front-to-back, side-to-side or surround.)
- The end product is usually a two-track mix.
- Tunes are mixed one at a time.
- Often a time-consuming job- 6 hours to mix a 3-minute song is not unusual.

Processing Gear Quirks: Pultec EQP-1A (1950s)

- Classic hardware processors often had quirks that were exploited by recording engineers.
- Example: Pultec EQ response curve with simultaneous bass boost and bass cut:



Pultec response curve obtained by simultaneously applying bass boost and bass cut. The violet curve has more bass cut than the yellow.

Processing Gear Quirks: 1176 Limiter (1967-present)

- Radio buttons select the compression ratio (see photo). But due to a design flaw, all four radio buttons could be depressed together!
- Distortion and lag time increase, plus compression curve becomes complex, giving more “pop” to transients.
- Known as “British mode.”
- Software emulations include this important quirk.

Typical Mixing Session

- Note floating windows for Plugin controls, mixer pane along bottom, tracks with audio waveforms top center.
- Sessions get cluttered fast!

Automation

- Automation allows a DAW to record movements of controls while audio is playing, then “play them back” later. This allows on-the-fly changes that would be impossible for a human with only two hands.
- Only the most expensive hardware mixers have automation. ($500,000+?)
- Latency while recording automation is usually not an issue since smooth, slow changes are the norm.

Typical Latency Tolerance

- ~250 ms: Tolerable for mixing and mixer automation recording, but somewhat sluggish.
- ~100 ms: Reasonable for mixing and mixer automation recording.
- 20-30 ms: Tolerable when playing an electronic keyboard with slow attack sounds (e.g. bowed strings.)
- 10-20 ms: Sluggish response when playing a piano virtual instrument. Electronic drums even more sluggish.
- ≤ 10 ms: Good response. Latencies below 10 ms offer little to no discernible improvement but greatly increase CPU burn.
- Values vary by individual sensitivity and preference.

Mastering

Limiters and compressors should be treated just like firearms.

- Roger Nichols, 8-time Grammy award winning engineer

Mastering

- Technically should be called Pre-Mastering.
- The job of preparing finished mixes for release on the appropriate media:
  - CD
  - Vinyl records (singles and albums)
  - Streaming Media Sites (iTunes, YouTube, Spotify, Soundcloud)
  - Terrestrial Radio
  - Satellite Radio
  - Broadcast TV
  - As embedded content (e.g. in a movie, video game, etc.)
- More complex today than ever because of multiple delivery systems. Most projects are released on (and optimized for) multiple media types.
- Arguably the most arcane skill in the production chain.

Mastering for CD

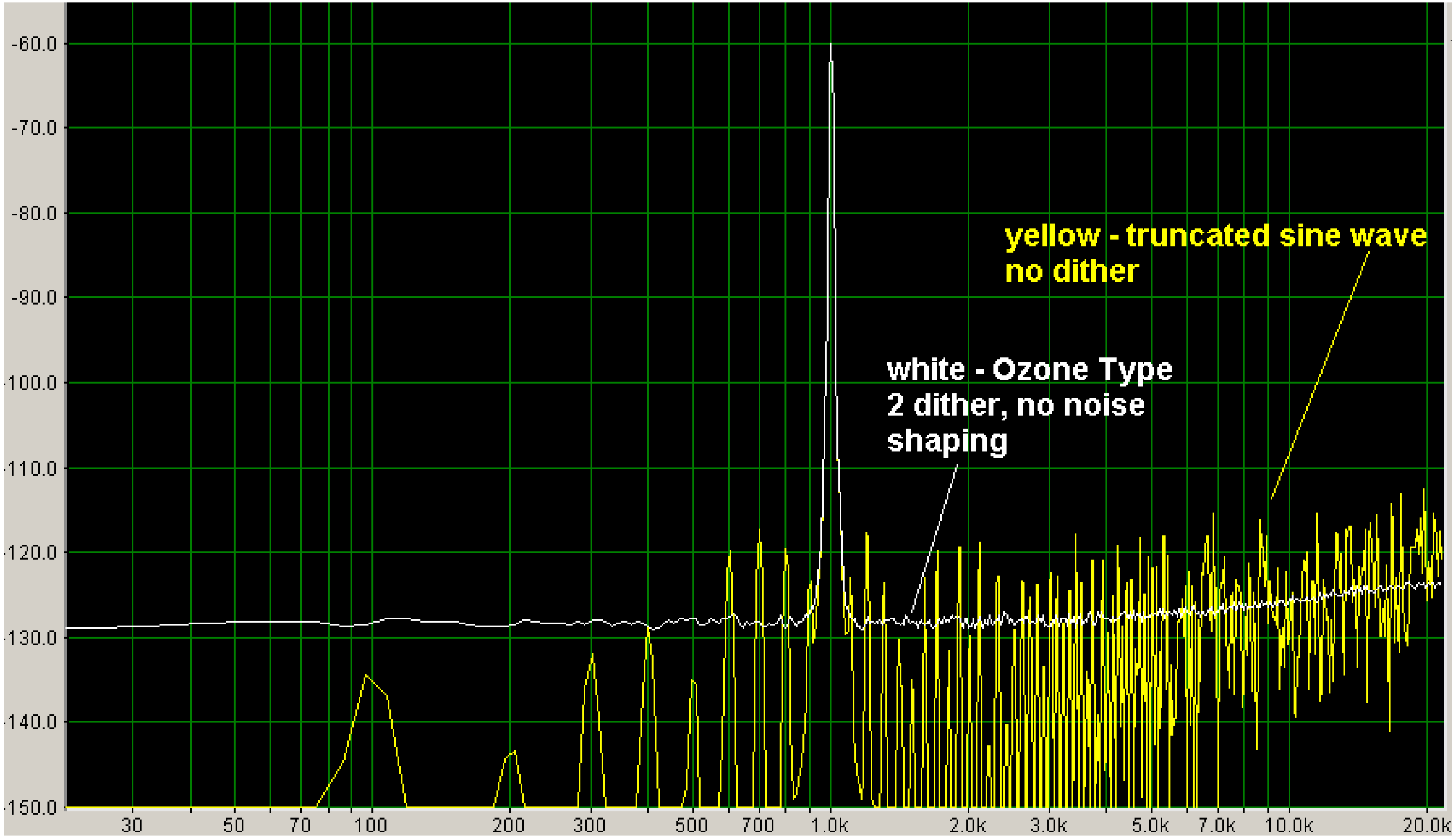
- The mastering engineer must:
  - Determine the order of tracks for albums.
  - Determine inter-track silence times.
  - Process all album tracks as needed for consistency, especially loudness and equalization.
  - Insert required metadata (e.g. PQ subcodes).
  - Supply timings and track numbers.
  - Dither 24- or 32-bit masters to 16 bits.
  - Prepare master image for replication.

Dithering for CD

- CD format is 16 bit linear, 44.1 KHz sample rate.
- For 16-bit product delivery, 32- or 24-bit masters must be dithered, not simply truncated.
- Dithering essentially *raises* the noise floor by replacing correlated noise (resulting from truncation) with less-offensive uncorrelated noise.
- Dithering algorithms are the subject of ongoing psychoacoustic research.

Truncated vs Dithered Sine Wave

- Yellow trace is truncated sine wave spectrum, white is dithered.
- Note dithered baseline is ~ -130 dB.



“Bad Dithering” Effects

- Spikes in yellow trace probably due to correlations between dither noise source and program material.

Mastering For Streaming Audio

- The largest audio content provider is YouTube.
  - More audio streams than all other providers combined, including illegal file-sharing sites.
  - 400 hours of content are uploaded to YouTube each minute (February 2017).
  - Consumes as much bandwidth as the entire Internet did in 2000.
  - 15 billion users per month. 800 million unique users.
- Mastering engineers need to keep aware of changing specs and trends for streaming services like YouTube, Spotify, Pandora, iTunes Radio, Soundcloud, Bandcamp, … and any significant new services that crop up.
- Streaming services have the potential to offer increased audio quality over current distribution media like CDs.

Volume Normalization

- Different delivery systems have different specs for loudness. Non-conforming content will be forced into conformance by automated processes that *will* reduce sound quality.
- Today's average pop music CD has less dynamic range than Edison's wax cylinders!
- Guidelines were established starting in 2011 as a response to the "loudness wars" that began in the 1990s.

Medium	Loudness mandate
ATSC (US Broadcast TV) CALM	-24 LUFS
EBU R-128	-23 LUFS
iTunes and iTunes radio	-16 LUFS
Game Audio Initiative (R-128)	-23 LUFS

Increased Bit Depth

- CD audio is resolution is frozen at 16 bits, but streaming services can provide 24 bit audio easily.
- Pro audio is generally recorded at 24 bit (integer) resolution, or 32-bit floating point.
- 24 bit audio offers increased dynamic range and avoids loss of precision from dithering 24 bit masters to 16 bits.
- Streaming media players generally handle 24-bit audio lossless formats fine today.
- Mainstream streaming audio media could surpass CD quality within a year.

Questions?