

Reference, adaptive filter, double-talk, residual suppressor: every AEC stage on one page.

The four stages, in order

1. Delay estimate	Line up reference & echo in time	20-200 ms; re-locks on route change
2. Adaptive filter	Predict echo, subtract from mic	NLMS; learns the room, ~35-45 dB ERLE
3. Double-talk detector	Freeze filter when both talk	Else it trains on your voice & diverges
4. Residual suppressor	Attenuate leftover echo	Mops up speaker distortion; can chop voice

How hard is the echo? (by device)

DEVICE	DIFFICULTY	WHY
Wired headset / earbuds	Trivial	Mic is sealed from speaker - almost no echo
Laptop, moderate volume	Easy	Short stable path; filter stays converged
Laptop, high volume	Moderate	More speaker distortion = more residual
Open-room speakerphone	Hard	Long reverberant path; double-talk clips
Bluetooth speaker / headset	Hardest	Variable loop delay breaks the estimate

Measuring it: ERLE & the room model

ERLE = echo level before cancellation - echo level after cancellation

Echo at 0 dB in, -40 dB out -> ERLE = 0 - (-40) = 40 dB (10,000x weaker).

Filter taps = echo-path duration x sample rate = 0.128 s x 48000 = ~6,144 taps.

Target: push residual echo below room noise, where the ear stops noticing it.

Remember

- Echo = your speaker's sound re-entering your mic and returning to the far end, delayed.
- The reference (what you are about to play) is the one thing the device knows - AEC subtracts a prediction built from it.
- Past ~25-30 ms one-way delay you need a real canceller, not hand-tuned attenuation (ITU-T G.131).
- Double-talk is the #1 complaint: test with BOTH people talking over each other, not polite turns.
- Echo for ~1 s after moving the laptop or plugging in headphones = the delay estimate and filter re-converging.
- Bluetooth is worst on every axis: long, distorted, and a delay that keeps moving. Recommend headphones.