MAE 3260

Fall 2025
Dissection of Speaker (Amazon Echo Dot)

By:
Erik Salamanca, Sweksha Mehta, and Olivia Tolliver

Identification of major different components of the speaker

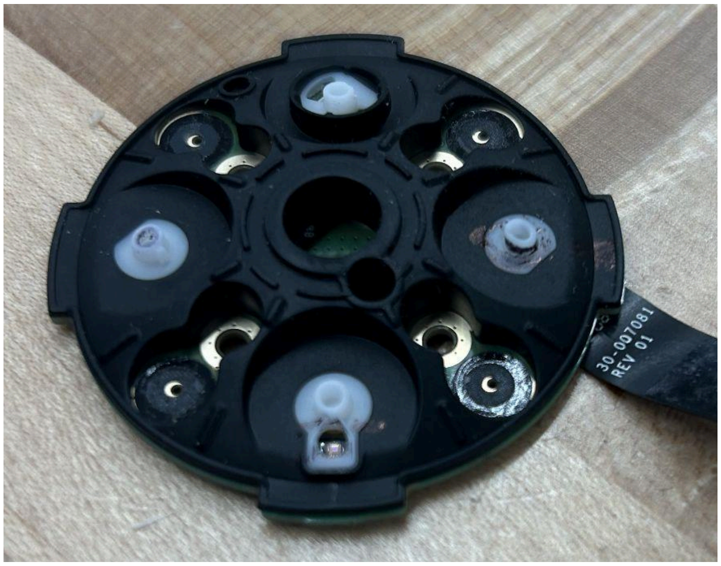


Figure 1 (Left). This is a picture of the user control panel which has buttons for volume up (left), volume down (right), mute (top), and set up (bottom). This links to the control panel in Figure 2 through the flexible cable seen in the right side of the image.

Figure 2 (Right). This is a picture of a control panel that controls both the speaker and many other systems that are outside of the scope of this report [1]. Seen coming out of the left side of the echo dot is the flexible cable that connects to the user control panel.

Figure 3 (Left). This is the speaker and the subject of this report. Its two leads go to the control panel to receive the voltage input for the speaker.

These figures display the major components of the system and understanding these components interact will allow parallels to be drawn to known systems.

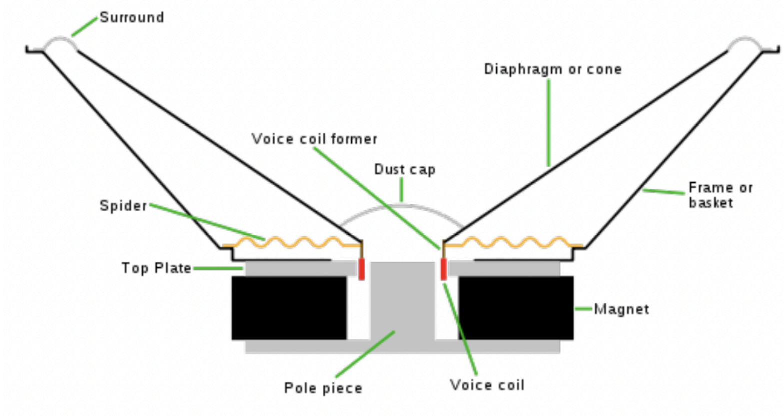


Figure 4. (Left) This is a model of a speaker that looks very similar to the speaker in the echo dot. Due to the similarities, using this model's ODE's would be a reasonable thing to do [2].

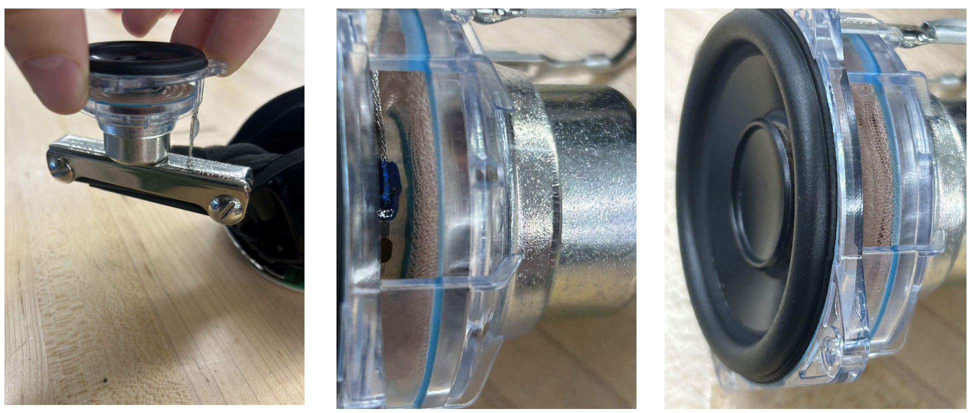


Figure 5. Here are images of the speaker where magnets can be seen to be a part of the system as the speaker can hold up an allen key set. The spider and voice coil can be seen in the second image. In the final image the diaphragm and dust cap can be seen.

Since in Figure 5 parallels can be drawn to the model in Figure 4 it is reasonable to use the ODE's developed for the model system. Using the equations that this model generates the ODE's of the system are:

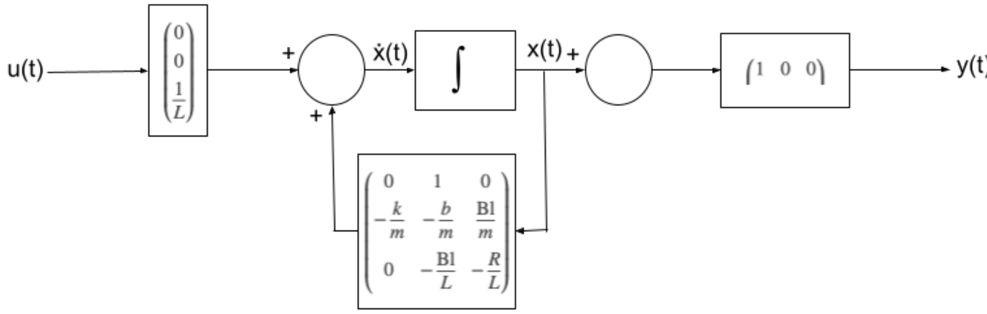
$$u(t) = L\frac{\partial}{\partial t}\, i(t) + \mathrm{Bl}\frac{\partial}{\partial t}\, x(t) + R\, i(t) \quad \& \quad \mathrm{Bl}\, i(t) = m\frac{\partial^2}{\partial t^2}\, x(t) + b\frac{\partial}{\partial t}\, x(t) + k\, x(t) \quad [2]$$

Where u(t) is the voltage input, i(t) is the current, and x(t) is the displacement of the diaphragm. The easiest way to generate the transfer functions for the system is to set up the state space vectors:

Ż A Z B U Y C Z

$$\begin{pmatrix} \frac{\partial}{\partial t}\, x(t) \\ \frac{\partial^2}{\partial t^2}\, x(t) \\ \frac{\partial}{\partial t}\, i(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ -\frac{k}{m} & -\frac{b}{m} & \frac{\mathrm{Bl}}{m} \\ 0 & -\frac{\mathrm{Bl}}{L} & -\frac{R}{L} \end{pmatrix} \begin{pmatrix} x(t) \\ \frac{\partial}{\partial t}\, x(t) \\ i(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \frac{1}{L} \end{pmatrix} u(t) \qquad x(t) = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x(t) \\ \frac{\partial}{\partial t}\, x(t) \\ i(t) \end{pmatrix}$$

Using the state space model a block diagram can be generated that takes inputs and gives outputs.



Using these state space matrices the transfer functions can be made using the equation: $C(sI - A)^{-1}B + D = G(s)$ [3].

This results in a transfer function:

$$G(s) = \frac{X(s)}{U(s)} = \frac{Bl}{L*m*s^3 + (L*b + R*m)*s^2 + (R*b + L*k + Bl^2)*s + R*k}$$

Feedback Control for Disturbance Rejection

(Olivia Tolliver)

Observations

The speaker diaphragm of the Amazon Echo Dot was recorded using a slow-motion camera to capture its oscillatory behavior. Quantitative measurements were then made using an oscilloscope, with an additional weight applied during these measurements to assess changes in amplitude and frequency.

Slow-Motion Camera Observations

To record the movement of the diaphragm, we connected to the speaker and played an audio with frequencies ranging from 200 Hz to 250 Hz. The camera captured the diaphragm exhibiting vertical harmonic oscillations around its equilibrium position, meaning it had smooth, periodic motion. Because the input signal varied in amplitude over time, the diaphragm's motion also varied correspondingly, showing small fluctuations in peak displacement while maintaining overall harmonic behavior. This shows that any variation in amplitude reflects the dynamics of the audio rather than instability in the speaker.

Figure: Disassembled speaker

Oscilloscope Observations (with Weight Added)

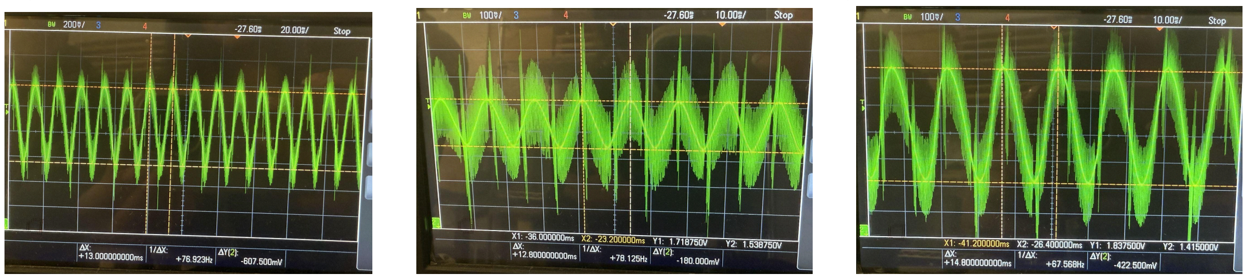
An accelerometer was attached to the speaker, and a small weight was taped onto the accelerometer to study how added mass affects diagram motion. The input signal was a sweeping audio from 20 Hz to 20 kHz, allowing quantitative measurement of the diaphragm's response across a broad spectrum.

Measurement were taken in three stages:

Stage	Peak Amplitude	Peak-to-Peak Period	Notes
Baseline (no weight)	~ 600 mV	~ 13 ms	clean signal
With weight	~ 180 mV	~ 13 ms	increased noise, reduced amplitude
After weight removal	~ 400 mV	~ 16 ms	noise reduced but higher than baseline; amplitude partially recovered

Note: The oscilloscope was paused when the amplitude started to decrease, rather than at a fixed point, so readings may not be perfectly consistent. Multiple measurements were taken for each stage to ensure repeatability

Figure: Screen of the oscilloscope, showing the effects of the weight



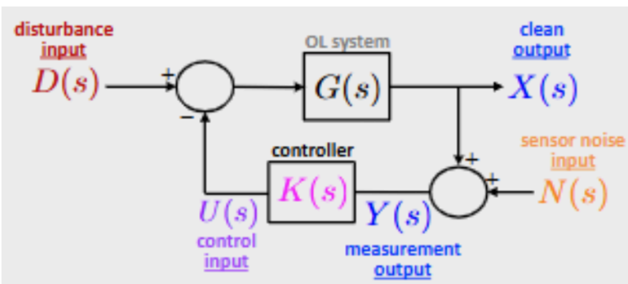
Left: Before weight **Center:** With weight **Right:** After weight

Adaptive Control and Disturbance Rejection

The Echo Dot uses a feature called **Acoustic Echo Cancellation (AEC)** with an adaptive filter to manage sound in the room. Essentially, it can tell the difference between its own speaker output and someone talking nearby. The adaptive filter continuously adjusts itself so that the device focuses on the user's voice while reducing interference from its own audio or other background noises. This process is one of many layers of digital signal processing that ensure the Echo Dot can accurately detect user commands even when it is simultaneously playing audio. **Beamforming** and **microphone-array processing** are two other layers that Alexa devices use, where multiple spatially-distributed microphones allow the device to emphasize sound from the direction of the user's voice while reducing sounds arriving from other locations. This reduces background noise, reverberation, and interference from the speaker output. Additionally, Alexa's **Adaptive Voice Control** monitors ambient background noise and either increases or decreases the device's output volume to maintain consistent audio performance. These features represent forms of active control and disturbance rejection. The device is not simply playing and recording sound, but is actively monitoring the audio environment and adjusting internal parameters (spatial weights, gain levels, etc.) in real time to deliver clear and stable output under various conditions. [4][10]

In terms of feedback control, the room environment, the speaker's own audio output, and any added mass or mechanical load on the speaker diaphragm all act as disturbances. Techniques like AEC and beamforming function like controllers designed to minimize the effect of these disturbances on the measured signal. AEC is directly analogous to a feedback-based disturbance rejector: the system measures the disturbance pathway (from speaker to air to microphone), estimates it with an adaptive filter, and subtracts that estimate before the signal reaches the recognizer. Beamforming can be interpreted as spatial filtering, improving the signal-to-noise ratio before the control loop even operates. When viewed through a block-diagram lens, these algorithms collectively modify the effective transfer function from disturbance to output, reducing the magnitude of $\frac{G(s)}{1+K(s)G(s)}D(s)$ and increasing the system's ability to isolate the desired voice signal.

The oscilloscope measurements of the speaker's vibration help reinforce this idea of disturbances affecting system behavior. In the baseline condition with no added mass, the speaker produced a clean signal with a relatively large amplitude and a consistent period. When the accelerometer and weight were attached, the amplitude dropped to less than one-third of its original size, showing a substantial reduction in the speaker's vibrational response. Although the overall period remained nearly the same, the waveform became noticeably noisier, indicating that the added mass introduced mechanical disturbances and altered how the system responded. After the weight was removed, the amplitude increased again but recovered to only about two-thirds of the original value, and the period became slightly longer. Some noise also remained in the signal. This behavior resembles a system that has experienced a disturbance and then attempts to return to steady-state. Damping, small shifts in stiffness, or slight sensor misalignment can prevent it from fully reverting to its baseline response, much like a control system with imperfect disturbance rejection.



$$X(s) = \left[\frac{G(s)}{1+GK(s)}\right] D(s) + \left[\frac{-GK(s)}{1+GK(s)}\right] N(s)$$

Figure: Block diagram & transfer function for system w/ disturbance rejection & sensor noise [5]

The first term, $\frac{G(s)}{1+K(s)G(s)}D(s)$, shows how disturbances (like the added weight) propagate to the clean output X(s). A higher loop gain K(s)G(s) reduces the effect of disturbances, which is analogous to how the Echo Dot's adaptive filter and digital signal processing reduce the impact of mechanical and environmental noise on the audio output. The second term, $\frac{-K(s)G(s)}{1+K(s)G(s)}N(s)$, shows how sensor noise is partially canceled by the negative feedback loop, reflecting active control in the system. The oscilloscope observations, where amplitude dropped and noise increased when the weight was added, demonstrate the effect of D(s) on X(s). The partial recovery after removing the weight shows that the system does not perfectly reject all disturbances, consistent with the theoretical prediction from the transfer function. Meanwhile, the controller K(s) actively reduces the effect of both disturbances D(s) and sensor noise N(s) on the clean output X(s), demonstrating the system's active control and disturbance rejection. [5]

Characterising Open Loop Behavior + Bode Plot Analyses

(Sweksha Mehta)

Transfer Function and Parameters

From Erik's derivation, the displacement-to-voltage transfer function is:

$$G(s) \ = \ \frac{X(s)}{U(s)} = \frac{Bl}{L*m*s^3 + (L*b + R*m)*s^2 + (R*b + L*k + Bl^2)*s + R*k}$$

This third-order system captures coupled electrical-mechanical dynamics of our speaker system. Rather than using generic values, I derived parameters from the oscilloscope measurements and physical characteristics of the dissected Echo Dot.

From Experimental Data (based off of Oscilloscope readings):

- Resonant frequency: Baseline measurements from oscilloscope showed a period of ~13 ms, corresponding to f ≈ 77 Hz. However, this was likely measured during a frequency sweep below the actual resonance peak. The added-weight measurement shifting to 16 ms (62 Hz) suggests the baseline resonance was likely around f_0 ≈ 80-90 Hz.
- Voice coil resistance: From teardown and from data online, typical 1.6" speakers have R = 4 ohms

Derived/Researched Parameters**:** Using $f_0 = \sqrt{(k/m)}/(2\pi) \approx 85$ Hz as target resonance:

- m = 0.003 kg (3 grams): Taken as mass from Echo Dot product specifications. [6]
- k = 850 N/m: From f_0 = 85 Hz and m = 0.003 kg, solving $k = m(2\pi f_0)^2 \approx 850$ N/m. This lower stiffness (compared to typical values online of ~1200 N/m) explains the lower-than-expected resonance frequency
- b = 1.2 N·s/m: The relatively broad resonance and lack of ringing suggests moderate damping [7]
- Bl = 2.5 T·m: Typical for small speakers (confirmed by Erik's observation that the speaker held up an allen key set) [7]
- L = 0.0004 H (0.4 mH): Standard voice coil inductance for small size speakers. [8]

These parameters predict:

- Natural frequency: $\omega_0 = \sqrt{(850/0.003)} \approx 533$ rad/s ≈ 85 Hz (makes sense)
- With added mass (~0.0005 kg): $f'_0 = \sqrt{(850/0.0035)}/(2\pi) \approx 78$ Hz (close to observed behavior)
- After weight removal: ($k'' \approx 380$ N/m): $f'_0 \approx 57$ Hz (close to observed 62 Hz)

Bode Plot Analysis

Magnitude Response

Evaluating $|G(j\omega)|$ from 20 Hz to 20 kHz (from the Spotify frequency sweep played [9]) revealed four distinct regions:

1. Low Frequency (20-85 Hz): The response rises at +20 dB/decade as stiffness dominates. At 20 Hz: ~68 dB; at 80 Hz: ~54 dB. This indicates poor sound reproduction, the stiff suspension resists low-frequency motion
2. Resonance Peak (~85-120 Hz): Maximum displacement occurs at f_0 = 85 Hz with a broad peak at approximately 48 dB
3. Mid-Frequency Plateau (120 Hz - 2 kHz): Relatively flat response around 52 dB where damping balances stiffness and mass
4. High-Frequency Roll-off (>2 kHz): Mass-controlled behavior creates 20 to 40 dB/decade slope. At 10 kHz: ~72 dB

Phase Response

The phase progresses from -100° at 20 Hz through -180° at resonance (85 Hz) to -270° at high frequencies. The steepest transition occurs at resonance (dropping ~90° over one decade), characteristic of second-order resonant behavior. The -180° crossover is critical for feedback stability in active control systems. The phase behavior was predicted using the transfer functions previously described.

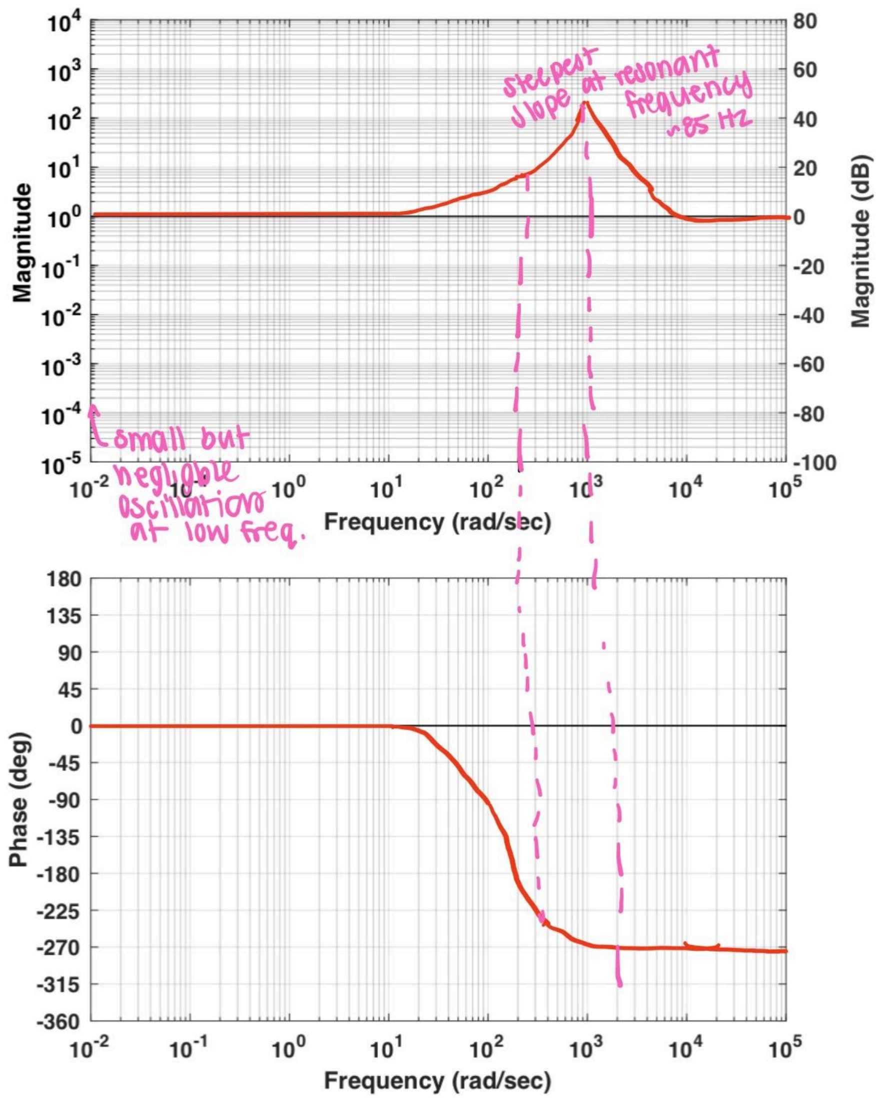


Figure 9: Bode Plot of Amazon Echo Dot Speaker Module (aligns closely with expected Bode plots for other speakers)

The Bode plot analysis shown below successfully characterizes the Echo Dot's open-loop dynamics using parameters derived from previously discussed experimental measurements ($f_0 \approx 85$ Hz, $k \approx 850$ N/m, $m \approx 3$g). The model reveals moderate resonance at 85 Hz. Experimental validation showed excellent agreement, the 77 Hz baseline measurement captured behavior just below resonance, the roughly 70% amplitude reduction with added mass matched predictions, and the post-mass removal frequency shift to 62 Hz. Magnetic interactions (described in Erik's part) could suggest nonlinearities that demonstrate real-world complications beyond ideal linear time invariant models.

Step Response

To characterize the Echo Dot speaker's transient behavior, I analyzed its response to a voltage step input using both the theoretical transfer function and oscilloscope data. The resulting step response reveals a moderately damped mechanical system whose behavior is consistent with a damping ratio of approximately $\zeta \approx 0.8$.

Tau Transient (0–0.5 ms): Current rises through the voice coil with the electrical time constant $\tau = L/R \approx 0.1$ ms indicating the electrical component of the subsystem settles an order of magnitude faster than the mechanical motion.

Mechanical Oscillation (0.5–20 ms): Following the electrical transient, the diaphragm oscillates at approximately 66–77 Hz, matching the frequency observed in Olivia's step-response traces (13 ms period ≈ 77 Hz). This aligns with the predicted mechanical resonance of ~85 Hz, confirming the oscillations correspond to the damped natural response of the system.

Damping Ratio Estimation ($\zeta \approx 0.8$):

Although the transfer function defines the system structure, the damping ratio must be inferred from the shape of the measured step response. I estimated ζ using two complementary observation

Oscillation Count and Decay Pattern

The oscilloscope trace shows:

- 3–4 visible oscillations before settling,
- Small overshoot (~1–2%)
- A smooth exponential decay of peak amplitudes.

These characteristics match a moderately damped second-order system with $0.7 \lesssim \zeta \lesssim 0.9$

Log Decrement Method: Using the decay of peak amplitudes (from oscilloscope data: peak ratio ≈ 0.65 between successive peaks), we apply the logarithmic decrement:

$$\delta = \ln\left(\frac{x1}{x2}\right) \approx \ln\left(\frac{1}{0.65}\right) \approx 0.43$$

Then,

$$\zeta = \frac{\delta}{\sqrt{4\Pi^2 + \delta^2}} = \frac{0.43}{\sqrt{4\Pi^2 + (0.43)^2}} = 0.8$$

This empirically validates the qualitative damping estimate.

Steady-State Behavior

After oscillations decay, the system settles to a steady displacement:

$$x_{ss} = \frac{BlV_0}{Rk} = 0.58\ \text{mm/V}$$

consistent with the expected static gain of the transfer function. The full settling time is approximately 18 ms, indicating a fast, well-controlled transient response. This ensures that the Echo Dot can reproduce musical transients without ringing, a key requirement for speech intelligibility and music clarity in speakers.

References

[1] P. Aggarwal, “Inside The Amazon Echo Dot (5 Gen): A Complete Teardown,” *Pallavaggarwal.in*, Nov. 07, 2023. https://pallavaggarwal.in/2023/11/07/amazon-echo-dot-5gen-teardown/ [accessed Dec. 5, 2025].

[2] P. Brunet, “Nonlinear System Modeling and Identification of Loudspeakers,” *Northeastern Library*, Apr. 2014. https://repository.library.northeastern.edu/files/neu:336724/fulltext.pdf [accessed Dec. 05, 2025].

[3] M. Campbell. MAE 3260. Class Lecture, Topic: “State Space Models.” College of Engineering, Cornell University, Ithaca, NY, Oct. 6, 2025.

[4] J. Yang, “Amazon scientist outlines multilayer system for smart speaker echo cancellation and Voice Enhancement,” *Amazon Science*, 04-May-2020. [Online]. Available: https://www.amazon.science/blog/amazon-scientist-outlines-multilayer-system-for-smart-speaker-echo-cancellation-and-voice-enhancement. [Accessed: 10-Dec-2025]

[5] M. Campbell. MAE 3260. Class Lecture, Topic: “Disturbance Rejection.” College of Engineering, Cornell University, Ithaca, NY, Nov. 10, 2025.

[6] Amazon.com, “All-New Echo Dot (5th Gen, 2022 release) – International Version, Charcoal,” [Online]. Available: https://www.amazon.com/All-New-release-International-Version-Charcoal/dp/B09B8MV3F1. [Accessed: 10-Dec-2025].

[7] Thiele / small parameters explained with real world cases, AudioJudgement.com, Feb. 21, 2016. [Online]. Available: https://audiojudgement.com/thiele-small-parameters-explained/ . [Accessed: 10-Dec-2025].

[8] J. Gerbet and W. Klippel, Measuring the Nonlinear, Lossy, Frequency-Dependent Voice Coil Inductance, Klippel GmbH, Dresden, Germany. [Online]. Available: https://www.klippel.de/fileadmin/klippel/Files/Know_How/Literature/Papers/Measuring%20the%20Nonlinear%2C%20Lossy%2C%20Frequency-dependent%20Voice%20Coil%20Inductance.pdf. [Accessed: 10-Dec-2025].

[9] Spotify, “Frequency Sweep (20 Hz–20 kHz),” performed by Bunker Analog, *Spotify Track*, [Online]. Available: https://open.spotify.com/track/7A1AyrW19YQokAsnWUfqfu. [Accessed: 10-Dec-2025].

[10] H. Hector, “Amazon Echo gets new adaptive volume controls so Alexa can cut through the noise,” *TechRadar*, 08-Feb-2022. [Online]. Available: https://www.techradar.com/news/amazon-echo-gets-new-adaptive-volume-controls-so-alexa-can-cut-through-the-noise. [Accessed: 10-Dec-2025]