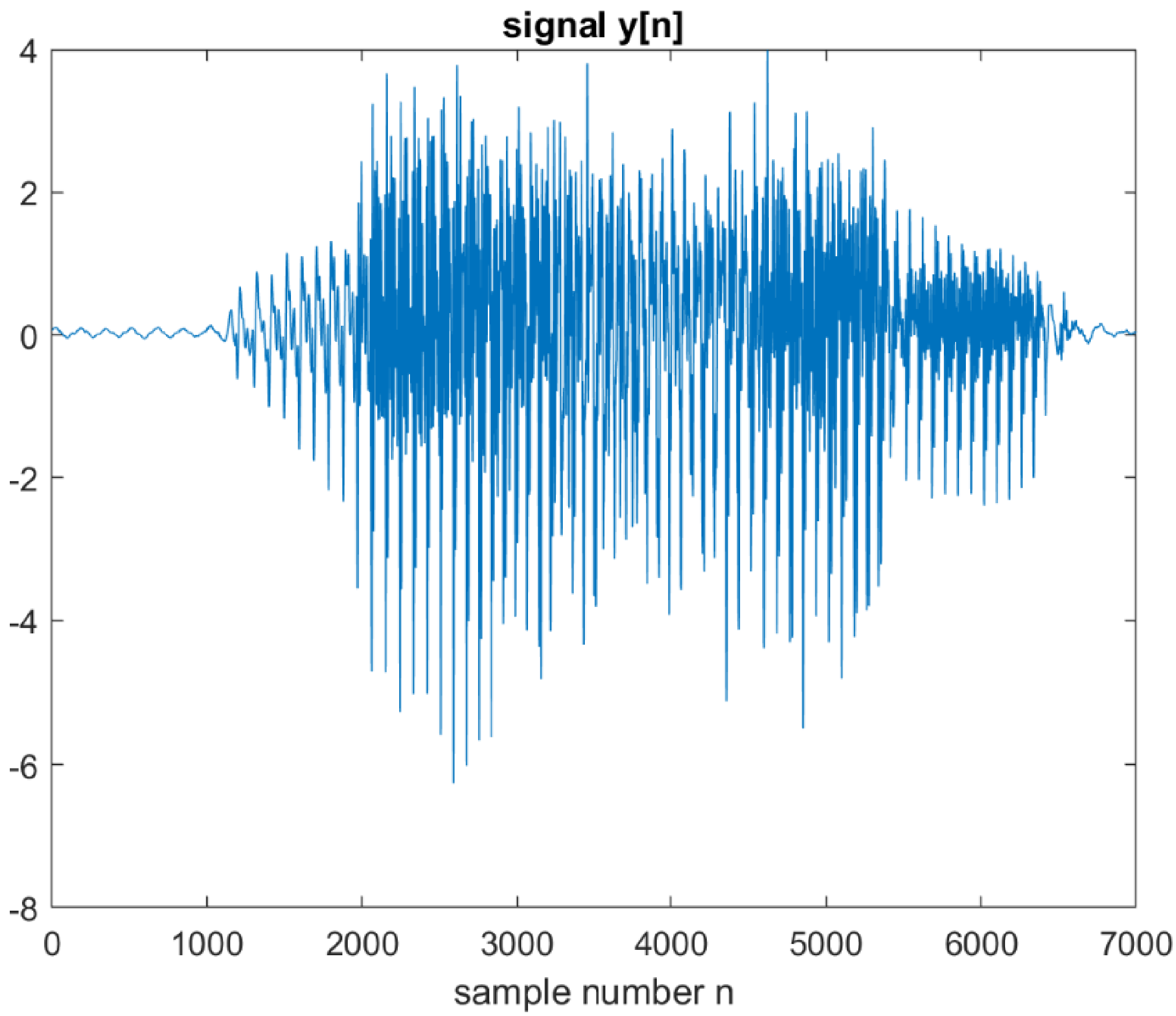
Julia Frothingham
EGR 320, Voss
Fall 2021

Lab 1 - Echo Cancellation

1. Listen to the speech file. Describe what you hear. Make a plot of y[n]. Given the sampling rate of 8192 Hz, what is the duration in seconds of the audio signal? In equation 1 y[n] = x[n]+ax[n-N], what do each of the variables represent? Explain how equation 1 models the signal of speech with an echo.

I hear a distorted audio clip of someone saying the words "line up," but it sounds unclear, like there's audio layered on top of it.

Figure 1. The original signal y[n] with an echo.



Sampling rate = 8192 Hz = 8192 samples/second
(8192 samples/second)(1/7000 samples) = **1.17 seconds duration**

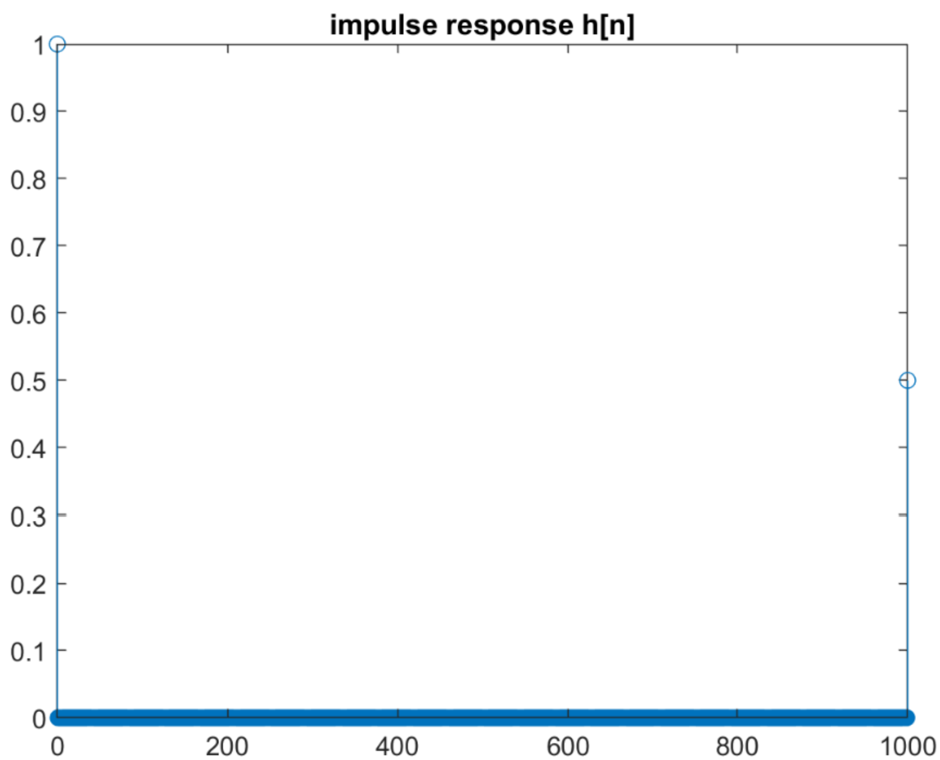
In Equation 1 y[n] = x[n]+ax[n-N], y[n] represents the signal as recorded, with an echo present and overlapping the original input. x[n] represents this original input, the sound with no echo. ax[n-N] represents the echo. Its amplitude is larger or smaller than the amplitude of the input by a factor of a, and it is shifted to the right by a factor of N, but otherwise has the same shape as x[n].

Speech with an echo consists of the original sound followed by the same sound repeated at a different volume a certain amount of time after the original sound. This equation models that. The constant N models how soon after a sound is made the echo begins, and the constant a models the change in volume of the echo compared to the original sound.

2. Determine and plot the impulse response of the echo system. Store this impulse response in the vector h[n] for 0 <= n <= 1000 and plot h[n]. Explain your plot of h[n], including why it makes sense and how it is consistent with Equation 1.

In this case, N=1000 and a=0.5. So Equation 1 has the form y[n] = x[n]+0.5x[n-1000]. The impulse response h[n] is the output of the system described by Equation 1 for an input equal to the impulse function. Letting d[n] be the unit impulse, the impulse response would be:
h[n] = d[n] - 0.5d[n-100].

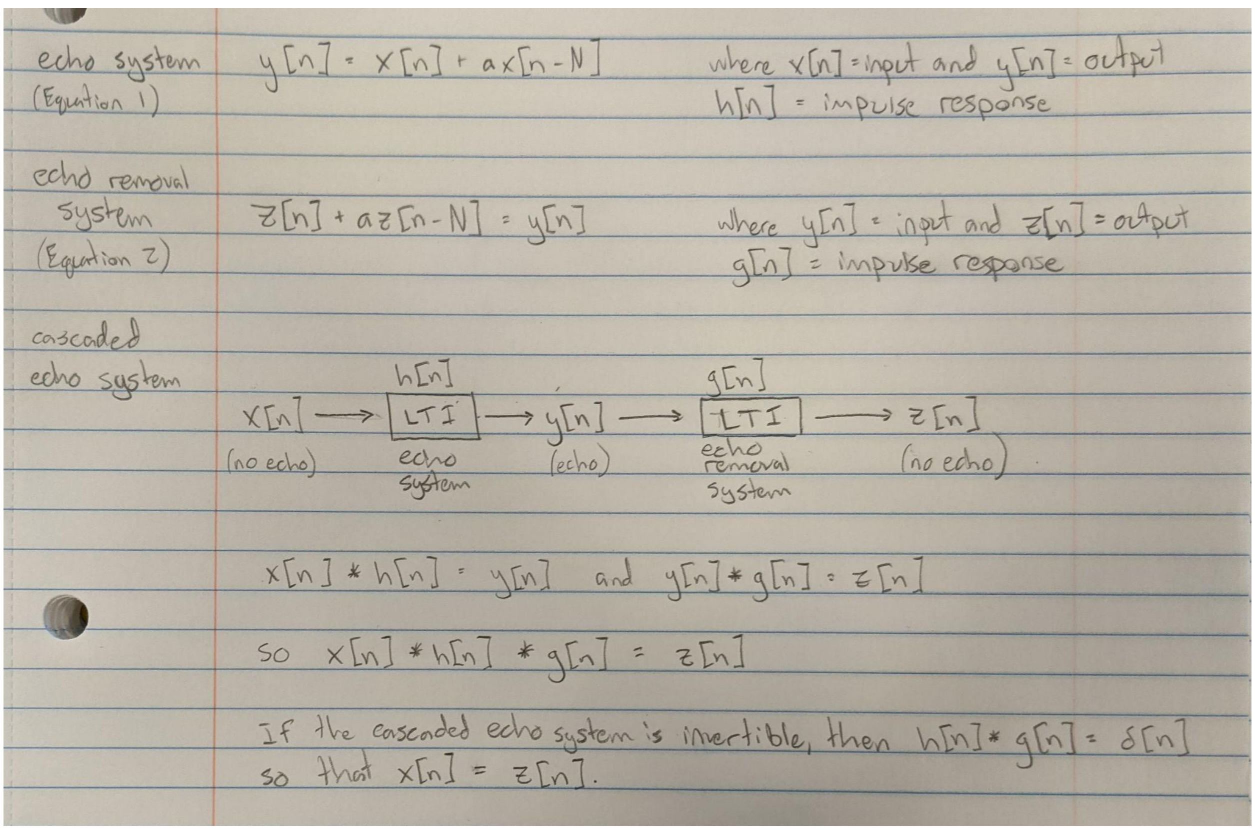
Figure 2. The impulse response h[n] of the echo system.



The plot of h[n] shows the impulse function and then a second impulse function scaled by 0.5 and shifted to the right by 1000. This makes sense and is consistent with Equation 1 because if the original input to the system was an impulse, the response of the system would be the original input followed by a signal of the same shape but shifted to the right by 1000 and with an amplitude multiplied by 0.5. The impulse response includes the input and the input's echo.

3. Consider an echo removal system described by the difference equation z[n] + az[n - N] = y[n], where y[n] is the input and z[n] is the output which has the echo removed. Draw a diagram that relates equations 1 and 2 through x[n], y[n], z[n], h[n], and g[n]. In terms of your diagram, describe what is meant by an inverse system and describe how the echo could be removed in terms of h[n] convolved with g[n].

Figure 3. Relation between echo system and echo removal system.

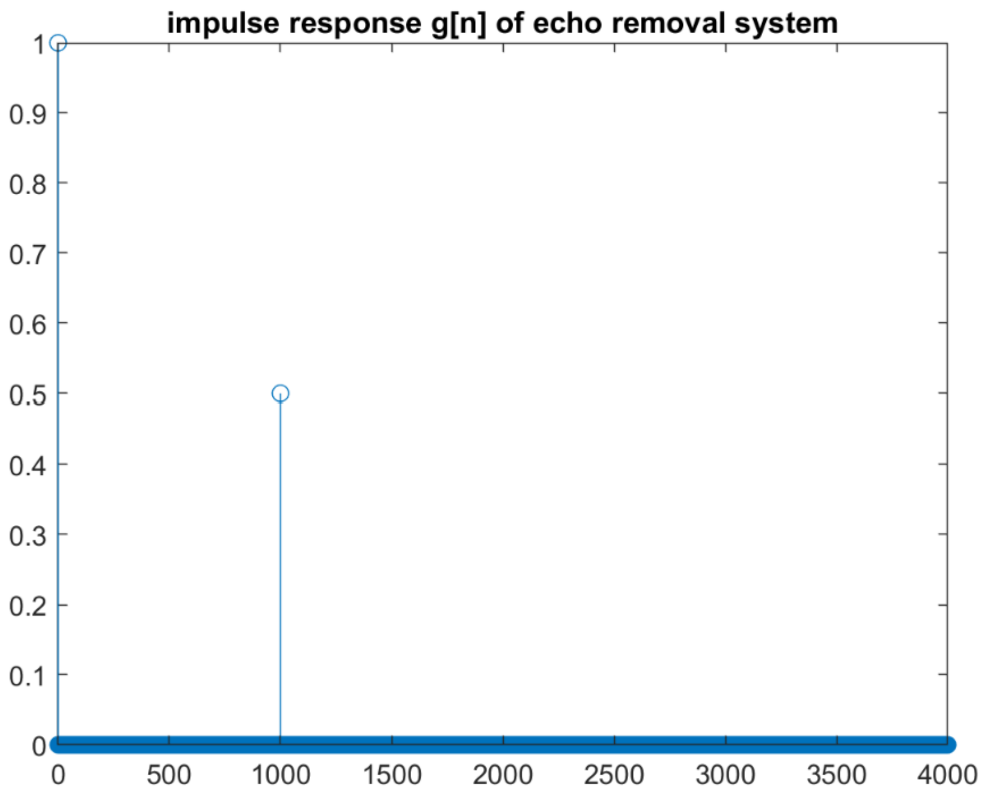


An inverse system is one that takes in the output of a previous system and returns the input of that previous system as its output. In the diagram above, the echo removal system is the inverse system to the echo system. If h[n] convolved with g[n] equals the unit impulse, then the statement x[n] convolved with h[n] convolved with g[n] equals z[n] becomes x[n] convolved with the unit impulse equals z[n]. Because the unit impulse is the identity function for convolution operations, this would mean that x[n] was equal to z[n]. So the output of the echo removal system would successfully be the signal with no echo.

4. Determine and plot the impulse response g[n] of the echo removal system in Equation 2. Describe your plot.

Let d[n] be the unit impulse.
g[n] = d[n] + 0.5d[n-1000]

Figure 4. Impulse response g[n] of the echo removal system.

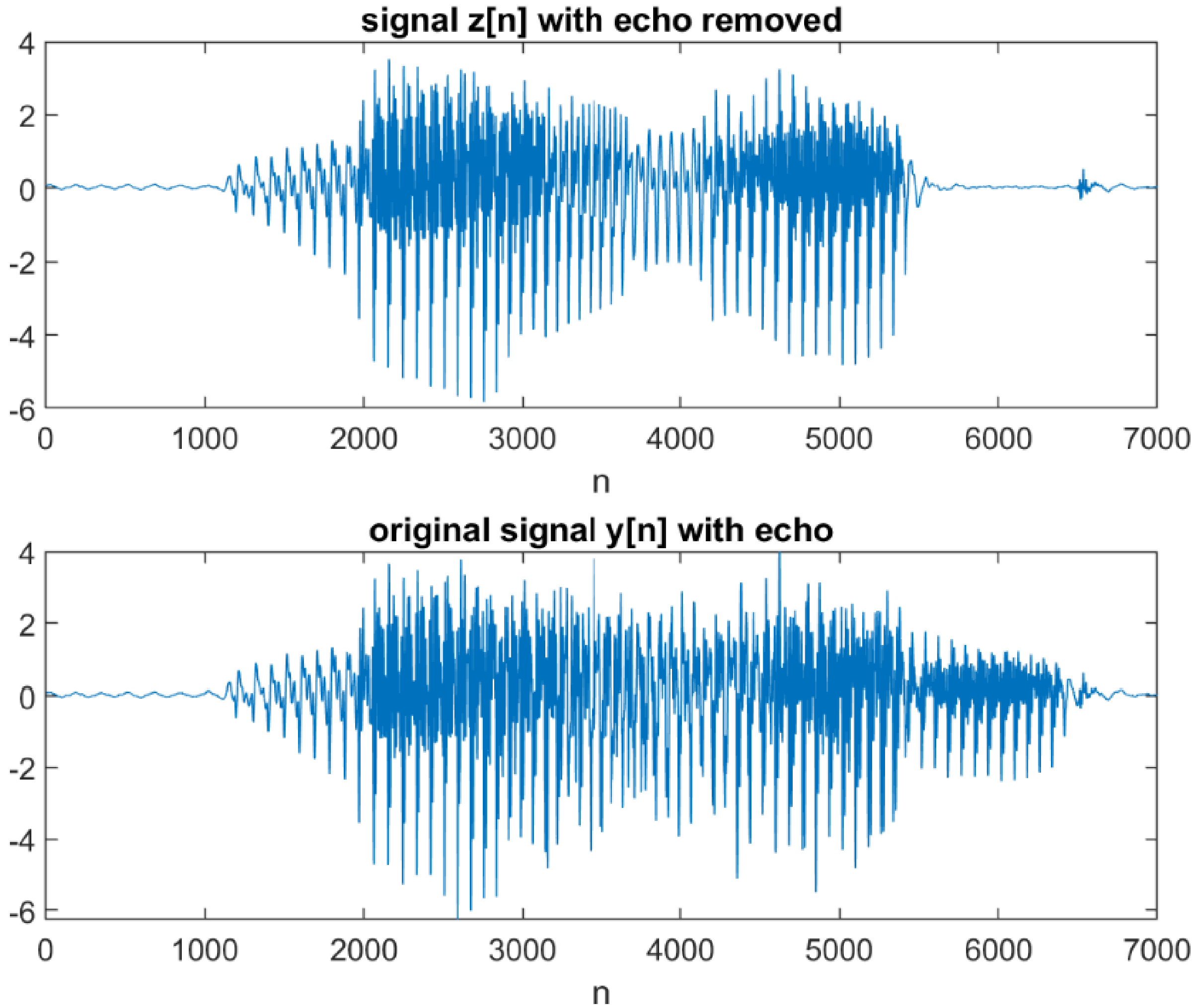


The plot shows the impulse response of the echo removal system. It consists of the unit impulse and then the unit impulse shifted to the right by 1000 and scaled by 0.5.

5. Implement the echo removal system. Use y[n] as an input to equation 2 described by g[n] and listen to the output z[n]. Plot and listen to the output z[n]. Do you still hear an echo? Make adjacent (upper and lower) plots that compare the signals y[n] and z[n] and describe how the two differ and why. What do you notice about the echo?

When listening to z[n], I hear a voice saying the words “line up.” It is much clearer than the signal y[n], and the sound cuts off neatly after the “p” in “up” rather than continuing. I do not still hear an echo.

Figure 5. Comparison of signals y[n] and z[n].

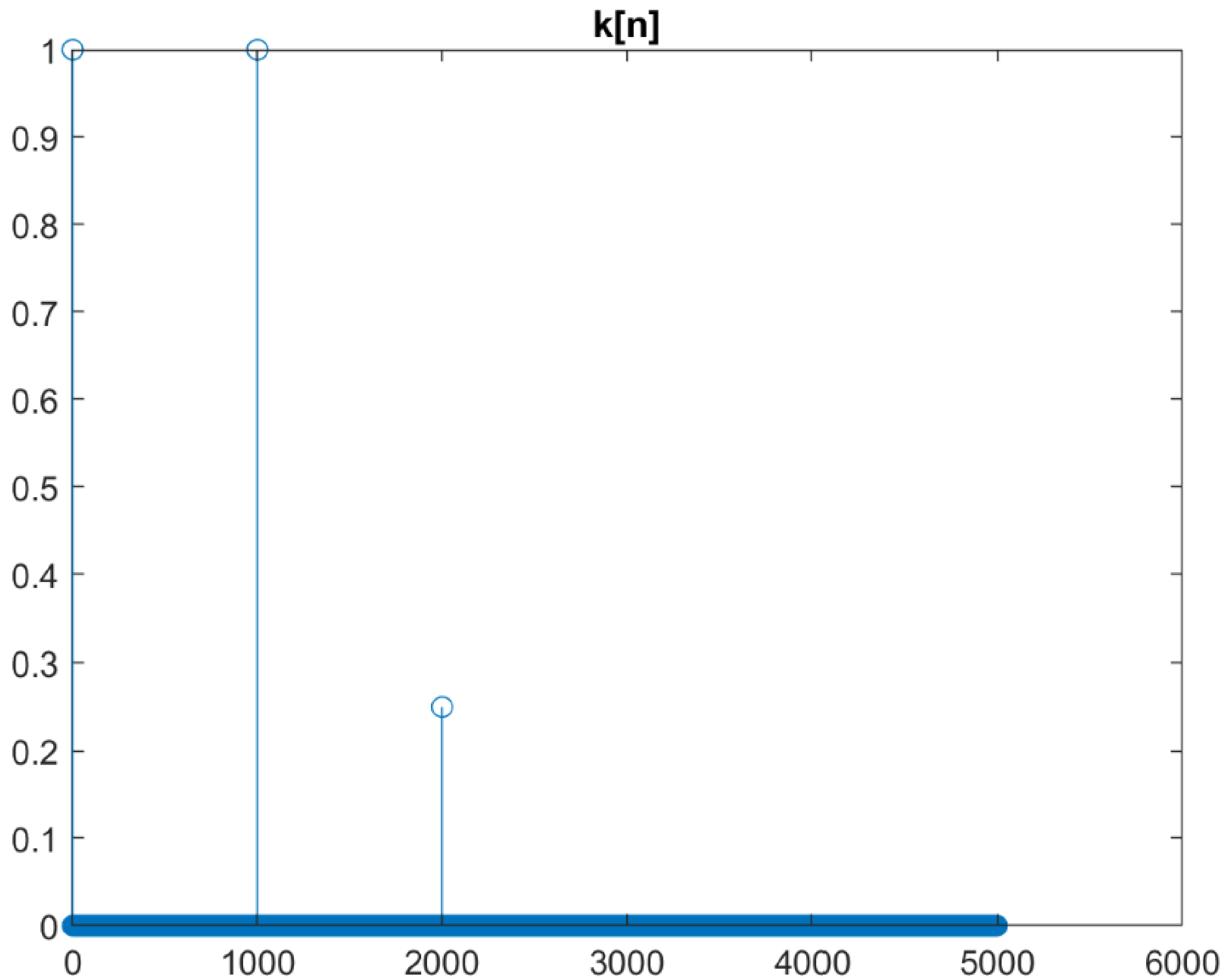


Both plots show the same start to the soundwave. However, z[n] displays two clear regions of greater amplitude corresponding to each of the words said - “line” and “up” - with a pause in between them. Because of its echo, y[n] does not have such clear regions - the repeated echo of the phrase begins before the original phrase is finished, which makes the overall audio sound muddy, and the total duration of the soundwave is longer than the original signal. The overall amplitude of the signal y[n] is also greater, as the echo has an additive effect on the audio.

6. Calculate in Matlab the overall impulse response of the cascaded echo system: k[n] = h[n] convolved with g[n]. Why might you expect k[n[to be a unit impulse? Why is k[n] not a unit impulse? Is the system g[n] truly an inverse system? Why or why not?

I would expect k[n] to be a unit impulse because if it was a unit impulse, then z[n] would equal x[n], with no echo present. I could not hear an echo in the playback of z[n], so I would have thought that the echo was successfully removed and that z[n] did equal x[n].

Figure 6. Overall impulse response k[n] of cascaded echo system.



It turns out that k[n] is not a unit impulse. It is a unit impulse followed by a unit impulse shifted to the right by 1000 followed by another impulse scaled by between 0.2 and 0.3 and shifted to the right by 2000.

It could be that k[n] is not a unit impulse because the original system described by Equation 1 was oversimplified or slightly off in its coefficients. So for a system that is described perfectly by Equation 1, the system described by Equation 2 would be an inverse system and k[n] would be the unit impulse.

Since k[n] is not a unit impulse, this is not an inverse system - its output does not equal the input of the system that feeds into it. Instead, it's an approximation. Listening to z[n], it's a pretty good approximation - the echo present in y[n] is not noticeable.