Z-transforms : Part I

1 Relation to Discrete-Time Fourier Transform

Consider the following discrete system, written three different ways:

$$
\begin{aligned}
y[n] &= b_{-1}y[n+1] + b_1 y[n-1] + a_{-1}x[n+1] + a_0 x[n] + a_2 x[n-2] \\[2mm]
Y(z) &= b_{-1}zY(z) + b_1 z^{-1}Y(z) + a_{-1}zX(z) + a_0 X(z) + a_2 z^{-2}X(z) \\[2mm]
H(z) &= \frac{Y(z)}{X(z)} = \frac{a_{-1}z + a_0 + a_2 z^{-2}}{-b_{-1}z + 1 - b_1 z^{-1}}
\end{aligned}
\tag{1}
$$

Simple substitution finds the Z-transform for a discrete system represented by a linear constant coefficient difference equation (LCCDE). Simply replace $y[n]$ with $Y(z)$, $x[n]$ with $X(z)$, and shifts of n_0 with multiplication by z^{n_0}. That's almost all there is to it.

Set

$$z = re^{j\hat{\omega}}$$

and let $r = 1$ for the moment. Then a shift in time by n_0 becomes a multiplication in the Z-domain by $e^{j\omega n_0}$. This should look familiar given what you know about Fourier analysis. Now here's the formula for the Z-transform shown next to the discrete-time Fourier transform of $x[n]$:

$$
\begin{aligned}
Z\text{-transform}: \quad X(z) &= \sum_{n=-\infty}^{\infty} x[n]z^{-n} \\[2mm]
\text{DTFT}: \quad X(e^{j\hat{\omega}}) &= \sum_{n=-\infty}^{\infty} x[n]e^{-j\hat{\omega}n},
\end{aligned}
$$

where we have used the notation $X(e^{j\hat{\omega}})$ instead of the equivalent $X(\hat{\omega})$, to emphasize similarity with the Z-transform . Substituting $z = re^{j\hat{\omega}}$ in the Z-transform ,

$$X(z) = \sum_{n=-\infty}^{\infty} (x[n]r^{-n})e^{-j\hat{\omega}n},$$

reveals that the Z-transform is just the DTFT of $x[n]r^{-n}$. If you know what a Laplace transform is, $X(s)$, then you will recognize a similarity between it and the Z-transform in that the Laplace transform is the Fourier transform of $x(t)e^{-\sigma t}$. Hence the Z-transform generalizes the DTFT

in the same way that the Laplace transform generalizes the Fourier transform. Whereas the Laplace transform is used widely for continuous systems, the Z-transform is used widely in design and analysis of discrete systems.

You may have noticed that in class we've already sneaked in use of the "Z-plane" in talking about the "unit circle." The "Z-plane" contains all values of z, whereas the unit circle contains only $z = e^{j\hat{\omega}}$. The Z-transform might exist anywhere in the Z-plane; the DTFT can only exist on the unit circle. One loop around the unit circle is one period of the DTFT.

(For those who know Laplace transforms, which are plotted in the S-plane, there is a nonlinear mapping between the S-plane and the Z-plane. The $Im\{s\}$ axis maps onto the unit circle, the left half plane ($\sigma < 0$) maps inside the unit circle, and the right half plane ($\sigma > 0$) maps outside the unit circle. Poles and zeros in the Laplace transform are correspondingly mapped to poles and zeros in the Z-transform .)

2   Discrete transfer function

Equation (1) gave the *transfer function*, $H(z)$, for a particular system. In general the form for LCCDE systems is

$$\sum_{l=-N}^{N} a_l y[n-l] \;=\; \sum_{k=-M}^{M} b_k x[n-k]$$

which transforms to

$$\sum_{l=-N}^{N} a_l z^{-l} Y(z) \;=\; \sum_{k=-M}^{M} b_k z^{-k} X(z)$$

$$H(z) \;=\; \frac{Y(z)}{X(z)} \;=\; \frac{\displaystyle\sum_{k=-M}^{M} b_k z^{-k}}{\displaystyle\sum_{l=-N}^{N} a_l z^{-l}}$$

Note that the limits M, N do not impose symmetry on the number of terms in either direction. We could have $b_{-1} \neq b_1$, etc. One could also have $N = 0$ with M finite, yielding an FIR system, or $N = 0$ and $M = \infty$ yielding an IIR system, or an IIR system can be created with $N > 0$ and $M = 0$. There are many possibilities, but in each case the transfer function $H(z)$ completely characterizes the linear shift invariant system. The following relationships hold:

$$y[n] \;=\; x[n] * h[n]$$
$$Y(z) \;=\; X(z)H(z)$$

We can learn to look at $H(z)$ and infer properties of the system. Is it low-pass, single-pole, FIR, IIR, stable, etc.? In order to do this, we need to consider its poles, zeros, and (like the Laplace transform) regions of convergence.
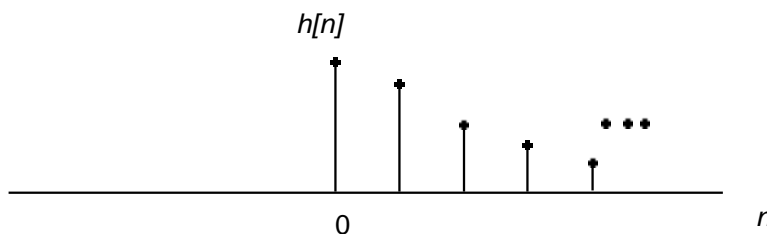
3   Poles, Zeros, Regions of Convergence

Example 1: Infinite-length right-sided

Consider the causal sequence:
$$h_1[n] = a^n u[n]$$

This impulse response might be used to model signals which decay over time, like echos.

h[n]

0   n

Consider its Z-transform :

$$
\begin{aligned}
H_1(z) &= \sum_{n=-\infty}^{\infty} a^n u[n] z^{-n} \\
&= \sum_{n=0}^{\infty} a^n z^{-n} \\
&= \sum_{n=0}^{\infty} (az^{-1})^n.
\end{aligned}
\tag{2}
$$

Remember the following useful result from mathematics of series summations:

$$\sum_{n=0}^{\infty} \gamma^n = \frac{1}{1-\gamma}, \quad |\gamma| < 1.$$

Notice this result is *only* true when $|\gamma| < 1$. Forgetting this constraint is dangerous. When $|\gamma| \geq 1$ then the series does not converge, and we say "it does not exist" or "it blows up." To proceed with (2) we must therefore have $|az^{-1}| < 1$, or $|z| > |a|$. This completes the result:

$$a^n u[n] \quad \overset{Z}{\longleftrightarrow} \quad \frac{1}{1 - az^{-1}}, \quad |z| > |a|.$$

The region in the Z-plane where the transform exists is called the *region of convergence* (ROC), e.g., $|z| > |a|$ for this example. Even though we can evaluate $\frac{1}{1-az^{-1}}$ for $|z| \leq |a|$, it is usually inappropriate to do so since this expression for the Z-transform is only defined within its ROC.

Let's draw a picture of this. First we need to think about the poles and zeros of $H_1(z)$. Poles are the values of z for which the Z-transform is ∞ and zeros are the values of z for which

$$|z| > |a|$$
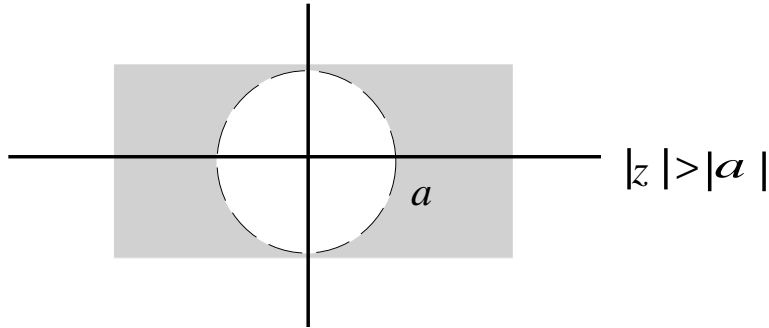
the Z-transform is 0. Every Z-transform has the same number of poles as it has of zeros. It is good practice to count them up when you think you've found them all. For the example above there is one pole and one zero, and the ROC is outside the pole:

$$H_1(z) = \frac{1}{1 - az^{-1}}, \quad |z| > |a|. \qquad \text{pole: } z = a$$
$$\text{zero: } z = 0$$

The zero at $z = 0$, outside the ROC, comes only from the expression $\frac{1}{1-az^{-1}}$ without considering its ROC. If you are troubled that there is a finite value (namely zero) found outside the ROC then you are currently in good company – because the staff is still troubled by this as well: how can the function be said to be zero outside its ROC when it doesn't even exist outside its ROC? Note we do not have trouble saying there is a pole outside the ROC, because a pole is where the function "blows up" which is just what you'd expect when you're not in the ROC. Nonetheless, all the standard texts put that zero there, outside the ROC. We will keep you posted if we convince the authors of the standard texts not to do that, or if they convince us not to worry about it.

Does the DTFT exist for $h_1[n]$? For the Laplace transform, the Fourier transform existed if the ROC included the $j\omega$ axis. For the Z-transform the DTFT exists if the ROC includes the unit circle. If this is true, we say the system is "stable," i.e., $\sum_{n=-\infty}^{\infty} |h[n]| < \infty$, or the impulse response is "absolutely summable." For $h_1[n]$, we see that the DTFT exists if $|a| < 1$.

Example 2: Infinite-length left-sided

Consider the anti-causal sequence obtained by looking at the previous example through a mirror:

$$
\begin{aligned}
h_2[n] &= -a^n u[-n - 1] \\
H_2(z) &= -\sum_{n=-\infty}^{\infty} a^n u[-n - 1] z^{-n} \\
&= -\sum_{n=-\infty}^{-1} a^n z^{-n} \\
&= -\sum_{p=1}^{\infty} a^{-p} z^{p}
\end{aligned}
$$

4

$$|z| < |a|$$
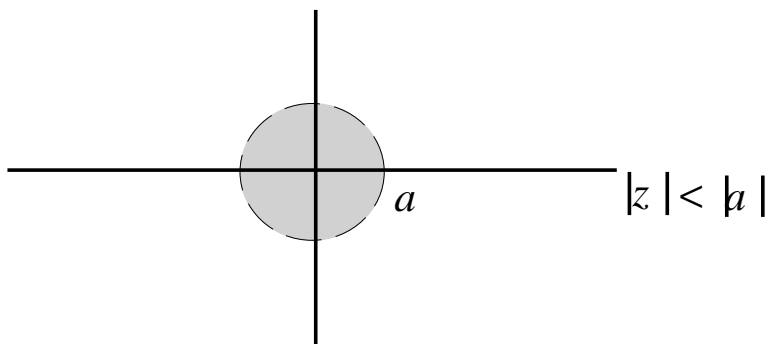
$$
\begin{aligned}
&= 1 - \sum_{p=0}^{\infty} (a^{-1}z)^p \\
&= 1 - \frac{1}{1 - a^{-1}z}, \quad |a^{-1}z| < 1 \\
&= \frac{1 - az^{-1}}{1 - az^{-1}} - \frac{-az^{-1}}{1 - az^{-1}}, \quad |z| < |a| \\
&= \frac{1}{1 - az^{-1}}, \quad |z| < |a|.
\end{aligned}
$$

Simplifying the last step gives:

$$
-a^n u[-n-1] \quad \xleftrightarrow{Z} \quad \frac{1}{1 - az^{-1}}, \quad |z| < |a|.
$$

The only difference between this Z-transform and the one in (2) is the ROC. For the anti-causal case we have the same poles and zeros, but the picture is shaded *inside* the pole:

$$
H_2(z) = \frac{1}{1 - az^{-1}}, \quad |z| < |a|.
\qquad
\begin{array}{l}
\text{pole: } z = a \\
\text{zero: } z = 0
\end{array}
$$

In general, an infinite single-pole sequence will have an ROC *inside* the pole if it is left-sided, and *outside* the pole if it is right-sided. Looking at the ROC for $h_2[n]$ tells us its DTFT will only exist if $|a| > 1$.

What is the effect of time-shifting on the Z-transform ? Consider shifting the right-sided $h_1[n]$ right by one sample, $\hat{h}_1[n] = a^{n-1}u[n-1]$, so that it only contains nonzero terms when $n \geq 1$. The Z-transform is

$$
\hat{H}_1(z) = \frac{z^{-1}}{1 - az^{-1}}, \quad |z| > |a|
\qquad
\begin{array}{l}
\text{pole: } z = a \\
\text{zero: } z = \infty.
\end{array}
$$

Shifting the sequence has moved the zero out to infinity, but has not affected the pole. Furthermore, all subsequent shifts to the right will result in additional zeros at infinity (and corresponding poles at the origin.)

The zeros do not affect the ROC. Intuitively this is pleasing – shifting does not affect convergence of the sequence.

Example 3: Infinite-length two-sided.

Let's add the two one-sided examples above, allowing the possibility for them to have different amplitudes:

$$h_3[n] = -b^n u[-n-1] + a^n u[n]$$

Everything we've been doing in these notes is linear so the Z-transforms also add. However, to satisfy both ROC's requires taking their intersection. If $a = b$ then the intersection is the empty set – so the new Z-transform has no ROC, i.e., $H_3(z)$ doesn't exist. The picture is a sequence decaying on one side of the origin and blowing up on the other. The only hope of the Z-transform converging is if $a < b$, i.e., the ROC: $|a| < |z| < |b|$.

Example 4: Finite length Consider a finite-length sequence:

$$\begin{aligned} h_4[n] &= \delta[n] + a_1\delta[n-1] + a_2\delta[n-2] \\ H_4(z) &= 1 + a_1 z^{-1} + a_2 z^{-2} \end{aligned}$$

Where is the ROC? If $z = 0$ we have division by zero, so we see there is a pole at $z = 0$. The zeros can be found by solving the roots of the quadratic. There will be two solutions, hence two zeros. Since the number of poles must be the same as the number of zeros, we need to find a second pole. This will be at $z = 0$ also.

Because we can factor any polynomial in z, we can always find its poles and zeros, e.g., $1 + 4z^{-2} = (1 + 2z^{-1})(1 - 2z^{-1})$ has zeros: $z = 2, -2$ and a "double pole" at $z = 0$. In fact, every finite length sequence will converge, so it will not have any poles other than possibly at $z = 0$ or $z = \infty$. It will only have poles at $z = 0$ if it contains negative powers of z (causal terms) and it will only have poles at $z = \infty$ if it contains non-causal terms. The ROC will exist everywhere except possibly at $z = 0, \infty$. The ROC for any FIR filter is $0 < |z| < \infty$.

Four cases have been introduced above, corresponding to four cases of ROC's. These four cases are summarized with the properties of an ROC:

Properties of an ROC:

1. Bounded by circles (dependent only on $|z|$).

2. Connected and bounded by poles or by ∞. (It can't contain a pole!)

3. If $h[n]$ is absolutely summable, then the ROC contains the unit circle, the system has a DTFT and is said to be "stable."

$$|a| < |z| < |b|$$
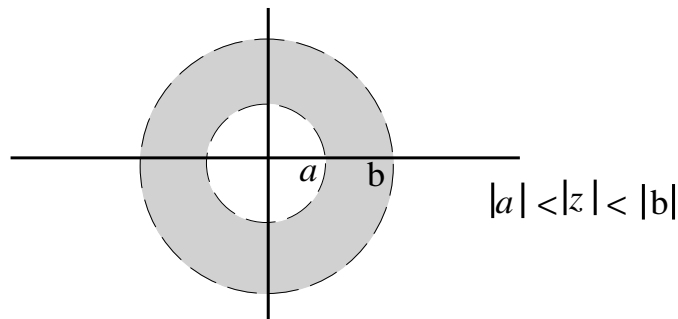
4. A stable and causal sequence has all its poles inside the unit circle.

5. Right-sided infinite sequence: ROC lies outside the outermost pole, $R_- < |z| < \infty$. If also causal, then the ROC includes $|z| = \infty$.

6. Left-sided infinite sequence: ROC lies inside the innermost pole, $0 < |z| < R_+$. If also anti-causal, then the ROC includes $|z| = 0$.

7. Two-sided infinite sequence: ROC lies between the two poles, $R_- < |z| < R_+$. The ROC does not include $|z| = 0$ or $|z| = \infty$.

8. Finite-length sequence: Always converges. ROC always includes $0 < |z| < \infty$. ROC includes 0 if sequence is anti-causal, or includes ∞ if causal.

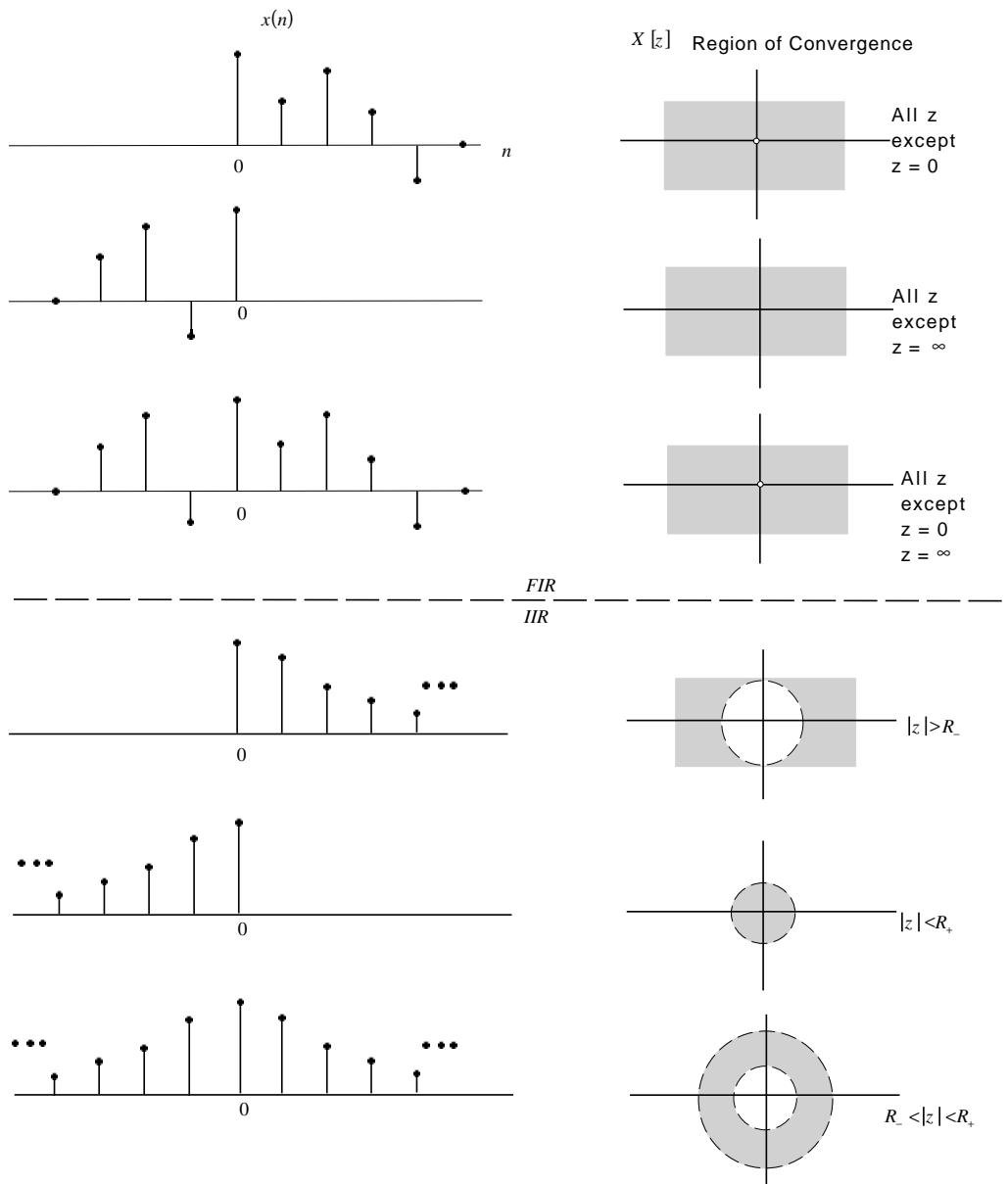The three infinite length cases, and the finite length case (with 3 examples) are summarized pictorially in Fig. 1.

Figure 1: Example ROC's for finite and infinite length sequences.