05 The Z-Transform

05.01 Definition of the Z-Transform

The Z-Transform is a generalization of the discrete-time Fourier transform. It exists for many sequences for which the DTFT does not exist. It is a function of the complex variable z.
Given a sequence $g(n)$, the Z-Transform is expressed as

$$G(z) = \sum_{n=-\infty}^{+\infty} g(n) z^{-n},$$

where $z = \text{Re }(z) + j\text{Im }(z)$ represents a continuous complex variable.
We will denote the Z-Transform of $g(n)$ also as:

$$\mathcal{Z}\{g(n)\} = G(z) = \sum_{n=-\infty}^{+\infty} g(n) z^{-n}.$$

If we consider $z = r \cdot e^{j\omega}$,

$$G(z) = \sum_{n=-\infty}^{+\infty} g(n) r^{-n} e^{-j\omega n}.$$

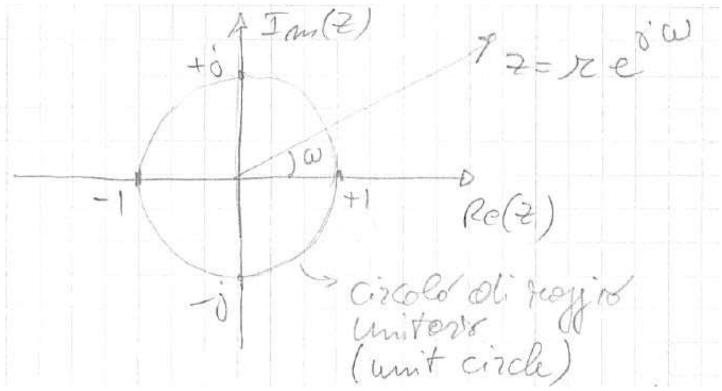
This expression can be interpreted as the Discrete-Time Fourier Transform (DTFT) of the modified sequence $\{g(n)r^{-n}\}$.
When we compare the definition of $G(z)$ with the DTFT of $g(n)$:

$$G(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} g(n) e^{-j\omega n}$$

we observe that

$$G(e^{j\omega}) = G(z)\Big|_{z=e^{j\omega}},$$

assuming these transforms exist.
We can provide a geometrical interpretation to the Z-Transform by considering a point z in the complex plane.



Assigned r and ω, the point $z = re^{j\omega}$ in the complex plane lies at the end of a vector with a magnitude of r and an angle of ω radians with respect to the real axis.
The set of points with $|z| = 1$ and consequently with $z = e^{j\omega}$, forms a circle with radius one (unit circle) in the z plane.
For $r = 1$, $z = e^{j\omega}$ and the Z-Transform $G(z)$ evaluated on the unit circle yields the DTFT of $g(n)$, assuming that this DTFT exists.
For $z = 1$, $G(z) = G(e^{j0})$, i.e., the value $G(e^{j\omega})$ for $\omega = 0$.
For $z = j$, $G(z) = G(e^{j\frac{\pi}{2}})$, i.e., the value $G(e^{j\omega})$ for $\omega = \frac{\pi}{2}$.
For $z = -1$, $G(z) = G(e^{j\pi})$, i.e., the value $G(e^{j\omega})$ for $\omega = \pi$.
For $z = -j$, $G(z) = G(e^{j\frac{3\pi}{2}})$, i.e., the value $G(e^{j\omega})$ for $\omega = \frac{3\pi}{2}$.
By moving along the unit circle from $z = 1$, till $z = -1$, till $z = 1$ again, we obtain all values of $G(e^{j\omega})$ for $0 \leq \omega \leq 2\pi$. As observed earlier, the DTFT is periodic with a period of 2π. Thus, by moving along the unit circle in either a clockwise or anti-clockwise direction, we can compute $G(e^{j\omega})$ for all values of ω within the range $-\infty < \omega < +\infty$.
As with the DTFT, the Z-Transform converges only when specific conditions are met, and typically, it converges within a specific region of the complex z plane.
For a given sequence, the set $\mathcal{R}$ of z values for which the Z-Transform converges is referred to as the *Region of Convergence (ROC)*. In particular, it can be proven that the Z-Transform converges for all values of z for which the sequence $g(n)r^{-n}$ is absolutely summable (with $r = |z|$):

$$\sum_{n=\infty}^{+\infty} \left|g(n)r^{-n}\right| < +\infty.$$

Even if the sequence $g(n)$ is not absolutely summable, $g(n)r^{-n}$ could still be absolutely summable. In other words, even if the sequence $g(n)$ does not have does not have a DTFT, it may have a Z-Transform within a certain region of the complex z plane.
Furthermore, the condition of absolute summability indicates that if there exists a specific $z = re^{j\omega}$ for which the Z-Transform exists (i.e., for which $\sum_{n=\infty}^{+\infty} |g(n)r^{-n}| < +\infty$), then the Z-Transform exists throughout the entire circle with a radius of r.
In general, the Region of Convergence $\mathcal{R}$ forms an annular region in the complex z plane:

$$R_{g-} < |z| < R_{g+}, \qquad \text{with } 0 \leq R_{g-} \leq R_{g+} \leq +\infty.$$

We will observe that different sequences can share the same Z-Transform; however, in such cases, they may have distinct regions of convergence. Therefore, it is crucial to explicitly specify the region of convergence for the Z-Transform.

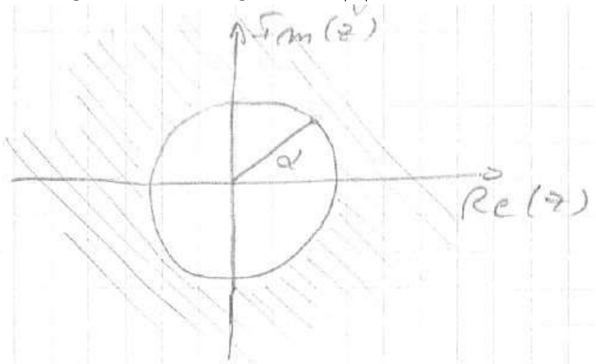
Example:
Let's compute the Z-Transform of the causal sequence $x(n) = \alpha^n \mu(n)$:

$$X(z) = \sum_{n=-\infty}^{+\infty} \alpha^n \mu(n) z^{-n} = \sum_{n=0}^{+\infty} \alpha^n z^{-n}.$$

This series converges if and only if $|\alpha z^{-1}| < 1$, and it converges to

$$X(z) = \frac{1}{1 - \alpha z^{-1}}.$$

The region of convergence is $|z| > \alpha$.



For $\alpha = 1$, $x(n) = \mu(n)$, thus $\mu(z) = \dfrac{1}{1 - z^{-1}}$ with a region of convergence $|z| > 1$.
Whenever the sequence is causal, the region of convergence lies outside a circle centered at the origin of the z plane.

Example:
Let us compute the Z-Transform of the exponential acausal sequence

$$x(n) = -\alpha^n \mu(-n-1)$$

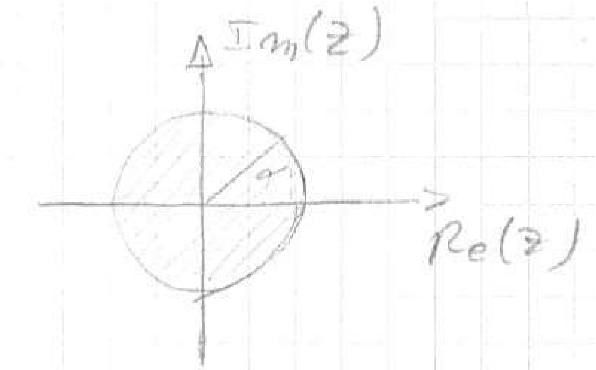
$$X(z) = \sum_{n=-\infty}^{+\infty} x(n) z^{-n} = \sum_{n=-\infty}^{-1} -\alpha^n z^{-n} =$$

(consider $n = -m$)

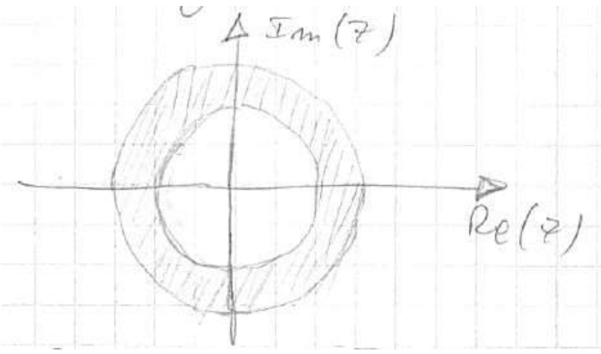
$$= \sum_{m=1}^{\infty} -\alpha^{-m} z^m = -\alpha^{-1} z \sum_{m=0}^{+\infty} \alpha^{-m} z^m =$$

$$= \frac{-\alpha^{-1} z}{1 - \alpha^{-1} z} = \frac{1}{1 - \alpha z^{-1}}$$

provided $|\alpha^{-1} z| < 1$.
We have obtained the same Z-Transform of the previous causal sequence, but now the region of convergence is $|z| < \alpha$.



Note that acausal sequences always have a region of convergence that is a circle centered at the origin of the z plane.
Two-sided sequences consist of a causal part and an acausal part, and their Z-Transform, if it converges at all, has an annular region of convergence.



This can be easily explained, as the causal part has a region of convergence $|z| > R_1$, and the acausal part has a region of convergence $|z| < R_2$. If $R_1 < R_2$, the two-sided sequence has a region of convergence $R_1 < |z| < R_2$. On the contrary, if $R_2 > R_1$ the Z-Transform does not converge.

Example:
Let's compute the Z-Transform of the finite length sequence

$$x(n) = \begin{cases} \alpha^n & M \leq n \leq N-1, \\ 0 & \text{elsewhere.} \end{cases}$$

$$X(z) = \sum_{n=M}^{N-1} \alpha^n z^{-n} =$$
$$= \alpha^M z^{-M} + \alpha^{M+1} z^{-M-1} + \ldots + \alpha^{N-1} z^{-N+1} =$$
$$= \alpha^M z^{-M} \sum_{m=0}^{N-M-1} \alpha^m z^{-m} =$$
$$= \alpha^M z^{-M} \frac{1 - \alpha^{N-M} z^{-(N-M)}}{1 - \alpha z^{-1}} =$$
$$= \frac{\alpha^M z^{-M} - \alpha^N z^{-N}}{1 - \alpha z^{-1}}$$

Since the series is composed of a finite number of terms, the region of convergence coincides with the entire z plane, except possibly from $z = 0$ or $z = \infty$.
In particular, for $N - 1 \geq M \geq 0$, the region of convergence is the entire z plane except at $z = 0$.
If $N - 1 > 0$, $M < 0$, the region of convergence is the entire z plane except at $z = 0$ and $z = \infty$.
If $M \leq N - 1 < 0$, the region of convergence is the entire z plane except at $z = \infty$.

If the region of convergence comprises the unit circle, the sequence admits the DTFT.
Nevertheless, even if a sequence has DTFT, there is no guarantee that the sequence also has a Z-Transform.
For example, the ideal low-pass sequence $h_{LP}(n)$, defined as

$$H(e^{j\omega}) = \begin{cases} 1 & |\omega| < \omega_c, \\ 0 & \text{otherwise} \end{cases},$$

admits DTFT, but it does not have a Z-Transform. Indeed, $h_{LP}(n) r^{-n}$, for any r, is not absolutely summable.

All the Z-Transforms we will consider in this course will always be rational functions of z^{-1}, i.e., the ratio of two polynomials:

$$H(z) = \frac{P(z)}{Q(z)} = \frac{p_0 + p_1 z^{-1} + \ldots + p_M z^{-M}}{d_0 + d_1 z^{-1} + \ldots + d_N z^{-N}}$$

The polynomial degree of the numerator is M, and that of the denominator is N.
An alternative form represents $H(z)$ as the ratio of two polynomials in z:

$$H(z) = z^{N-M} \frac{p_0 z^M + p_1 z^{M-1} + \ldots + p_M}{d_0 z^N + d_1 z^{N-1} + \ldots + d_N}$$

Both relations can be written in a factorized form as[1]

$$H(z) = \frac{p_0}{d_0} \frac{\prod_{l=1}^{M} \left(1 - \xi_l z^{-1}\right)}{\prod_{l=1}^{N} \left(1 - \lambda_l z^{-1}\right)} = \frac{p_0}{d_0} z^{N-M} \frac{\prod_{l=1}^{M} (z - \xi_l)}{\prod_{l=1}^{N} (z - \lambda_l)}$$

where ξ_l and λ_l are the roots of the numerator and denominator polynomials in z.
When $z = \xi_l$, we have $H(\xi_l) = 0$. Thus, the roots of the numerator ξ_l are referred to as the *zeros* of $H(z)$.
When $z = \lambda_l$, we have $H(\lambda_l) = \infty$. Thus, the roots of the denominator λ_l are referred to as the *poles* of $H(z)$.
Note that in our Z-Transform $H(z)$ we have M zeros and N poles.
Moreover, if $N > M$ we have additional $(N - M)$ zeros in the origin; if $M > N$ we have additional $(M - N)$ poles in the origin.
Note also that a rational Z-Transform can be fully characterized by the positions of its poles λ_l and zeros ξ_l, along with the constant term $\frac{p_0}{d_0}$.
It is interesting to observe that the region of convergence of a rational Z-Transform is delimited by the location of its poles.
If $H(z)$ has N poles, each with a different magnitude, we will have $N+1$ possible regions of convergence.
For each of these regions of convergence, we have a different sequence $h(n)$ that shares the same Z-Transform $H(z)$.

[1] $\prod_{i=1}^{M} x_i = x_1 \cdot x_2 \cdot \ldots \cdot x_M$.

05.02 The inverse Z-Transform

Let us denote $\mathcal{C}$ a closed contour in the region of convergence that encloses the origin (for example, $\mathcal{C}$ could be a circle centered in the origin within the region of convergence). The expression for the Inverse Z-Transform is given by:

$$g(n) = \frac{1}{2\pi j}\oint_{\mathcal{C}} G(z)z^{n-1}\mathrm{d}z,$$

where $\oint_{\mathcal{C}}$ indicates the integration performed on the closed contour $\mathcal{C}$.
In general, it is challenging to directly evaluate the integral along a closed contour. Additionally, the integral needs to be computed for every time value n.
Fortunately, we will only consider rational Z-Transforms and causal sequences. Under these conditions, there are simple methods that allow us to compute the Z-Transform.

A first method for computing the Z-Transform is the *tabular method*. Very often, the Z-Transform has a simple expression or can be decomposed into the sum of simple expressions. Many books provide tables that offer inverses for these elementary transforms.

$\mathcal{Z}\{x(n)\}$	$x(n)$	R.O.C.
1	$\delta(n)$	z-plane
z^{-N}	$\delta(n-N)$	z-plane - $\{z=0\}$
$\frac{1}{1-z^{-1}}$	$\mu(n)$	$\lvert z\rvert > 1$
$\frac{1}{1-\alpha z^{-1}}$	$\alpha^n\mu(n)$	$\lvert z\rvert > \lvert\alpha\rvert$
$\frac{\alpha z^{-1}}{(1-\alpha z^{-1})^2}$	$n\alpha^n\mu(n)$	$\lvert z\rvert > \lvert\alpha\rvert$
$\frac{1}{(1-\alpha z^{-1})^2}$	$(n+1)\alpha^n\mu(n)$	$\lvert z\rvert > \lvert\alpha\rvert$

For example, if we have

$$H(z) = \frac{0.5z}{z^2 - z + 0.25} \qquad \text{R.O.C. } |z| > 0.5$$

Then,

$$H(z) = \frac{0.5z}{(z-0.5)^2} = \frac{0.5z^{-1}}{(1-0.5z^{-1})^2}.$$

From the table, we can deduce that $h(n) = n\,0.5^n\mu(n)$.

A second method for the Z-Transform inversion is that of the *partial fraction expansion*.
If a sequence is causal, the region of convergence is external to a circle centered in the origin. We can compute the inverse Z-Transform of a rational function $G(z)$ by decomposing it into partial fractions, inverting each of the partial fractions, and summing up all the resulting sequences. We exploit the linearity property of the Z-Transform, which we will study later:
“Let $g(n)$ and $h(n)$ be two sequences with Z-Transform $G(z)$ and $H(z)$, respectively. For any constant

α and β, the Z-Transform of $\alpha g(n) + \beta h(n)$ is $\alpha G(z) + \beta H(z)$ with the region of convergence being the intersection of the regions of convergence of $G(z)$ and $H(z)$."

We consider $G(z) = \frac{P(z)}{D(z)}$, where $P(z)$ is a polynomial of degree M and $D(z)$ is a polynomial of degree N.

If $M \geq N$, the rational function $\frac{P(z)}{D(z)}$ is called an *improper* fraction.

If $M < N$, the rational function is called a *proper* fraction.

If the fraction is improper, we can decompose it as the sum of a polynomial in z^{-1} plus a proper fraction:

$$G(z) = \sum_{l=0}^{M-N} \eta_l z^{-l} + \frac{P_1(n)}{D(n)},$$

where the degree of $P_1(n)$ is lower than N. To obtain this expression, it suffices to divide $P(z)$ by $D(z)$ (we will cover polynomial division later).

The inverse-transform of $\sum_{l=0}^{M-N} \eta_l z^{-l}$ is known, as it is $\sum_{l=0}^{M-N} \eta_l \delta(n-l)$. Thus, we concentrate our attention only on the inversion of a proper rational function. We will distinguish two cases:

- simple poles,
- multiple poles.

Simple poles:

In most cases, $G(z)$ is a proper rational function with simple poles, i.e., with all distinct poles.

Let us denote λ_k, $1 \leq k \leq N$, the N poles of $G(z)$. Then, $G(z)$ can be rewritten in the following form:

$$G(z) = \sum_{l=1}^{N} \frac{\rho_l}{1 - \lambda z^{-1}},$$

where the ρ_l constants are called *residuals* and can be easily computed using:

$$\rho_l = (1 - \lambda_l z^{-1}) G(z) \Big|_{z=\lambda_l}$$

(i.e., by removing the factor $(1 - \lambda_l z^{-1})$ from the denominator and computing the resulting value for $z = \lambda_l$).

Each term $\frac{\rho_l}{1 - \lambda_l z^{-1}}$ has a simple inverse: $\rho_l \, \lambda^n \, \mu(n)$. Thus, by exploiting the linearity of the Z-Transform, we get

$$g(n) = \sum_{l=1}^{N} \rho_l \, \lambda_l^n \, \mu(n).$$

Example:

$$H(z) = \frac{1 + 2.0z^{-1}}{(1 - 0.2z^{-1}) \cdot (1 + 0.6z^{-1})}$$

$$H(z) = \frac{\rho_1}{1 - 0.2z^{-1}} + \frac{\rho_2}{1 + 0.6z^{-1}}$$

$\lambda_1 = 0.2$ and $\lambda_2 = -0.6$.

$$\rho_1 = \frac{1+2.0z^{-1}}{1+0.6z^{-1}}\Big|_{z=0.2} = \frac{1+2\cdot(0.2)^{-1}}{1+0.6\cdot(0.2)^{-1}} = \frac{11}{4} = 2.75$$

$$\rho_2 = \frac{1+2.0z^{-1}}{1-0.2z^{-1}}\Big|_{z=-0.6} = \frac{1+2\cdot(-0.6)^{-1}}{1-0.2\cdot(-0.6)^{-1}} = \frac{0.6-2.0}{0.6+0.2} = -1.75$$

Thus, we obtain

$$h(n) = 2.75\,(0.2)^n\,\mu(m) - 1.75\,(-0.6)^n\,\mu(n)$$

Another way to compute the residues is to express the numerator of $G(z)$ as a function of the residues ρ_i and then impose the equality of the numerator polynomials. This way, we obtain a system of equations in the unknown ρ_i which, once solved, allows us to derive the partial fraction expansion of $G(z)$.

In the previous example:

$$H(z) = \frac{\rho_1}{1-0.2z^{-1}} + \frac{\rho_2}{1+0.6z^{-1}} =$$

$$= \frac{\rho_1(1+0.6z^{-1}) + \rho_2(1-0.2z^{-1})}{(1-0.2z^{-1})\cdot(1+0.6z^{-1})} =$$

$$= \frac{\rho_1 + \rho_2 + (0.6\rho_1 - 0.2\rho_2)z^{-1}}{(1-0.2z^{-1})\cdot(1+0.6z^{-1})}$$

Imposing:

$$\rho_1 + \rho_2 + (0.6\rho_1 - 0.2\rho_2)z^{-1} = 1 + 2z^{-1}$$

it must be:

$$\begin{cases} \rho_1 + \rho_2 = 1 \\ 0.6\rho_1 - 0.2\rho_2 = 2 \end{cases}$$

Solving this system of equations, we obtain $\rho_1 = 2.75$ and $\rho_2 = -1.75$.

When the rational function has real coefficients, the complex poles always appear in complex-conjugate pairs, and the corresponding residues are also complex-conjugate. Thus, they can be combined into a single fraction of the second order:

$$\frac{\rho_l}{1-\lambda_l z^{-1}} + \frac{\rho_l^*}{1-\lambda_l^* z^{-1}} = \frac{p_{0,l} + p_{1,l}z^{-1}}{1 + d_{1,l}z^{-1} + d_{2,l}z^{-2}}$$

Multiple poles:

If $G(z)$ is a rational function with multiple poles (double poles, triple poles, or even of higher order), the partial fraction expansion takes on a different form.

Suppose $z = \nu$ is a pole of multiplicity L (i.e., in the denominator there are L roots equal to ν), and that the other poles are simple. We have

$$G(z) = \sum_{l=1}^{N-L} \frac{\rho_l}{1-\lambda_l z^{-1}} + \sum_{i=1}^{L} \frac{\gamma_i}{(1-\nu z^{-1})^i}$$

where the constants γ_i (which are no longer called residues) are computed with the formula:

$$\gamma_i = \frac{1}{(L-i)!(-\nu)^{L-i}} \frac{\mathrm{d}^{L-i}}{\mathrm{d}(z^{-1})^{L-i}} \left[(1-\nu z^{-1})^L G(z)\right]\Big|_{z=\nu},$$

for $1 \leq i \leq L$.

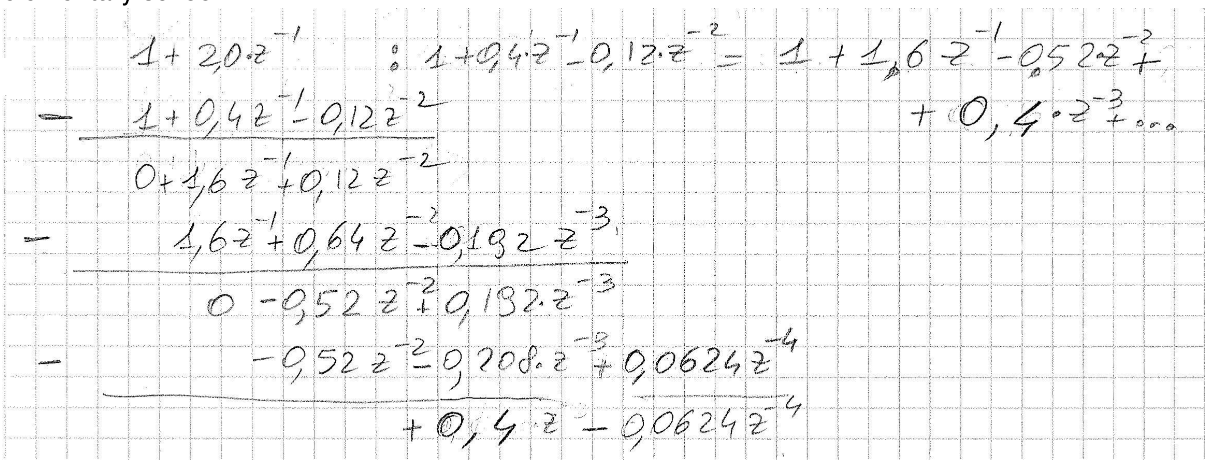
Thus, the computation of γ_i requires eliminating the factor $(1-\nu z^{-1})^L$ from the denominator of $G(z)$ and computing the derivatives with respect to z^{-1} from the first to the $(L-1)$-th. The values of these derivatives and of $(1-\nu z^{-1})^L G(z)$ computed for $z=\nu$ provide the γ_i values. The other residues are computed as before. Eventually, from the tables reported in many books, it is easy to find the inverse transform of $\frac{\gamma_i}{(1-\nu z^{-1})^i}$.

Inverse Z-Transform via long division method.

Another method for computing the inverse of a causal Z-Transform is the *long division*.

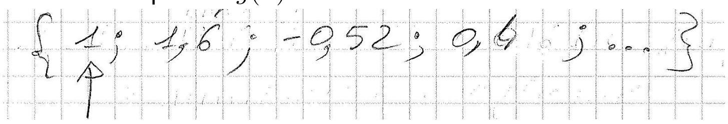
The objective is to obtain the expansion of $G(z)$ in the form of a power series in z^{-1}. If we possess such a series, we immediately know the inverse transform. Indeed, the coefficient that multiplies z^{-n} is the term $g(n)$ of the inverse transform.

If $G(z)$ is rational, we can express the numerator and denominator as polynomial in z^{-1}. Then, we can derive the power series expansion by means of the (long) division of the numerator by the denominator. The division of polynomials applies the same methodology as the divisions you have studied in elementary school:



(How many times does 1 go into 1? 1. Thus we subtract from $1+0.2z^{-1}$ the polynomial $1\cdot(1+0.4z^{-1}-0.12z^{-2})$ and we get the remainder $1.6z^{-1}+0.12z^{-2}$. How many times does 1 go into $1.6z^{-1}$? $1.6z^{-1}$. Thus, we subtract from $1.6z^{-1}+0.12z^{-2}$ the polynomial $1.6z^{-1}\cdot(1+0.4z^{-1}-0.12z^{-2})$ and get the remainder $-0.52z^{-2}+0.192z^{-3}$. How many times does 1 go into $-0.52z^{-2}$? $-0.52z^{-2}$. And so on.)

Thus the sequence $g(n)$ is :

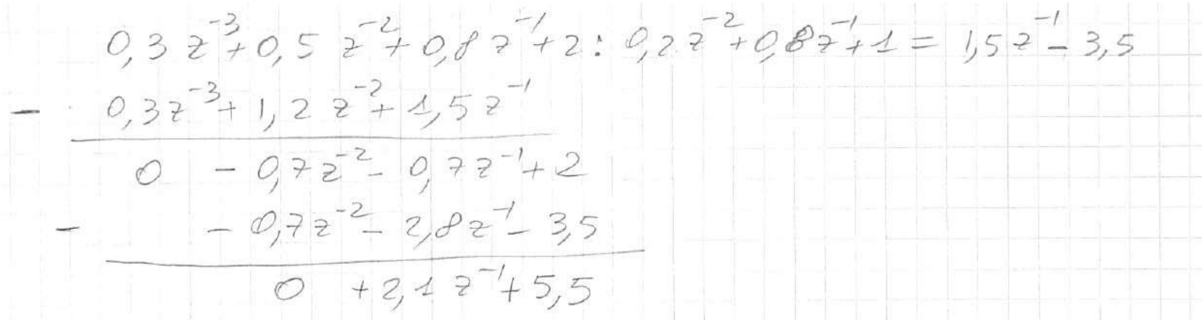


The division between polynomials is also used in the case of the partial fraction expansion for passing from an improper $G(z)$ to an expression composed by a polynomial in z^{-1} (which has a trivial inverse) plus a proper rational function.

For example, consider:

$$G(z) = \frac{2 + 0.8z^{-1} + 0.5z^{-2} + 0.3z^{-3}}{1 + 0.8z^{-1} + 0.2z^{-2}}.$$

We must divide the two polynomials till we get a remainder of degree lower than the denominator. Since we only want to reduce the order in z^{-1} of the numerator, in this case, we apply the same division method we have seen but, compared with the long division, we must reverse the order of the two polynomials:



(Note that $(1.5z^{-1} - 3.5) \cdot (0.2z^{-2} + 0.8z^{-1} + 1) + (2.1z^{-1} + 5.5) = 0.3z^{-3} + 0.5z^{-2} + 0.8z^{-1} + 2$).
Thus,

$$G(z) = -3.5 + 1.5z^{-1} + \frac{5.5 + 2.1z^{-1}}{1 + 0.8z^{-1} + 0.2z^{-2}},$$

and we have

$$g(n) = -3.5\delta(n) + 1.5\delta(n-1) + \rho_1 \lambda_1^n \mu(n) + \rho_2 \lambda_2^n \mu(n),$$

with λ_1 and λ_2 the roots of $1 + 0.8z^{-1} + 0.2z^{-2}$, and ρ_1 and ρ_2 the corresponding residues.

05.03 Properties of the Z-Transform

Also the Z-Transform has important properties. In the following we will consider

$$g(n) \stackrel{\mathcal{Z}}{\longleftrightarrow} G(z) \qquad \text{with ROC } \mathcal{R}_g$$

$$h(n) \stackrel{\mathcal{Z}}{\longleftrightarrow} H(z) \qquad \text{with ROC } \mathcal{R}_h$$

where $\mathcal{R}_g$ is the annular region $R_{g-} < |z| < R_{g+}$, and $\mathcal{R}_h$ is the annular region $R_{h-} < |z| < R_{h+}$.

Conjugation theorem
The Z-Transform of $g^*(n)$ is $G^*(z^*)$ with ROC $\mathcal{R}_g$.
Proof:

$$\mathcal{Z}\{g^*(n)\} = \sum_{n=-\infty}^{+\infty} g^*(n) z^{-n} =$$

$$= \left[\sum_{n=-\infty}^{+\infty} g(n)(z^*)^{-n} \right]^* = G^*(z^*).$$

Moreover, since $|z^*| = |z|$, $G^*(z^*)$ converges for all the z values for which $G(z)$ converges.

Q.E.D.

Time-reversal

The Z-Transform of $x(n) = g(-n)$ is $X(z) = G(1/z)$ with ROC: $R_{g+}^{-1} < |z| < R_{g-}^{-1}$.

Proof:

$$\sum_{n=-\infty}^{+\infty} g(-n)z^{-n} = \sum_{m=-\infty}^{+\infty} g(m)z^{m} =$$

$$= \sum_{m=-\infty}^{+\infty} g(m)\left(z^{-1}\right)^{-m} = G(1/z).$$

$X(z) = G(1/z)$ converges for all z for which $1/z$ belongs to the ROC of $G(z)$. Thus, for all z for which

$$\frac{1}{R_{g+}} < |z| < \frac{1}{R_{g-}}.$$

Q.E.D.

Linearity

The Z-Transform of the sequence $\alpha g(n) + \beta h(n)$ is $\alpha G(z) + \beta H(z)$, with region of convergence that is the intersection between the two ROCs $\mathcal{R}_g$ and $\mathcal{R}_h$, i.e., $\mathcal{R}_g \cap \mathcal{R}_h$.

(If the two regions of convergence do not overlap the Z-Transform of $\alpha g(n) + \beta h(n)$ does not exist!)

Exercise:

Let us compute the Z-Transform of

$$x(n) = r^n \cos(\omega_0 n)\mu(n),$$

$r \in \mathbb{R}$.

Note that

$$x(n) = r^n \frac{e^{j\omega_0 n} + e^{-j\omega_0 n}}{2}\mu(n) =$$

$$= \frac{1}{2}r^n e^{j\omega_0 n}\mu(n) + \frac{1}{2}r^n e^{-j\omega_0 n}\mu(n).$$

Since $\alpha^n\mu(n)$ has Z-Transform $\frac{1}{1-\alpha z^{-1}}$ with ROC $|z| > \alpha$,

$$X(z) = \frac{1/2}{1 - re^{j\omega_0}z^{-1}} + \frac{1/2}{1 - re^{-j\omega_0}z^{-1}} =$$

$$= \frac{1}{2}\frac{1 - re^{-j\omega_0}z^{-1} + 1 - re^{j\omega_0}z^{-1}}{(1 - re^{j\omega_0}z^{-1})\cdot(1 - re^{-j\omega_0}z^{-1})} =$$

$$= \frac{1 - r\cos(\omega_0)z^{-1}}{1 - re^{j\omega_0}z^{-1} - re^{-j\omega_0}z^{-1} + r^2z^{-2}} =$$

$$= \frac{1 - r\cos(\omega_0)z^{-1}}{1 - 2r\cos(\omega_0)z^{-1} + r^2z^{-2}}$$

with ROC $|z| > r$.

Time-shifting

The Z-Transform of the delayed sequence $x(n) = g(n - n_0)$ (with n_0 integer) is $X(z) = z^{-n_0}G(z)$. The region of convergence of $X(z)$ is the same as $G(z)$ apart from $z = 0$ for $n_0 > 0$ or $z = \infty$ for $n_0 < 0$ (because z^{-n_0} adds $|n_0|$ poles in $z = 0$ or $z = \infty$).

Proof:

$$\sum_{n=-\infty}^{+\infty} g(n - n_0)z^{-n} = \sum_{m=-\infty}^{+\infty} g(m)z^{-(m+n_0)} =$$

$$= \sum_{m=-\infty}^{+\infty} g(m)z^{-m} \cdot z^{-n_0} = G(z)z^{-n_0}.$$

Q.E.D.

Multiplication for an exponential sequence

The Z-Transform of the sequence $x(n) = \alpha^n g(n)$ is $X(Z) = G\left(\frac{z}{\alpha}\right)$. The region of convergence is $|\alpha|\mathcal{R}_g$.

Proof:

$$\sum_{n=-\infty}^{+\infty} \alpha^n g(n)z^{-n} = \sum_{n=-\infty}^{+\infty} g(n)\left(\frac{z}{\alpha}\right)^n = G\left(\frac{z}{\alpha}\right).$$

$X(z) = G\left(\frac{z}{\alpha}\right)$ converges for any value z for which $\frac{z}{\alpha} \in \mathcal{R}_g$ and thus when

$$R_{g-}|\alpha| < |z| < R_{g+}|\alpha|.$$

Q.E.D.

Convolution Theorem

The Z-Transform of the convolution sum between two sequences $g(n) \circledast h(n)$ is $G(z) \cdot H(z)$ with a region of convergence that is the intersection between the ROCs of $G(z)$ and $H(z)$, i.e., $\mathcal{R}_g \cap \mathcal{R}_h$.

Proof:

$$X(z) = \sum_{n=-\infty}^{+\infty} x(n)z^{-n} =$$

$$= \sum_{n=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} g(k)h(n-k)z^{-n} =$$

$$= \sum_{k=-\infty}^{+\infty} g(k) \sum_{n=-\infty}^{+\infty} h(n-k)z^{-n} =$$

$$= \sum_{k=-\infty}^{+\infty} g(k) \sum_{m=-\infty}^{+\infty} h(m)z^{-(m+k)} =$$

$$= \sum_{k=-\infty}^{+\infty} g(k)z^{-k}H(z) = G(z)H(z).$$

Moreover, $X(z)$ converges for any value z for which $\sum_{m=-\infty}^{+\infty} h(m)z^{-m}$ and $\sum_{k=-\infty}^{+\infty} g(k)z^{-k}$ converge. Thus, the ROC of $X(z)$ is $\mathcal{R}_g \cap \mathcal{R}_h$.

Q.E.D.

05.04 The Transfer Function

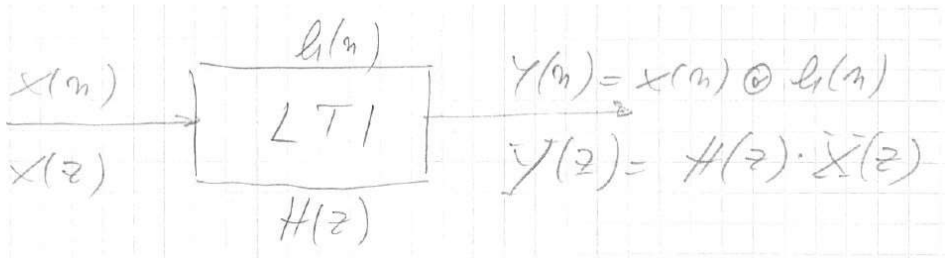
The convolution theorem leads to the concept of *Transfer Function* of an LTI system.
Let us consider an LTI system with impulse response $h(n)$. The input-output relation of the system is

$$y(n) = h(n) \circledast x(n) = \sum_{k=-\infty}^{+\infty} h(k)x(n-k)$$

with $x(n)$ and $y(n)$ the input and output of the LTI system, respectively.
Let us indicate with $X(z)$, $Y(z)$, and $H(z)$ the Z-transform of $x(n)$, $y(n)$, and $h(n)$, respectively. According to the convolution theorem, the input-output relation in the Z-Transform domain becomes:

$$Y(z) = H(z)X(z),$$

with $H(z) = \sum_{n=-\infty}^{+\infty} h(n)z^{-n}$.



The Z-Transform of the impulse response, $H(z)$, is commonly called *Transfer Function*. Similar to the impulse response, it completely characterizes the LTI system.

$$H(z) = \frac{Y(z)}{X(z)},$$

$H(z)$ is given by the ratio between the Z-Transform of $y(n)$ and $x(n)$.
The transfer function can be seen as a generalization of the frequency response concept. If the region of convergence of $H(z)$ comprises the unit circle:

$$H(e^{j\omega}) = H(z)\Big|_{z=e^{j\omega}}.$$

Thus, the frequency response can be obtained by evaluating the transfer function on the unit circle.
The transfer function can be obtained from the frequency response by *analytic continuation*:

$$H(z) = H(e^{j\omega})\Big|_{\omega=\frac{1}{j}\ln(z)}.$$

We are particularly interested in two types of discrete-time LTI systems:

- FIR filters,
- IIR filters (described by a finite difference equation).

FIR filters

FIR filters have a finite impulse response:

$$y(n) = \sum_{m=-N_1}^{N_2} h(m)x(n-m)$$

with $h(n) = 0 \; \forall n < N_1$, and $\forall n > N_2$.
$h(n)$ has length $N_1 + N_2 + 1$.

$$H(z) = \sum_{n=N_1}^{N_2} h(n)z^{-n} =$$

$$= h(-N_1)z^{+N_1} + h(-N_1+1)z^{+N_1-1} + \ldots + h(0) + h(1)z^{-1} + \ldots + h(N_2-1)z^{-N_2+1} + h(N_2)z^{-N_2}.$$

Properties of FIR filters

- The region of convergence of $H(z)$ coincides with the entire complex plane, possibly excluding 0 and ∞.
- FIR filters are guaranteed to be stable in the BIBO sense:

$$\sum_{n=-\infty}^{+\infty} |h(n)| = \sum_{n=-N_1}^{N_2} |h(n)| < +\infty.$$

- FIR filters are always realizable. Even if the impulse response has non-null terms for $n < 0$, introducing a delay z^{-N_1} ensures realizability.
- For causal filters, the transfer function is given by:

$$H(z) = h(0) + h(1)z^{-1} + h(2)z^{-2} + \ldots + h(N-1)z^{N-1}.$$

  Thus, the transfer function is a polynomial in z^{-1}, earning them the name *filters with only zeros*.
- Since the input-output relation is non-recursive, quantization errors propagate only from the input to the output in FIR filters.
- We will prove that FIR filters can have linear phase response.

IIR filters

The IIR filters we are interested in are described by a finite difference equation:

$$y(n) = b_0x(n) + b_1x(n-1) + \ldots + b_Mx(n-M) - a_1y(n-1) - a_2y(n-2) - \ldots - a_Nx(n-N).$$

For these filters, it is easy to compute the transfer function using the linearity property and the time-shift property of the Z-Transform. By transforming both members:

$$Y(z) = b_0X(z) + b_1z^{-1}X(z) + b_2z^{-2}X(z) + \ldots b_Mz^{-M}X(z) - a_1z^{-1}Y(z) - a_2z^{-2}Y(z) - \ldots - a_Nz^{-N}Y(z),$$

$$\left(1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_N z^{-N}\right) Y(z) = \left(b_0 + b_1 z^{-1} + \ldots + b_M z^{-M}\right) X(z)$$

Thus,

$$H(z) = \frac{Y(z)}{X(z)} = \frac{b_0 + b_1 z^{-1} + \ldots + b_M z^{-M}}{1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_N z^{-N}}.$$

The transfer function is a rational function in z^{-1}. It is also:

$$H(z) = z^{M-N} \cdot \frac{b_0 z^M + b_1 z^{M-1} + \ldots + b_M}{z^N + a_1 z^{N-1} + \ldots + a_N}.$$

The roots of the numerator polynomial are called *zeros* of the system. The roots of the denominator polynomial are called *poles*. The denominator polynomial is also called the *characteristic polynomial*.

Properties of IIR filters

- IIR filters or systems can be stable or unstable in BIBO sense. They are stable if and only if

$$\sum_{n=-\infty}^{+\infty} |h(n)| < +\infty.$$

  IIR systems are stable if and only if the region of convergence of the transfer function includes the unit circle. Indeed,

$$|H(z)| = \left| \sum_{n=-\infty}^{+\infty} h(n) z^{-n} \right| \leq \sum_{n=-\infty}^{+\infty} \left| h(n) z^{-n} \right| = \sum_{n=-\infty}^{+\infty} |h(n)||z^{-n}|$$

  On the unit circle $|z^{-n}| = 1$:

$$|H(z)|\Big|_{z=e^{j\omega}} = |H(e^{j\omega})| \leq \sum_{n=-\infty}^{+\infty} |h(n)|.$$

  Thus, if this sum is finite, $|H(z)|\Big|_{z=e^{j\omega}}$ is finite, and the transfer function converges on the unit circle. Conversely, if $H(z)$ converges for $z = e^{j\omega}$, then $\sum_{n=\infty}^{+\infty} |h(n)|$ is finite, and the filter is stable.

- A causal IIR filter is stable if and only if its poles fall within the unit circle.
  Indeed, in case $h(n)$ is causal, the region of convergence of $H(z)$ is external to the circle that encloses all poles. Since for the BIBO stability, the unit circle must belong to the region of convergence, then the poles must all fall within the unit circle.
  We can justify this property also from the observation that (in the partial fraction expansion) to any pole outside the unit circle corresponds an exponentially increasing sequence:

$$H(z) = \frac{1}{1 - pz^{-1}} \longrightarrow h(n) = p^n \mu(n),$$

  if $|p| > 1$ then $h(n) \to +\infty$.

- The input-output relationship of these filters is always recursive. For this reason, we will see that the quantization errors recirculate.

- We will also see that in causal IIR filters the phase response cannot be linear and that these filters introduce a phase distortion.

Despite these adverse properties, IIR filters adapt better than FIR filters to the realization of the specifications imposed on low-pass, high-pass, and pass-band filters. IIR filters can implement these specifications with a much lower number of coefficients than FIR filters. This property derives from the fact that the transfer function of IIR filters has zeros and poles, while that of FIR filters has only zeros. We will return to these filters later in our course.

For more information study:

S. K. Mitra, ”Digital Signal Processing: a computer based approach,” 4th edition, McGraw-Hill, 2011
Chapters 6.1-6.4, pp. 277-297
Chapter 6.5, pp. 297-303
Chapter 6.7, pp. 308-311 and pp. 315-319

