5 Toeplitz Operators

There are two signal spaces which will be important for us.

- **Semi-infinite signals:** Functions $x \in \ell_2(\mathbb{Z}_+, \mathbb{R})$. They have a Fourier transform $g = Fx$, where $g \in H_2$; that is, $g : \mathbb{D} \to \mathbb{C}$ is analytic on the open unit disk, so it has no poles there.
- **Bi-infinite signals:** Functions $x \in \ell_2(\mathbb{Z}, \mathbb{R})$. They have a Fourier transform $g = Fx$, where $g \in L_2(\mathbb{T})$. Then $g : \mathbb{T} \to \mathbb{C}$, and g may have poles both inside and outside the disk.

5.1 Causality and Time-invariance

Suppose G is a bounded linear map $G : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ given by

$$y_i = \sum_{j \in \mathbb{Z}} G_{ij} u_j$$

where G_{ij} are the coefficients in its matrix representation. The map G is called ***time-invariant*** or ***shift-invariant*** if it is ***Toeplitz***, which means

$$G_{i-1,j} = G_{i,j+1}$$

that is G is constant along diagonals from top-left to bottom right. Such matrices are ***convolution*** operators, because they have the form

$$G = \begin{bmatrix} \ddots & & & & & \\ & a_0 & a_{-1} & a_{-2} & & \\ & a_1 & \boxed{a_0} & a_{-1} & a_{-2} & \\ & a_2 & a_1 & a_0 & a_{-1} & \\ & & a_2 & a_1 & a_0 & \\ & & & & & \ddots \end{bmatrix}$$

Here the box indicates the $0, 0$ element, since the matrix is indexed from $-\infty$ to ∞. With this matrix, we have $y = Gu$ if and only if

$$y_i = \sum_{k \in \mathbb{Z}} a_{i-j} u_j$$

We say G is ***causal*** if the matrix G is *lower triangular*. For example, the matrix

$$G = \begin{bmatrix} \ddots & & & & & \\ & a_0 & & & & \\ & a_1 & \boxed{a_0} & & & \\ & a_2 & a_1 & a_0 & & \\ & a_3 & a_2 & a_1 & a_0 & \\ & & & & & \ddots \end{bmatrix}$$

is both causal and time-invariant. These definitions extend in the natural way to operators on $\ell_2(\mathbb{Z}_+)$; for example

$$H = \begin{bmatrix} b_0 & & & & \\ b_1 & b_0 & & & \\ b_2 & b_1 & b_0 & & \\ b_3 & b_2 & b_1 & b_0 & \\ & & & & \ddots \end{bmatrix}$$

is both causal and time-invariant on $\ell_2(\mathbb{Z}_+)$. The causal and time-invariant operators are precisely those of interest to us; a linear time-invariant state-space system gives rise to such an operator.

We'll occasionally have need to characterize such operators without reference to the matrix representation. To do so, for each $t \in \mathbb{Z}$ define the projection $P_t : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ by

$$y = P_t u \qquad \text{if} \qquad y_s = \begin{cases} u_s & \text{if } s \leq t \\ 0 & \text{otherwise} \end{cases}$$

Then one defines $G : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ to be ***causal*** if $P_t G P_t = P_t G$ for all $t \in \mathbb{Z}$, and one can show that this holds iff G is lower-triangular.

5.2 The Shift Operator

We define the ***shift operator*** $L : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ by

$$(Lx)_k = x_{k-1}$$

also called the ***delay operator*** or the ***right-shift***. In matrix form L is given by

$$L = \begin{bmatrix} \ddots & & & & & \\ & 0 & & & & \\ & 1 & \boxed{0} & & & \\ & & 1 & 0 & & \\ & & & 1 & 0 & \\ & & & & & \ddots \end{bmatrix}$$

Then $L^* = L^{-1}$. We can use this to give a coordinate-free definition of time-invariance as follows. A map $G : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ is called ***time-invariant*** if

$$L^* G L = G$$

and this holds iff G is Toeplitz, since it means that

$$GL = LG$$

By definition, the matrix for G has i, j entry equal to

$$G_{ij} = \langle e_i, G e_j \rangle$$

the above equation means that

$$G_{i-1,j} = G_{i,j+1}$$

so G is Toeplitz. Notice that we need a slightly different argument on $\ell_2(\mathbb{Z}_+)$, since the right-shift on $\ell_2(\mathbb{Z}_+)$ is not invertible.

5.3 Multiplication Operators

Our primary interest is linear time-invariant maps on ℓ_2. We have seen that this means that the matrix representation is Toeplitz, that is, that these are convolution maps. These arise by multiplication in the frequency domain.

For functions $g : \mathbb{T} \to \mathbb{C}$ define the infinity-norm to be

$$\|g\|_\infty = \operatorname*{ess\,sup}_{\theta \in [0,2\pi]} |g(e^{j\theta})|$$

Here ess sup means the essential supremum; in other words, $M = \|g\|_\infty$ implies that

$$|g(e^{j\theta})| \leq M$$

at almost all θ, excluding at most a set of measure zero. Then define $L_\infty(\mathbb{T})$ to be

$$L_\infty(\mathbb{T}) = \left\{ g : \mathbb{T} \to \mathbb{C} \;\middle|\; g \text{ is Lesbesgue measurable and } \|g\| \text{ is finite} \right\}$$

Suppose $g \in L_\infty(\mathbb{T})$. Define the ***multiplication operator*** M_g by

$$(M_g f)(\lambda) = g(\lambda) f(\lambda)$$

So we write $y = M_g f$ to mean $y = gf$, the idea being that this reminds us that M_g is a linear map.

The following result is very simple, but important enough that we'll state it as a theorem.

Theorem 1. *Suppose $g \in L_\infty(\mathbb{T})$. Then $M_g : L_2(\mathbb{T}) \to L_2(\mathbb{T})$ and*

$$\|M_g\| = \|g\|_\infty$$

Proof. Suppose $u \in L_2(\mathbb{T})$, and $y = M_g u$. Then

$$\begin{aligned} \|y\|_2^2 &= \frac{1}{2\pi} \int_0^{2\pi} |y(e^{j\theta})|^2 \, d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} |g(e^{j\theta})|^2 |u(e^{j\theta})|^2 \, d\theta \\ &\leq \|g\|_\infty^2 \frac{1}{2\pi} \int_0^{2\pi} |u(e^{j\theta})|^2 \, d\theta \\ &= \|g\|_\infty^2 \|u\|_2^2 \end{aligned}$$

and so $y \in L_2(\mathbb{T})$ and the induced 2-norm of M_g satisfies

$$\|M_g\| \leq \|g\|_\infty$$

To show equality, we need to show that for all $\varepsilon > 0$ there exists $u \in L_2(\mathbb{T})$ with $\|u\| = 1$ and

$$\|M_g u\|_2 \geq \|g\|_\infty - \varepsilon$$

We'll assume that g is *continuous.* The result holds more generally, but requires some measure theory to prove. The idea of the proof is as follows. Suppose $e^{j\theta_0}$ is the point on $\mathbb{T}$ at which the infinity norm is achieved

$$|g(e^{j\theta_0})| = \max_{\theta \in [0,2\pi]} |g(e^{j\theta})|$$

and we know there is such a point since g is continuous. Then let u be

$$u(e^{j\theta}) = \begin{cases} \sqrt{\pi/\delta} & \text{if } \theta \in [\theta_0 - \delta, \theta_0 + \delta] \\ 0 & \text{otherwise} \end{cases}$$

Then $\|u\| = 1$ and as we can show that by choosing δ small we can make $\|M_g u\|$ as close as we like to $\|g\|_\infty$. The details are left as an exercise. ■

We now turn to the correspondence between multiplication in the frequency domain and convolution in the time domain. From now on, we'll use the hat notation for Fourier transforms. That is, if $g \in \ell_2(\mathbb{Z})$ then we'll let $\hat{g} = Fg$.

The first thing to notice is that $L_\infty(\mathbb{T}) \subset L_2(\mathbb{T})$. So if we have $\hat{g} \in L_\infty(\mathbb{T})$, it is also an element of $L_2(\mathbb{T})$ and so it has well-defined Fourier coefficients, given by $g = F^* \hat{g}$.

Theorem 2. *Suppose $\hat{g} \in L_\infty(\mathbb{T})$, with its associated multiplication operator $M_{\hat{g}}$ and let $g = F^*\hat{g}$. Then $F^* M_{\hat{g}} F$ is Toeplitz, and is given by*

$$F^* M_{\hat{g}} F = \begin{bmatrix} \ddots & & & & & \\ & g_0 & g_{-1} & g_{-2} & & \\ & g_1 & \boxed{g_0} & g_{-1} & g_{-2} & \\ & g_2 & g_1 & g_0 & g_{-1} & \\ & & g_2 & g_1 & g_0 & \\ & & & & & \ddots \end{bmatrix}$$

Proof. As usual, let e_i be the i'th standard basis vector in $\ell_2(\mathbb{Z})$. Let $G = F^* M_g F$. Then $G : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$, and

$$\begin{aligned} G_{pq} &= \langle e_p, F^* M_{\hat{g}} F e_q \rangle \\ &= \langle F e_p, M_{\hat{g}} F e_q \rangle \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{g}(e^{j\theta}) e^{j(q-p)k\theta} \, d\theta \\ &= \langle \phi_{p-q}, \hat{g} \rangle \\ &= g_{p-q} \end{aligned}$$

and so G_{pq} depends has the specified form as desired. ■

Suppose $u \in \ell_2(\mathbb{Z})$, a signal defined on bi-infinite time. Then we take it's λ-transform, $\hat{u} = Fu$, to give the function $\hat{u} \in L_2(\mathbb{T})$. Then we apply the multiplication operator M_g, to give $\hat{y} = M_G\hat{u}$, and let $y \in \ell_2(\mathbb{Z})$ the the inverse transform of $\hat{y}$ given by $y = F^*\hat{y}$. The resulting map from u to y is F^*M_GF, which is Toeplitz.

This result also exhibits the *impulse response* g corresponding to the multiplication operator $M_{\hat{g}}$, since if we apply the input

$$e_0 = (\dots, 0, 0, 0, \boxed{1}, 0, 0, 0, \dots)$$

we receive output

$$g = (\dots, g_{-3}, g_{-2}, g_{-1}, \boxed{g_0}, g_1, g_2, g_3, \dots)$$

Further, this result tells us the *induced 2-norm* of $M_{\hat{g}}$, since

$$\begin{aligned} \|F^*M_{\hat{g}}F\| &= \sup_{u\neq 0, u\in\ell_2(\mathbb{Z})} \|F^*M_{\hat{g}}Fu\|_2 \\ &= \sup_{u\neq 0, u\in\ell_2(\mathbb{Z})} \|M_{\hat{g}}Fu\|_2 \\ &= \sup_{\hat{u}\neq 0, \hat{u}\in L_2(\mathbb{T})} \|M_{\hat{g}}\hat{u}\|_2 \\ &= \|M_{\hat{g}}\| \\ &= \|\hat{g}\|_\infty \end{aligned}$$

So the maximum absolute value of the transfer function on the unit circle, which is the maximum of the magnitude Bode plot, is the induced 2-norm of the convolution map, over bi-infinite signals. The proof of Theorem 2 also gives an indication of the worst case signal, which is almost a sinusoid.

We have shown above that every $\hat{g} \in L_\infty(\mathbb{T})$ has a corresponding Toeplitz operator on $\ell_2(\mathbb{Z})$ which is bounded. In fact, the converse is also true; given a bounded Toeplitz operator on $\ell_2(\mathbb{Z})$, there is a corresponding function $g \in L_\infty(\mathbb{T})$ which generates it. We will not need this result in this course.

5.4 Polynomial Transfer Functions

One of the most important results about Fourier series is as follows.

Theorem 3. *Let $\hat{g} : \mathbb{T} \to \mathbb{C}$ be the function $\hat{g}(\lambda) = \lambda$. Then*

$$F^*M_{\hat{g}}F = L$$

In other words, in Fourier coordinates the delay is just multiplication by λ. Hence the choice of notation L for the shift. Now if $\hat{g}$ is a polynomial, the corresponding Toeplitz map is clear. Suppose $\hat{g} : \mathbb{T} \to \mathbb{C}$ is

$$\hat{g}(\lambda) = g_{-n}\lambda^{-n} + g_{-n+1}\lambda^{-n+1} + \cdots + g_{-1}\lambda^{-1} + g_0 + g_1\lambda + \cdots + a^m\lambda^m$$

Then the corresponding Laurent operator is

$$F^*M_{\hat{g}}F = g_{-n}L^{-n} + g_{-n+1}L^{-n+1} + \cdots + g_{-1}L^{-1} + g_0 + g_1L + \cdots + a^mL^m$$

which is

$$F^* M_{\hat{g}} F = \begin{bmatrix} & & & \ddots & & & & & \\ g_m & \cdots & g_1 & g_0 & g_{-1} & \cdots & g_{-n} & & \\ & g_m & \cdots & g_1 & \boxed{g_0} & g_{-1} & \cdots & g_{-n} & \\ & & g_m & \cdots & g_1 & g_0 & g_{-1} & \cdots & g_{-n} \\ & & & & & & \ddots & & \end{bmatrix}$$

a banded matrix. In other words, polynomial transfer functions correspond to banded convolution maps. Such maps are called ***finite-impulse-response*** (FIR) systems.

5.5 Properties of Multiplication Operators

(i) **Inverses**: Suppose $\hat{g}$ is continuous. Then $M_{\hat{g}}$ is invertible if and only if $\hat{g}(e^{j\theta}) \neq 0$ for all θ, and

$$(M_{\hat{g}})^{-1} = M_{\hat{h}}$$

where $h(e^{j\theta}) = 1/g(e^{j\theta})$.

(ii) **Commutativity**: For any $f, g, \in L_\infty(\mathbb{T})$

$$M_{\hat{f}} M_{\hat{g}} = M_{\hat{g}} M_{\hat{f}}$$

Note that this implies that Toeplitz maps commute also.

(iii) **Adjoint**: Suppose $\hat{g} \in L_\infty(\mathbb{T})$ and $M_{\hat{g}} : L_2(\mathbb{T}) \to L_2(\mathbb{T})$ is the corresponding multiplication operator. Define the function $\tilde{g} \in L_\infty(\mathbb{T})$ by

$$\tilde{g}(e^{j\theta}) = \overline{\hat{g}(e^{j\theta})}$$

Then $M_{\hat{g}}^* = M_{\tilde{g}}$. Because then

$$\begin{aligned} \langle y, M_g x \rangle &= \frac{1}{2\pi} \int_0^{2\pi} \overline{y(e^{j\theta})} \hat{g}(e^{j\theta}) x(e^{j\theta}) \, d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} \overline{\tilde{g}(e^{j\theta}) y(e^{j\theta})} x(e^{j\theta}) \, d\theta \\ &= \langle M_{\hat{g}}^* y, x \rangle \end{aligned}$$

In particular, if $\hat{g}$ is the polynomial

$$\hat{g}(\lambda) = g_0 + g_1 \lambda + \cdots + g_n \lambda^n$$

then

$$\tilde{g}(\lambda) = \overline{g_0} + \overline{g_1} \lambda^{-1} + \cdots + \overline{g_n} \lambda^{-n} \tag{1}$$

Notice that for $\lambda \in \mathbb{T}$ we also have

$$\tilde{g}(\lambda) = \tilde{g}(\lambda) = \overline{g_0} + \overline{g_1} \overline{\lambda} + \cdots + \overline{g_n} \overline{\lambda}^n$$

and so this is also correct on the unit circle. However it is not correct on the unit disk, since equation (1) is the Fourier series expansion, and we construct the value of the function on the interior of the unit disk from this.

5.6 Rational Transfer Functions

For a simple rational, suppose $\hat{g}$ is given by

$$\hat{g}(\lambda) = \frac{1}{1 - a\lambda}$$

where $a \in \mathbb{C}$ and $|a| < 1$. Then we know that the inverse Fourier transform of $\hat{g}$ is

$$g = (\dots, 0, 0, 0, \boxed{1}, a, a^2, a^3, \dots)$$

and so the corresponding Toeplitz operator is

$$F^* M_g F = \begin{bmatrix} \ddots & & & & & \\ & 1 & & & & \\ & a^{-1} & \boxed{1} & & & \\ & a^{-2} & a^{-1} & 1 & & \\ & a^{-3} & a^{-2} & a^{-1} & 1 & \\ & & & & & \ddots \end{bmatrix}$$

which is *lower triangular*; i.e., g is the transfer function for a ***causal*** system.

We also know the inverse of this system. Define $\hat{h}$ by

$$\hat{h}(\lambda) = 1 - a\lambda$$

Since $\hat{h} \in L_\infty$, the map $M_{\hat{h}}$ is bounded, and

$$M_{\hat{g}} M_{\hat{h}} = I \qquad M_{\hat{h}} M_{\hat{g}} = I$$

so $M_{\hat{h}} = M_{\hat{g}}^{-1}$. Then the corresponding Toeplitz operators are also inverses of each other, since

$$\begin{aligned} F^* M_{\hat{g}} F &= F^* (M_{\hat{h}})^{-1} F \\ &= (F^* M_{\hat{h}} F)^{-1} \end{aligned}$$

and this means that

$$\begin{bmatrix} \ddots & & & \\ & 1 & & \\ & -a & \boxed{1} & \\ & & -a & 1 \\ & & & & \ddots \end{bmatrix}^{-1} = \begin{bmatrix} \ddots & & & & & \\ & 1 & & & & \\ & a^{-1} & \boxed{1} & & & \\ & a^{-2} & a^{-1} & 1 & & \\ & a^{-3} & a^{-2} & a^{-1} & 1 & \\ & & & & & \ddots \end{bmatrix} \qquad (2)$$

But this *only holds if* $|a| < 1$. Because if $|a| > 1$, then the inverse Fourier transform of $\hat{g}$ is

$$g = (\dots, -a^{-3}, -a^{-2}, -a^{-1}, \boxed{0}, 0, 0, \dots)$$

and so

$$F^* M_g F = \begin{bmatrix} \ddots & & & & & \\ & 0 & -a^{-1} & -a^{-2} & & \\ & & \boxed{0} & -a^{-1} & -a^{-2} & \\ & & & 0 & -a^{-1} & \\ & & & & 0 & \\ & & & & & \ddots \end{bmatrix}$$

which means

$$\begin{bmatrix} \ddots & & & \\ & 1 & & \\ & -a & \boxed{1} & \\ & & -a & 1 \\ & & & & \ddots \end{bmatrix}^{-1} = \begin{bmatrix} \ddots & & & & \\ & 0 & -a^{-1} & -a^{-2} & \\ & & \boxed{0} & -a^{-1} & -a^{-2} \\ & & & 0 & -a^{-1} \\ & & & & \ddots \end{bmatrix} \qquad (3)$$

So which of equations (2) and (3) holds depends on the magnitude of a. If $|a| < 1$ then

$$\frac{1}{1 - a\lambda}$$

has its pole outside the disk, and the corresponding Toeplitz operator is *causal*. On the other hand, if $|a| > 1$ then the pole is inside the disk, and the corresponding Toeplitz operator is *anti-causal*. Notice the bizarre feature that the inverse of a lower-triangular operator is upper triangular! This never happens in finite dimensions.

Algebraically, both of these equations are correct. And in control theory are used to (2) as the convolution map corresponding to $1/(1 - a\lambda)$ no matter what the magnitude of a; if $|a| < 1$ we simply say that the system is unstable. But on ℓ_2, we cannot have exponentially growing signals. In signal processing applications, if causality does not matter then we want to use the position of the poles to determine the causality of the map, not the stability. This is precisely the mathematics that is forced upon us by viewing transfer functions as multiplication operators on $\ell_2(\mathbb{Z})$.

There are other theories of transfer functions which allow us to interpret $1/(1 - a\lambda)$ as causal no matter what the magnitude of a. These are useful for studying unstable systems, but make it substantially harded to analyze the 2-norm and make use of least-squares techniques.

We can now find the corresponding Laurent map for any rational transfer function g; we simply use partial fractions.

5.7 More on Causality

We'll give another way to see the above. First, we need a useful result.

Theorem 4. *Suppose $A \in L(V, V)$ is a bounded linear map on the Banach space V, and $\|A\| < 1$. Then the series*

$$I + A + A^2 + A^3 + \cdots$$

converges, the operator $I - A$ is invertible, and

$$(I - A)^{-1} = I + A + A^2 + A^3$$

Proof. Let S_n be the partial sum

$$S_n = \sum_{k=0}^{n} A^k$$

Then for $m > n$ we have

$$\begin{aligned}
\|S_m - S_n\| &= \left\| \sum_{k=n+1}^{m} A^k \right\| \\
&\le \sum_{k=n+1}^{m} \|A^k\| \\
&\le \sum_{k=n+1}^{m} \|A\|^k \\
&= \|A\|^{n+1} \left(\frac{1 - \|A\|^{m-n}}{1 - \|A\|} \right) \\
&\le \frac{\|A\|^{n+1}}{1 - \|A\|}
\end{aligned}$$

which tends to zero as $n \to \infty$. Hence S is a Cauchy sequence, and since $L(V, V)$ is a Banach space this implies that the series converges.

The remainder of the lemma now follows, since composition with $I - A$ is continuous, and hence we can apply it term by term to show

$$(I - A)(I + A + A^2 + A^3 + \cdots) = I$$

and

$$(I + A + A^2 + A^3 + \cdots)(I - A) = I$$

as required. ■

Now consider again our pair of functions

$$\hat{g}(\lambda) = \frac{1}{1 - a\lambda} \qquad \hat{h}(\lambda) = 1 - a\lambda$$

We know that,

$$\begin{aligned}
F^* M_{\hat{g}} F &= (F^* M_{\hat{h}} F)^{-1} \\
&= (I + aL)^{-1}
\end{aligned}$$

and so if $|a| < 1$ then

$$F^* M_{\hat{g}} F = I + aL + a^2 L^2 + a^3 L^3 + \dots$$

from Theorem 4. So we have

$$
\begin{aligned}
F^* M_g F &= I + aL + a^2 L^2 + \dots \\
&= \begin{bmatrix}
\ddots & & & & & \\
& 1 & & & & \\
& a^{-1} & \boxed{1} & & & \\
& a^{-2} & a^{-1} & 1 & & \\
& a^{-3} & a^{-2} & a^{-1} & 1 & \\
& & & & & \ddots
\end{bmatrix}
\end{aligned}
$$

Or alternatively, if $|a| > 1$ then

$$
\begin{aligned}
(I + aL)^{-1} &= a^{-1} L^{-1} (I + a^{-1} L^{-1})^{-1} \\
&= a^{-1} L^{-1} (I + a^{-1} L^{-1} + a^{-2} L^{-2} + \dots) \\
&= \begin{bmatrix}
\ddots & & & & \\
& 0 & -a^{-1} & -a^{-2} & \\
& & \boxed{0} & -a^{-1} & -a^{-2} \\
& & & 0 & -a^{-1} \\
& & & & \ddots
\end{bmatrix}
\end{aligned}
$$