6 Products and Inverses of Toeplitz Operators

6.1 Invertibility

First, we have two very simple but important facts. First is that uniqueness of solutions is the same as emptiness of the null space. As we will see, this is almost equivalent to left-invertibility.

Lemma 1. *Suppose X and Y are vector spaces and $A : X \to Y$ is linear, and suppose* $\operatorname{null} A = \{0\}$*. Then for each $y \in \operatorname{range} A$, the vector $x \in X$ such that $y = Ax$ is unique.*

Proof. Suppose not; then there exists distinct x_1 and x_2 such that $y = Ax_1 = Ax_2$, so $A(x_1 - x_2) = 0$, so $x_1 - x_2 \in \operatorname{null} A$ which is a contradiction. ■

This immediately gives us a function $b : Y \to X$ such that

$$b(Ax) = x \qquad \text{for all } x \in X$$

which we can conveniently write as

$$b \circ A = I$$

But so far, nothing guarantees that this function b is linear or bounded.

And the second simple fact is that right invertibility is the same as existence of solutions is the same as fullness of the range space.

Lemma 2. *Suppose X and Y are vector spaces and $A : X \to Y$ is linear, and suppose* $\operatorname{range} A = Y$*. Then for each $y \in Y$ there exists $x \in X$ such that $y = Ax$.*

Proof. This is just by definition of the range. ■

Lemma 3. *Suppose X and Y are vector spaces and $A : X \to Y$ is linear, and suppose* $\operatorname{range} A = Y$ *and* $\operatorname{null} A = \{0\}$*. Then there exists a* ***unique*** *map $D : Y \to X$ such that*

$$DA = I \qquad AD = I$$

Further, D is linear.

Proof. From the previous two lemmas we know that there exists b and c such that

$$b \circ A = I \qquad A \circ c = I$$

and therefore both of the following hold

$$b \circ A \circ c = c \qquad b \circ A \circ c = b$$

Hence $b = c$. Call this function d. To see that it is linear, suppose $Ax_1 = y_1$ and $Ax_2 = y_2$. Then we know

$$A(\alpha x_1 + \beta x_2) = \alpha y_1 + \beta y_2$$

and so

$$\alpha x_1 + \beta x_2 = d(\alpha y_1 + \beta y_2)$$

and so

$$\alpha d(y_1) + \beta d(y_2) = d(\alpha y_1 + \beta y_2)$$

which means d is linear. ■

A is called ***invertible*** if $\operatorname{range} A = Y$ and $\operatorname{null} A = \{0\}$. We have the following important result, called the ***Banach Inverse Theorem***.

Theorem 4. *Suppose X and Y are Banach spaces, and $A : X \to Y$ is linear, bounded, and invertible. Then A^{-1} is bounded.*

Proof. The proof requires more machinery than we'll cover in these notes (the Baire Category theorem). See Gohberg p. 282 or Luenberger p. 149. ■

6.1.1 Linearity and Boundedness of One-Sided Inverses

In this section we need to work on a Hilbert space. You do not need to remember the proofs of this section, but the main results are useful. First, we'll state a result about existence of projectors.

Theorem 5. *Suppose H is a Hilbert space, and $S \subset H$ is a* ***closed*** *subspace. Then there exists a bounded linear map P such that*

$$P^2 = P \qquad \operatorname{range} P = S \qquad \operatorname{null} P = S^\perp$$

Proof. Again we need the Baire Category theorem. See Gohberg p. 286. ■

Now we can use this to give conditions under which the left inverse of a linear map with empty null space is linear and bounded.

Lemma 6. *Suppose X and Y are Hilbert spaces, and $A : X \to Y$ is a bounded linear map, with* $\operatorname{null} A = \{0\}$ *and* $\operatorname{range} A$ ***closed****. Then there exists a* bounded linear *map L such that $LA = I$.*

Proof. Using the previous lemma, let P be the projector $P : Y \to Y$ with

$$\operatorname{range} P = \operatorname{range} A \qquad \operatorname{null} P = (\operatorname{range} A)^\perp$$

And define B to be the linear map $B : X \to \operatorname{range} A$ which is just A with restricted domain. Since $\operatorname{range} A$ is closed, it is a Hilbert space, and so the Banach inverse theorem implies that B has a bounded inverse. Then let $L = B^{-1}P$. ■

A map $A : X \to Y$ is called ***bounded below*** if there exists $m > 0$ such that for all $x \in X$

$$\|Ax\| \geq m\|x\|$$

This gives a simple sufficient condition for the range to be closed.

Lemma 7. *Suppose X and Y are Hilbert spaces, and $A : X \to Y$ is a bounded linear map. Then A has a closed range and* $\operatorname{null} A = \{0\}$ *if and only if it is bounded below.*

In other words, the operators that are bounded below are precisely those that are left invertible. The proof is left as an exercise. For right invertibility, the situation is slightly easier.

Lemma 8. *Suppose X and Y are Hilbert spaces, and $A : X \to Y$ is a bounded linear map, with* $\operatorname{range} A = Y$*. Then there exists a bounded linear map R such that $AR = I$.*

Proof. Define $B : (\operatorname{null} A)^\perp \to Y$ to be the restriction of A to $(\operatorname{null} A)^\perp$. We will show that this map is invertible. First notice that $\operatorname{null} B = \{0\}$, since if $Bx = 0$ then we have $Ax = 0$ and so $x \in \operatorname{null} A$. But $\operatorname{null} A \cap (\operatorname{null} A)^\perp = \{0\}$, so this is impossible.

And we need to show that $\operatorname{range} B = Y$. Given any $y \in Y$, we can construct $x \in (\operatorname{null} A)^\perp$ as follows. Using the above lemma on existence of projectors, let P be a projector such that

$$\operatorname{range} P = \operatorname{null} A \qquad \operatorname{null} P = (\operatorname{null} A)^\perp$$

Then $A = A(I - P) + AP$, and since $AP = 0$ we know

$$A = A(I - P)$$

Now pick x such that $y = Ax$, and let $\hat{x} = (I - P)x$. Then $\hat{x} \in (\operatorname{null} A)^\perp$ and $A\hat{x} = y$. So $\operatorname{range} B = Y$ and $\operatorname{null} B = \{0\}$, and since $(\operatorname{null} A)^\perp$ is closed it is a Hilbert space, and so B is invertible. So let $R = B^{-1}$, then $AR = I$ as desired. ■

6.2 Toeplitz Operators on Semi-infinite Time

Let $S : \ell_2(\mathbb{Z}_+) \to \ell_2(\mathbb{Z}_+)$ be the forward shift operator

$$S = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & 1 & 0 & \\ \vdots & & & \ddots \end{bmatrix}$$

Then, just as in the bi-infinite time case, a linear map $A : \ell_2(\mathbb{Z}_+) \to \ell_2(\mathbb{Z}_+)$ is ***Teoplitz*** if

$$S^* A S = A$$

Notice that on bi-infinite time, the shift operator L is invertible, and in fact unitary. However on $\ell_2(\mathbb{Z}_+)$ the shift operator S is not invertible, although since $S^*S = I$ it is an isometry and is left invertible.

Given a Toeplitz map B on $\ell_2(\mathbb{Z})$, we can construct a Toeplitz map on $\ell_2(\mathbb{Z}_+)$ simply by restriction and projection. For a vector $x \in \ell_2(\mathbb{Z})$, partition it as

$$x = \begin{bmatrix} x_- \\ x_+ \end{bmatrix}$$

where $x_+ \in \ell_2(\mathbb{Z}_+)$ and $x_- \in \ell_2(\mathbb{Z}_-)$, and $\mathbb{Z}_-$ is the set of negative integers. Then if $\hat{g} \in L_\infty(\mathbb{T})$ we have the bi-infinite Toeplitz map L_g given by

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

Then define S_g to be

$$S_g = \begin{bmatrix} g_0 & g_{-1} & & \\ g_1 & g_0 & g_{-1} & \\ g_2 & g_1 & g_0 & \\ \vdots & & & \ddots \end{bmatrix}$$

In other words, we have

$$S_g = \begin{bmatrix} 0 & I \end{bmatrix} L_g \begin{bmatrix} 0 \\ I \end{bmatrix}$$

where the partition is the same as that used for x. We call S_g the (semi-infinite) Toeplitz operator with transfer function g.

For example, suppose $g(\lambda) = a_{-1}\lambda^{-1} + a_0 + a_1\lambda + a_2\lambda^2$. Then the bi-infinite Toeplitz map is

$$L_g = F^* M_{\hat{g}} F = a_{_1} L^{-1} + a_0 I + a_1 L + a_2 L^2$$

and the semi-infinite Toeplitz map is

$$\begin{aligned} S_g &= \begin{bmatrix} 0 & I \end{bmatrix} (a_{_1} L^{-1} + a_0 I + a_1 L + a_2 L^2) \begin{bmatrix} 0 \\ I \end{bmatrix} \\ &= a_{_1} S^* + a_0 I + a_1 S + a_2 S^2 \end{aligned}$$

The induced-norm of a semi-infinite Toeplitz operator is exactly the same as that of the bi-infinite operator, as follows.

Theorem 9. *Suppose $\hat{g} \in L_\infty(\mathbb{T})$, and S_g is the corresponding Toeplitz map. Then*

$$\|S_g\| = \|\hat{g}\|_\infty$$

The proof of this theorem is similar to the proof used for the bi-infinite case. However, we need to be careful to construct a worst case input in $\ell_2(\mathbb{Z}_+)$ instead of in $\ell_2(\mathbb{Z})$. One way to do this is to appropriately shift and truncate the worst case input from $\ell_2(\mathbb{Z})$.

6.3 Causal Toeplitz Maps

We know that if the transfer function $\hat{g} \in L_\infty(\mathbb{T})$ then the corresponding Toeplitz operator S_g is bounded. And from the construction above, we also know that S_g is causal if $\hat{g} \in H_2$. So we'd like to define

$$H_\infty = H_2 \cap L_\infty(\mathbb{T}) \qquad \text{cannot quite do this\dots}$$

We cannot quite do this because there is a distinction between the set H_2 whose elements are analytic functions on the open unit disk and the set $\tilde{H}_2$ whose elements are functions on $\mathbb{T}$ which are Fourier transforms of signals in $\ell_2(\mathbb{Z}_+)$. These two sets correspond, however, and the same correspondence holds if we focus on bounded functions. We define

$$H_\infty = \{\, f : \mathbb{D} \to \mathbb{C} \mid f \text{ is analytic and } \|f\|_\infty \text{ is finite} \,\}$$

where

$$\|f\|_\infty = \sup_{\lambda \in \mathbb{D}} |f(\lambda)|$$

In particular, if $f \in H_\infty$ then $f \in H_2$. Conversely, if $f \in H_2$, let $g \in \tilde{H}_2$ be the function constructed by using the power-series coefficients of f as Fourier coefficients. Then $f \in H_\infty$ if

$$\operatorname*{ess\,sup}_{\lambda \in \mathbb{T}} |g(\lambda)| \qquad \text{is finite}$$

and

$$\|f\|_\infty = \operatorname*{ess\,sup}_{\lambda \in \mathbb{T}} |g(\lambda)|$$

To summarize, the transfer functions in H_∞ are those corresponding to causal and bounded Toeplitz operators.

6.4 Products of Semi-infinite Toeplitz Operators

Semi-infinite Toeplitz operators are not as simple as bi-infinite ones. The product of S_g and S_h is not necessarily Toeplitz, and the inverse of S_g is not necessarily Toeplitz either. However, there are two important cases when the product is Toeplitz.

Lemma 10. *Suppose $g \in L_\infty$ and $h \in H_\infty$. The $S_g S_h$ is Toeplitz, and $S_g S_h = S_{gh}$.*

Proof. Let $Q = \begin{bmatrix} 0 \\ I \end{bmatrix}$. Then we have

$$\begin{aligned}
S_{gh} &= Q^* F^* M_{\hat{g}\hat{h}} F Q \\
&= Q^* F^* M_{\hat{g}} F F^* M_{\hat{h}} F Q \\
&= Q^* \begin{bmatrix} ? & ? \\ ? & S_g \end{bmatrix} \begin{bmatrix} ? & 0 \\ ? & S_h \end{bmatrix} Q \\
&= Q^* \begin{bmatrix} ? & ? \\ ? & S_g S_h \end{bmatrix} Q \\
&= S_g S_h
\end{aligned}$$

where ? denotes irrelevant entries. ■

A very similar argument shows that if S_g is anticausal then $S_g S_h$ is Toeplitz.

One important consequence of these two properties is that we must be careful when translating between transfer functions and semi-infinite Toeplitz operators. The simplest case is when the transfer function is a trigonometric polynomial, such as

$$\hat{g}(\lambda) = a_{-m}\lambda^{-m} + \cdots + a_{-1}\lambda^{-1} + a_0 + a_1\lambda + \cdots + a_n\lambda^n$$

then clearly

$$L_g = a_{-m}L^{-m} + \cdots + a_{-1}L^{-1} + a_0 + a_1 L + \cdots + a_n L^n$$

and simply multiplying on by Q^* on the left and Q on the right gives

$$S_g = a_{-m}(S^*)^m + \cdots + a_{-1}S^* + a_0 + a_1 S + \cdots + a_n S^n$$

But we often write polynomials and rational functions in factored form, such as

$$\begin{aligned}\hat{g}(\lambda) &= (\lambda - 2)(\lambda - 3)\\ &= (1 - 2\lambda^{-1})\lambda(\lambda - 3)\\ &= \lambda(1 - 2\lambda^{-1})(\lambda - 3)\\ &= \lambda^2 - 5\lambda + 6\end{aligned}$$

Of course these are all the same function; but if we simply replace λ by S and λ^{-1} by S^* these may give different linear operators. From the expanded form, we have

$$S_g = S^2 - 5S + 6$$

and the other simple replacements give

$$\begin{aligned}(S - 2)(S - 3) &= S^2 - 5S + 6\\ (1 - 2S^*)S(S - 3) &= S^2 - 5S + 6\\ S(1 - 2S^*)(S - 3) &= S^2 - 5S + 6SS^*\end{aligned}$$

Notice that the last one of these is not even Toeplitz.

The example makes clear that we can only directly replace λ by S and λ^{-1} by S^* in a factorization of $\hat{g}$ if we order the factors so that the anticausal terms appear before the causal ones. Then the above lemma implies that the resulting product will be Toeplitz.

6.5 Inverses of Semi-infinite Toeplitz Operators

We know that $1/(1 - a\lambda)$ may correspond to either a causal or an anticausal Toeplitz map, depending on whether $|a| < 1$.

Examples. We look at the following examples.

(a) Suppose $g(\lambda) = 1 - \alpha\lambda$ and $|\alpha| < 1$. Then

$$S_g = I - \alpha S$$

The Neumann series (Theorem 4 in Section 5) implies that, since $\|\alpha S\| < 1$, the map S_g is invertible and

$$S_g^{-1} = I + \alpha S + \alpha^2 S^2 + \alpha^3 S^3 + \dots$$
$$= \begin{bmatrix} 1 & & & & \\ \alpha & 1 & & & \\ \alpha^2 & \alpha & 1 & & \\ \alpha^3 & \alpha^2 & \alpha & 1 & \\ \vdots & & & & \ddots \end{bmatrix}$$

Therefore if $y \in \ell_2$, there always exists a unique $x \in \ell_2$ such that

$$y = S_g x$$

and since $S_g = I - \alpha S$ this equation is just

$$y_t = \begin{cases} x_t - \alpha x_{t-1} & \text{if } t > 0 \\ x_t & \text{if } t = 0 \end{cases}$$

which is more familiar as the state-space system

$$x_{t+1} = \alpha x_t + y_{t+1} \qquad\qquad x_0 = y_0$$

where y_0 is the initial condition and $y_1, y_2, \dots$ the input. Hence we have shown that if $y \in \ell_2$ then $x \in \ell_2$ also.

(b) Similarly, if $\hat{g}(\lambda) = 1 - \alpha\lambda^{-1}$ and $|\alpha| < 1$ then

$$S_g = I + \alpha S^*$$

and

$$S_g^{-1} = I + \alpha S^* + \alpha^2 (S^*)^2 + \alpha^3 (S^*)^3 + \dots$$
$$= \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \dots \\ & 1 & \alpha & \alpha^2 & \\ & & 1 & \alpha & \\ & & & 1 & \\ & & & & \ddots \end{bmatrix}$$

(c) Suppose $g(\lambda) = 1 - \alpha\lambda$ and $|\alpha| > 1$. Then

$$\begin{aligned} S_g &= I - \alpha S \\ &= -\alpha(I - \alpha^{-1}S^*)S \end{aligned}$$

Since $I - \alpha^{-1}S^*$ is invertible, and S is not invertible, the map S_g is not invertible.

(d) Suppose

$$\hat{g}(\lambda) = \lambda^{-1}(\lambda - \tfrac{1}{2})(\lambda - 3)$$

Then we rearrange this expression so that all terms are of the form $1 - a\lambda$ or $1 - a\lambda^{-1}$ with $|a| < 1$. This gives

$$\hat{g}(\lambda) = -3(1 - \tfrac{1}{2}\lambda^{-1})(1 - \tfrac{1}{3}\lambda)$$

and we order the terms so that the causal terms come last and the anti-causal ones come first, so

$$S_g = -3(I - \tfrac{1}{2}S^*)(I - \tfrac{1}{3}S)$$

Now, as above, both of the factors are invertible, so S_g is invertible and

$$S_g^{-1} = -\tfrac{1}{3}(I - \tfrac{1}{3}S)^{-1}(I - \tfrac{1}{2}S^*)^{-1}$$

However, now we have S_g^{-1} is the product of a causal operator on the left with an anticausal operator on the right, and so S_g^{-1} is *not* Toeplitz.

(e) Suppose

$$\hat{g}(\lambda) = (\lambda - \tfrac{1}{2})(\lambda - 3)$$

Then again writing factors as $1 - a\lambda$ or $1 - a\lambda^{-1}$ with $|a| < 1$ we have

$$\hat{g}(\lambda) = -3(1 - \tfrac{1}{2}\lambda^{-1})\lambda(1 - \tfrac{1}{3}\lambda)$$

Now we do one additional thing; as well as placing causal terms on the right, and anti-causal terms on the left, we treat any λ^k term specially and put it in the middle. Then

$$S_g = -3(1 - \tfrac{1}{2}S^*)S(1 - \tfrac{1}{3}S)$$

This has the form $S_g = ABC$ and since A and C are invertible we know that S_g is invertible if and only if B is invertible. But the middle term is S, which is not invertible, and so S_g is not invertible.

However, S_g is left invertible, because S is, and we have

$$A = -\tfrac{1}{3}(1 - \tfrac{1}{3}S)^{-1}S^*(1 - \tfrac{1}{2}S^*)^{-1}$$

is clearly a left inverse of S_g.

The general approach is therefore as follows. If

$$\hat{g}(\lambda) = \frac{\prod_{i=1}^{n_a}(\lambda - a_i)\prod_{j=1}^{n_b}(\lambda - b_j)}{\prod_{p=1}^{n_c}(\lambda - c_p)\prod_{q=1}^{n_q}(\lambda - d_q)}$$

where $|a_i| < 1$, $|c_i| < 1$, $|b_i| > 1$ and $|d_i| > 1$, then write it as

$$\hat{g}(\lambda) = \hat{g}_1(\lambda)\lambda^k\hat{g}_2(\lambda)$$

where

$$\hat{g}_1(\lambda) = \frac{\prod_{i=1}^{n_a}(1 - a_i\lambda^{-1})}{\prod_{p=1}^{n_c}(1 - c_p\lambda^{-1})}$$

and

$$\hat{g}_2(\lambda) = \frac{\prod_{j=1}^{n_b}(-b_j)(1 - b_j^{-1}\lambda)}{\prod_{q=1}^{n_d}(-d_q)(1 - d_q^{-1}\lambda)}$$

and $k = n_a - n_c$. This number is called the *winding number* of g, and is the number of zeros in $\mathbb{D}$ minus the number of poles in $\mathbb{D}$.

Then S_g is right-invertible if and only if $k \le 0$ and S_g is left-invertible if and only if $k \ge 0$. The operator S_g is invertible if and only if $k = 0$. Note that if $g \ne 0$ then S_g has to be either left or right invertible, (or both).

For the important special case where S_g is causal, then we know all the poles of $\hat{g}$ are outside $\bar{D}$, and so k is the number of zeros in $\mathbb{D}$. Then S_g is always left-invertible, and it's invertible if and only if it has no zeros in $\mathbb{D}$, that is, no unstable zeros.

6.6 Zeros

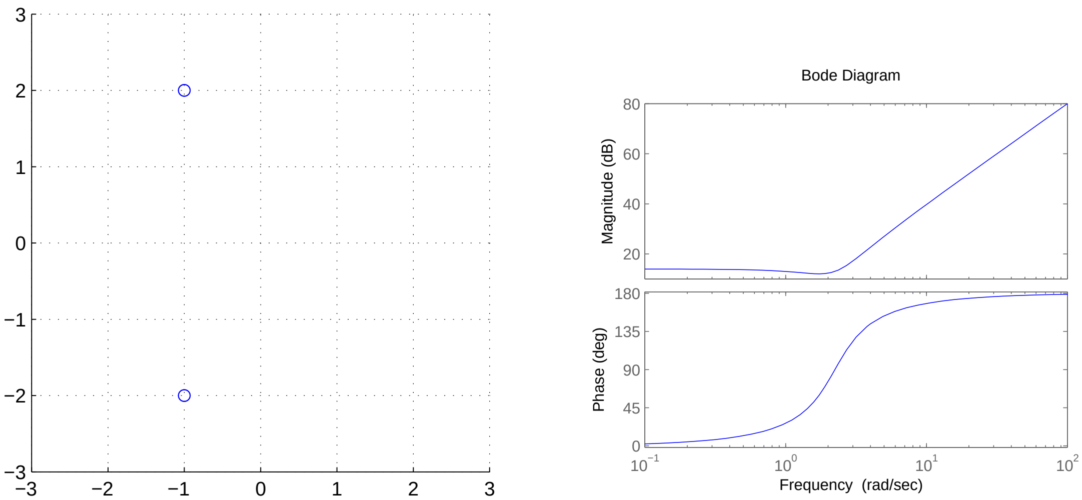
A system with no unstable zeros is called ***minimum phase*** or ***outer***.

Suppose we have a single-input single-output stable state-space system

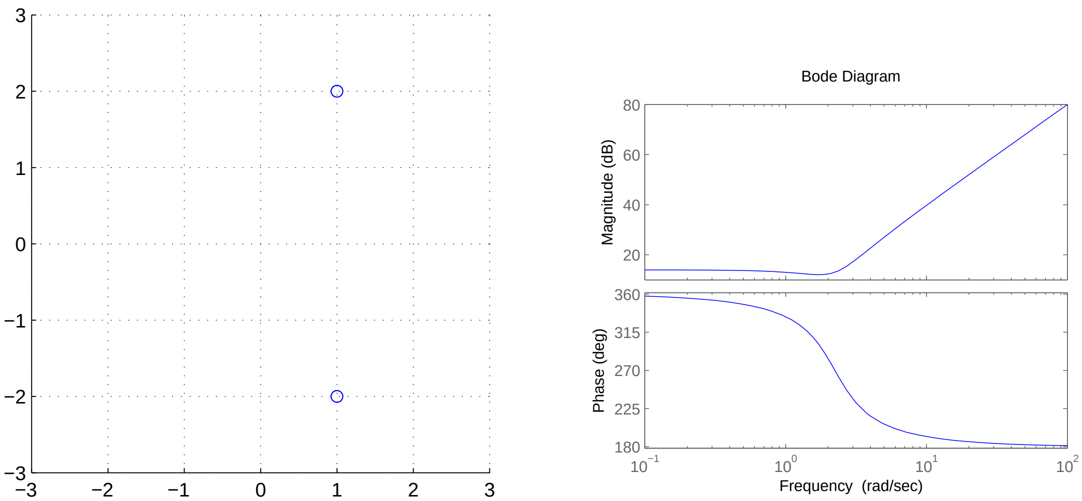
$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t \qquad x_0 = 0 \\ y_t &= Cx_t + Du_t \end{aligned}$$

where $B \in \mathbb{R}^{n\times 1}$ and $C \in \mathbb{R}^{1\times n}$, and if we measure $y_0, y_1, \ldots$. Then if there exists an input $u_0, u_1, \ldots$ generating that output, it is unique, since the corresponding Toeplitz system is left-invertible. And if the system has no unstable zeros, then we can generate any desired sequence $y_0, y_1, \ldots$ by applying the appropriate input. However, if there are unstable zeros, then there are output sequences in ℓ_2 which are unachievable with any input.

The bode plot of $s^2 + 2s + 5$ is below; this is a minimum-phase system, with stable zeros.



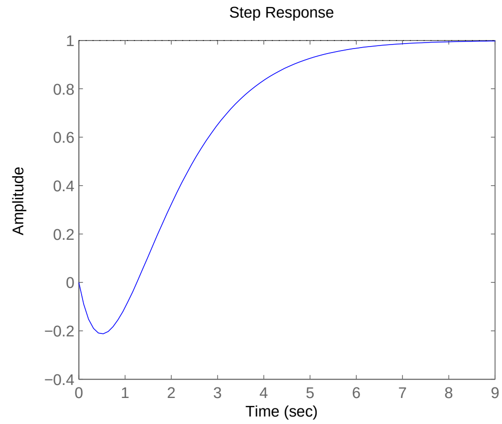
The bode plot of $s^2 - 2s + 5$ is below; this is a non-minimum-phase system, with unstable zeros.



Of all transfer functions with the same magnitude, the one with stable zeros has the smallest phase; i.e. the most positive phase. Of course, phase is an angle, so to understand what this means we need to be a little careful; start by evaluating the phase at $g(j\omega)$ with ω very large. Both systems have almost the same phase. Then as you decrease ω the phase of the min-phase system decreases, and the other one increases.

Note that zeros depend on the actuators and sensors, unlike poles. Some interpretations of unstable zeros are:

- Unstable zeros lead to more negative phase, so often less phase margin, hence the system is often harder to control.
- A system with unstable zeros often has a step response which starts by decreasing away from zero. For example, the step response of $(1 - s)/(s + 1)^2$ is below. This does not always happen, and the analysis is based on a simple second-order system.



- As we will see later in the coure, unstable RHP zeros impose interpolation constraints on stabilizing controllers.
- Unstable RHP zeros mean the system is not invertible as an operator on $\ell_2(\mathbb{Z}_+)$. And inversion is what controllers do.