APPPHYS217 Thursday 29 April 2010

Exponential response

An easy way to derive the relationship between $\Sigma = (A, B, C, D)$ and $H(s)$ is to consider the input-output response of the state-space model to an exponential input,

$$t \geq 0 \, : \, u(t) = \exp(st), \quad s = \sigma + j\omega.$$

Using the standard expression we have learned for the resulting evolution of the state vector $x(t)$,

$$\begin{aligned} x(t) &= \exp(At)\,x(0) + \exp(At)\int_0^t d\tau \exp(-A\tau)B\,u(\tau) \\ &\rightarrow \exp(At)\left\{x(0) + \int_0^t d\tau \exp(-A\tau)B\exp(s\tau)\right\} \\ &= \exp(At)\left\{x(0) + \int_0^t d\tau \exp((sI - A)\tau)B\right\}. \end{aligned}$$

Here we have rearranged terms inside the integrand in a way that looks trivial, but it's worth making note of the implicit assumptions about what commutes with what:

$$\begin{aligned} \exp(-A\tau)B\exp(s\tau) &= \exp(-A\tau)\{\exp(s\tau)I\}B \\ &= \exp(-A\tau)\exp(sI\tau)B \\ &= \exp((sI - A)\tau)B. \end{aligned}$$

At this point we are wondering how to integrate the expression with the matrix exponential. Note that as long as M is invertible, we can write

$$\begin{aligned} \int_0^t d\tau \exp(M\tau) &= \int_0^t d\tau \left\{I + M\tau + \frac{1}{2}(M\tau)^2 + \frac{1}{3!}(M\tau)^3 + \frac{1}{4!}(M\tau)^4 + \cdots\right\} \\ &= \left[\tau I + \frac{1}{2}M\tau^2 + \frac{1}{3!}M^2\tau^3 + \frac{1}{4!}M^3\tau^4 + \frac{1}{5!}M^4\tau^5 + \cdots\right]_0^t \\ &= \left[\tau M + \frac{1}{2}(M\tau)^2 + \frac{1}{3!}(M\tau)^3 + \frac{1}{4!}(M\tau)^4 + \frac{1}{5!}(M\tau)^5 + \cdots\right]_0^t M^{-1} \\ &= [\exp(M\tau) - I]_0^t M^{-1} \\ &= \exp(Mt)M^{-1} - M^{-1}. \end{aligned}$$

In our case we can use this result as long as $(sI - A)$ is invertible, *i.e.*, as long as s is not an eigenvalue of A. Assuming this holds, we can use $M \equiv (sI - A)$ and (leaving B outside the integrand)

$$\begin{aligned} x(t) &= \exp(At)\left\{x(0) + \int_0^t d\tau \exp((sI - A)\tau)B\right\} \\ &= \exp(At)\left\{x(0) + \exp((sI - A)t)(sI - A)^{-1}B - (sI - A)^{-1}B\right\} \\ &= \exp(At)\left\{x(0) - (sI - A)^{-1}B\right\} + \exp(At)\exp((sI - A)t)(sI - A)^{-1}B \\ &= \exp(At)\left\{x(0) - (sI - A)^{-1}B\right\} + \exp(st)(sI - A)^{-1}B \\ &= \exp(At)\left\{x(0) - (sI - A)^{-1}B\right\} + \left\{(sI - A)^{-1}B\right\}\exp(st). \end{aligned}$$

We can now easily compute the output signal,

$$\begin{aligned} y(t) &= Cx(t) + D\exp(st) \\ &= C\exp(At)\left\{x(0) - (sI-A)^{-1}B\right\} + \left\{C(sI-A)^{-1}B\right\}\exp(st) + D\exp(st). \end{aligned}$$

Hence we finally have the result that $y(t)$ corresponds to the free evolution of a modified initial condition (first term in curly-braces) plus a steady-state response that can be written

$$\left\{C(sI-A)^{-1}B + D\right\}\exp(st) \equiv H_{yu}(s)\exp(st).$$

If we assume that the initial condition is carefully set to

$$x(0) = (sI-A)^{-1}B,$$

we can ignore the initial-condition contribution to $y(t)$ and the transfer function completely summarizes the overall steady-state exponential response. Defining

$$\begin{aligned} a_{yu} &\equiv \mathrm{Re}[H_{yu}(s)], \quad b_{yu} \equiv \mathrm{Im}[H_{yu}(s)], \\ g_{yu} &\equiv |H_{yu}(s)| = \sqrt{a_{yu}^2 + b_{yu}^2}, \quad \varphi_{yu} \equiv \tan^{-1}\left(\frac{b_{yu}}{a_{yu}}\right), \end{aligned}$$

we have (as long as s is not an eigenvalue of A)

$$\begin{aligned} y(t) &= H_{yu}(s)\exp(st) \\ &= g_{yu}\exp(st + j\varphi_{yu}). \end{aligned}$$

We thus see that for s not an eigenvalue of A, the Laplace transfer function $H_{yu}(s)$ is very much like a frequency response, but for general exponential functions on $t \geq 0$. Don't forget however that we are making a special assumption about the initial conditions in writing this simple relation. In case $s = j\omega$ we recover the fact that the frequency response (gain and phase) of Σ correspond to the magnitude and phase of the complex-valued function $H_{yu}(j\omega)$. If we restrict our attention to the case of frequency response and A happens to be stable, note that we have more general grounds for ignoring the initial-condition contribution to $y(t)$ as $H_{yu}(j\omega)\exp(j\omega t)$ is guaranteed to dominate for t sufficiently large.

Poles of the transfer function; **input-output stability**

As was mentioned earlier, the input-output transfer function derived from a state-space model can always be written in the form of a rational function (the ratio of two polynomials in s),

$$H_{yu}(s) = \frac{n(s)}{d(s)}.$$

It was noted that the zeros of $d(s)$, which generally correspond to the complex eigenvalues of A, are special values of s called the poles of $H_{yu}(s)$, as it would appear to casual inspection that the transfer function should diverge there. Although the relation $y(t) = H_{yu}(s)\exp(st)$ is only valid for s not an eigenvalue of A (so that $sI - A$ is invertible), one might wonder whether the form of $H_{yu}(s)$ could imply that the input-output system has divergent response to an exponential input $u(t) = \exp(s_p t)$, where s_p is any pole of the transfer function. This would be especially worrisome if we were talking about a pole with negative real part!

Let's examine a SISO system with two-dimensional state space and $D = 0$. Writing

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad C = \begin{bmatrix} c_1 & c_2 \end{bmatrix},$$

we have

$$\begin{aligned} H_{yu}(s) &= C(sI - A)^{-1}B \\ &= \begin{bmatrix} c_1 & c_2 \end{bmatrix} \left(\begin{bmatrix} s - a_{11} & -a_{12} \\ -a_{21} & s - a_{22} \end{bmatrix} \right)^{-1} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \\ &= -\left(\frac{1}{a_{12}a_{21} - (s - a_{11})(s - a_{22})} \right) \begin{bmatrix} c_1 & c_2 \end{bmatrix} \begin{bmatrix} a_{22} - s & -a_{12} \\ -a_{21} & a_{11} - s \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \\ &= -\frac{b_2(-c_1a_{12} + c_2(-s + a_{11})) + b_1(-c_2a_{21} + c_1(-s + a_{22}))}{a_{12}a_{21} - (s^2 - a_{11}s - a_{22}s + a_{11}a_{22})} \\ &= \frac{s(-b_1c_1 - b_2c_2) + b_1c_1a_{22} - b_1c_2a_{21} - b_2c_1a_{12} + b_2c_2a_{11}}{s^2 - s(a_{11} + a_{22}) + a_{11}a_{22} - a_{12}a_{21}}. \end{aligned}$$

As promised this is a rational function, with the order of $n(s)$ less than or equal to the order of $d(s)$. Looking at $d(s)$ we recognize the characteristic polynomial of A,

$$\begin{aligned} \det(A - sI) &= \det\left(\begin{bmatrix} a_{11} - s & a_{12} \\ a_{21} & a_{22} - s \end{bmatrix} \right) \\ &= (a_{11} - s)(a_{22} - s) - a_{12}a_{21} \\ &= s^2 - s(a_{11} + a_{22}) + a_{11}a_{22} - a_{12}a_{21}. \end{aligned}$$

It would seem to follow from this (together with the exponential-response calculation above) that for any s with negative real part, which is ***not*** an eigenvalue of A, then $H_{yu}(s)\exp(st)$ should be bounded. But we are still not entirely sure what happens if s is an eigenvalue of A, stable or unstable.

To check this explicitly let's pick a specific form for the dynamics,

$$A = \begin{bmatrix} r & w \\ w & r \end{bmatrix}, \quad \lambda = r \pm w,$$

which can thus be forced to have real and distinct eigenvalues. Note that the eigenvectors of A are simply

$$r + w \leftrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \equiv v_+, \quad r - w \leftrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \equiv v_-.$$

For simplicity let's assume $x(0) = 0$, and consider the evolution of the state vector in response to $u(t) = \exp(st)$:

$$x(t) = \exp(At)\int_0^t d\tau\, \exp(-A\tau)B\exp(s\tau)$$
$$= \exp(At)\int_0^t d\tau\, \exp((sI - A)\tau)B.$$

If we now assume that $s = r + w$, which is an eigenvalue of A, we cannot use the general expression from above and therefore must explicitly calculate:

$$x(t) = \exp(At)\int_0^t d\tau\, \exp\left(\begin{bmatrix} w & -w \\ -w & w \end{bmatrix}\tau\right)B$$
$$= \exp(At)\int_0^t d\tau \begin{bmatrix} \frac{1}{2}e^{2w\tau} + \frac{1}{2} & -\frac{1}{2}e^{2w\tau} + \frac{1}{2} \\ -\frac{1}{2}e^{2w\tau} + \frac{1}{2} & \frac{1}{2}e^{2w\tau} + \frac{1}{2} \end{bmatrix}B.$$

Here the integral is to be done component-wise, so we just need to evaluate

$$\frac{1}{2}\int_0^t d\tau(1 \pm e^{2w\tau}) = \frac{1}{2}\left[\tau \pm \frac{1}{2w}e^{2w\tau}\right]_0^t$$
$$= \frac{1}{2}\left[t \pm \frac{1}{2w}e^{2wt} \mp \frac{1}{2w}\right].$$

Hence,

$$x(t) = \exp(At)\begin{bmatrix} \frac{1}{2}t + \frac{1}{2w}e^{2wt} - \frac{1}{2w} & \frac{1}{2}t - \frac{1}{2w}e^{2wt} + \frac{1}{2w} \\ \frac{1}{2}t - \frac{1}{2w}e^{2wt} + \frac{1}{2w} & \frac{1}{2}t + \frac{1}{2w}e^{2wt} - \frac{1}{2w} \end{bmatrix}B.$$

Noting

$$\exp(At) = \exp\left(\begin{bmatrix} rt & wt \\ wt & rt \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{2}e^{(r+w)t} + \frac{1}{2}e^{(r-w)t} & \frac{1}{2}e^{(r+w)t} - \frac{1}{2}e^{(r-w)t} \\ \frac{1}{2}e^{(r+w)t} - \frac{1}{2}e^{(r-w)t} & \frac{1}{2}e^{(r+w)t} + \frac{1}{2}e^{(r-w)t} \end{bmatrix},$$

we have

$$x(t) = \frac{1}{4}\begin{bmatrix} e^{(r+w)t} + e^{(r-w)t} & e^{(r+w)t} - e^{(r-w)t} \\ e^{(r+w)t} - e^{(r-w)t} & e^{(r+w)t} + e^{(r-w)t} \end{bmatrix}\begin{bmatrix} t + \frac{1}{w}e^{2wt} - \frac{1}{w} & t - \frac{1}{w}e^{2wt} + \frac{1}{w} \\ t - \frac{1}{w}e^{2wt} + \frac{1}{w} & t + \frac{1}{w}e^{2wt} - \frac{1}{w} \end{bmatrix}B$$
$$= \frac{1}{4}\begin{bmatrix} 2te^{rt}e^{tw} + \frac{2}{w}e^{rt}e^{tw} - \frac{2}{w}\frac{e^{rt}}{e^{tw}} & 2te^{rt}e^{tw} - \frac{2}{w}e^{rt}e^{tw} + \frac{2}{w}\frac{e^{rt}}{e^{tw}} \\ 2te^{rt}e^{tw} - \frac{2}{w}e^{rt}e^{tw} + \frac{2}{w}\frac{e^{rt}}{e^{tw}} & 2te^{rt}e^{tw} + \frac{2}{w}e^{rt}e^{tw} - \frac{2}{w}\frac{e^{rt}}{e^{tw}} \end{bmatrix}B$$
$$= \frac{1}{2}\begin{bmatrix} (t + \frac{1}{w})e^{(r+w)t} - \frac{1}{w}e^{(r-w)t} & (t - \frac{1}{w})e^{(r+w)t} + \frac{1}{w}e^{(r-w)t} \\ (t - \frac{1}{w})e^{(r+w)t} + \frac{1}{w}e^{(r-w)t} & (t + \frac{1}{w})e^{(r+w)t} - \frac{1}{w}e^{(r-w)t} \end{bmatrix}B.$$

For simplicity let's now assume

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \Rightarrow \quad y(t) = \frac{1}{2}\left(t + \frac{1}{w}\right)e^{(r+w)t} - \frac{1}{2w}e^{(r-w)t}.$$

If both $r + w$ and $r - w$ are negative (*e.g.*, $r = -2$, $w = 1$), meaning that both eigenvalues of A are negative, then we see that the output decays to zero as $t \to \infty$. If

$r + w$ is positive (*e.g.*, $r = 2,\ w = 1$), then $y(t)$ will diverge but we are not suprised because in this case $u(t)$ is itself unbounded. However, an odd situation can occur if $r + w$ is negative but the other eigenvalue $r - w$ is positive (*e.g.*, $r = 1,\ w = -2$). In this case we can be driving the system with input

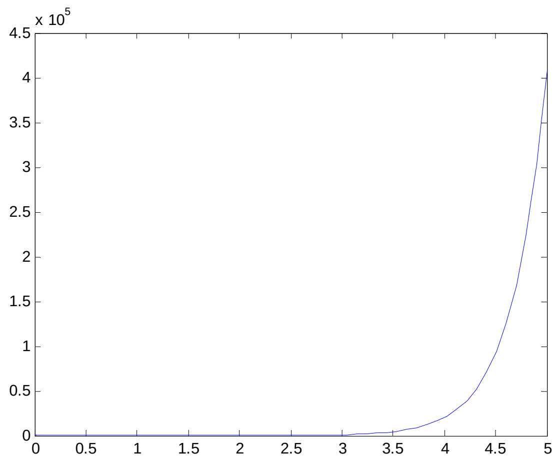
$$u(t) = \exp(st) = \exp((r+w)t) = \exp(-t),$$

whcih is a decaying exponential, and yet

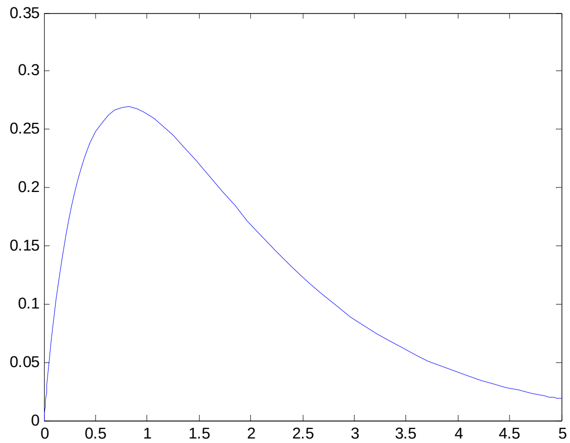
$$\begin{aligned} y(t) &= \frac{1}{2}\left(t + \frac{1}{w}\right)e^{(r+w)t} - \frac{1}{2w}e^{(r-w)t} \\ &= \frac{1}{2}\left(t + \frac{1}{w}\right)e^{-t} - \frac{1}{2w}e^{3t}, \end{aligned}$$

which clearly diverges as $t \to \infty$. This is therefore an example of a state-space system for which $y(t)$ is unbounded even for a bounded $u(t)$, and the divergence is clearly associated with the unstable eigenvalue of A.

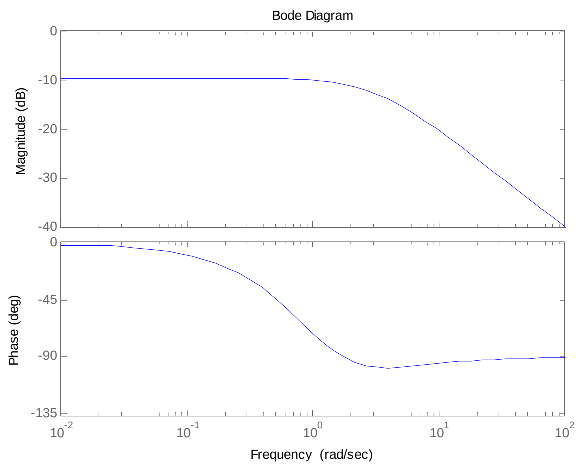
This result seems rather counter-intuitive, so let's try to convince ourselves that it's real (we didn't just make an algebra mistake). For example, we can use Matlab's **ODE45** integrator to solve an appropriate initial-condition problem. For $(r, w) = (1, -2)$:



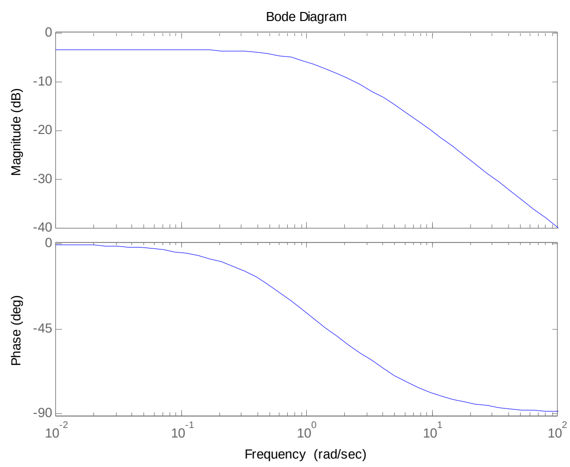
(This unstable system also has divergent impulse and step responses). But for $(r, w) = (-2, 1)$:



Note that if we check the frequency response of the $(r, w) = (1, -2)$ case:



and of the $(r, w) = (-2, 1)$ case:



Both of these look innocent enough, but the rather severe instability does not translate into any divergences in the frequency response input-output gain. This illustrates a very important point about linear input-output systems: 'pointwise' consideration of the frequency response doesn't tell you everything, and not all qualitative properties are reflected in an obvious way in the Bode plot. Apparently some important features are more subtly encoded in the 'shape' of the Bode plot over a range of frequencies.

In our example we haven't been sufficiently general to prove this, but it is a fact that asymptotically stable linear systems (all eigenvalues of A have negative real part) have bounded response to bounded inputs. Hence one only needs to worry about this sort of input-output instability when A has some unstable eigenvalues.

Earlier in the term we learned to associate eigenvalues of the A matrix with positive real part with instability. Indeed, continuing with our simple example with

$$A = \begin{bmatrix} r & w \\ w & r \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix},$$

$$r + w \leftrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \equiv v_+, \quad r - w \leftrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \equiv v_-,$$

in the absence of any control input u we have that an initial state

$$\begin{aligned} x(0) &= \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \frac{x_1(0)}{\sqrt{2}} \{v_+ + v_-\} + \frac{x_2(0)}{\sqrt{2}} \{v_+ - v_-\} \\ &= \frac{1}{\sqrt{2}} \{x_1(0) + x_2(0)\} v_+ + \frac{1}{\sqrt{2}} \{x_1(0) - x_2(0)\} v_-, \end{aligned}$$

will evolve as

$$\begin{aligned} x(t) &= \exp(At)x(0) \\ &= \frac{1}{\sqrt{2}} \{x_1(0) + x_2(0)\} e^{(r+w)t} v_+ + \frac{1}{\sqrt{2}} \{x_1(0) - x_2(0)\} e^{(r-w)t} v_-. \end{aligned}$$

We thus see that in general, unless $x(0)$ is equal to one of the eigenvectors, both eigenvalues will appear in the evolution. Hence in situations where even just one of the eigenvalues is positive, for example $r + w$, we would need very special conditions in order to avoid exponential divergence ($x(0)$ would need to be exactly proportional to the stable eigenvector).

Returning to the issue of exponential-input response, we noted above that we can use $y(t) = H_{yu}(-\alpha)\exp(-\alpha t)$ as long as $-\alpha$ is not an eigenvalue of A and neglecting the (modified) initial-value response. Unless all the eigenvalues of A have real parts that are more negative than $-\alpha$, we can only neglect the initial-value response if we make the specific choice

$$x(0) = (sI - A)^{-1}B.$$

In order to understand this let us consider the exponential-input response in the context of our current example system,

$$x(t) = \exp(At)\int_0^t d\tau \exp(-A\tau)Bu(\tau),$$

where

$$\begin{aligned}
\exp(-A\tau)B &= \begin{bmatrix} \frac{1}{2}e^{-(r+w)\tau} + \frac{1}{2}e^{-(r-w)\tau} & \frac{1}{2}e^{-(r+w)\tau} - \frac{1}{2}e^{-(r-w)\tau} \\ \frac{1}{2}e^{-(r+w)\tau} - \frac{1}{2}e^{-(r-w)\tau} & \frac{1}{2}e^{-(r+w)\tau} + \frac{1}{2}e^{-(r-w)\tau} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{2}e^{-\tau(r+w)} + \frac{1}{2}e^{-\tau(r-w)} \\ \frac{1}{2}e^{-\tau(r+w)} - \frac{1}{2}e^{-\tau(r-w)} \end{bmatrix} \\
&= \frac{1}{\sqrt{2}}e^{-\tau(r+w)}v_+ + \frac{1}{\sqrt{2}}e^{-\tau(r-w)}v_-,
\end{aligned}$$

and thus

$$\exp(At)\exp(-A\tau)Bu(\tau) = \left\{\frac{1}{\sqrt{2}}e^{(t-\tau)(r+w)}v_+ + \frac{1}{\sqrt{2}}e^{(t-\tau)(r-w)}v_-\right\}u(\tau).$$

If we now consider $u(t) = \exp(-\alpha t)$ with $-\alpha$ not equal to either of the eigenvalues $r + w$ or $r - w$,

$$\begin{aligned}
x(t) &= \left\{\frac{1}{\sqrt{2}}e^{t(r+w)}\int_0^t d\tau\, e^{-\tau(r+w)}e^{-\alpha\tau}\right\}v_+ + \left\{\frac{1}{\sqrt{2}}e^{t(r-w)}\int_0^t d\tau\, e^{-\tau(r-w)}e^{-\alpha\tau}\right\}v_- \\
&= \left\{\frac{1}{\sqrt{2}}e^{t(r+w)}\frac{1}{-\alpha-r-w}\left[e^{-t(\alpha+r+w)} - 1\right]\right\}v_+ + \left\{\frac{1}{\sqrt{2}}e^{t(r-w)}\frac{1}{-\alpha-r+w}\left[e^{-t(\alpha+r-w)} - 1\right]\right\}v_- \\
&= -\frac{e^{-\alpha t} - e^{t(r+w)}}{\sqrt{2}\,(\alpha+r+w)}v_+ - \frac{e^{-\alpha t} - e^{t(r-w)}}{\sqrt{2}\,(\alpha+r-w)}v_- \\
&= -\frac{e^{-\alpha t}}{\sqrt{2}}\left[\frac{v_+}{\alpha+r+w} + \frac{v_-}{\alpha+r-w}\right] + \frac{1}{\sqrt{2}}\left[\frac{e^{t(r+w)}}{\alpha+r+w}v_+ + \frac{e^{t(r-w)}}{\alpha+r-w}v_-\right].
\end{aligned}$$

If one of the eigenvalues has positive real part we see that there is a diverging contribution to $x(t)$. Recall however that we are supposed to set

$$\begin{aligned}
x(0) = (sI - A)^{-1}B &\to \begin{bmatrix} -\alpha - r & -w \\ -w & -\alpha - r \end{bmatrix}^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \frac{1}{(\alpha+r)^2 - w^2}\begin{bmatrix} -\alpha - r & w \\ w & -\alpha - r \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \frac{1}{(\alpha+r)^2 - w^2}\begin{bmatrix} -\alpha - r \\ w \end{bmatrix},
\end{aligned}$$

which decomposes as

$$
\begin{aligned}
x(0) &= \frac{1}{(\alpha+r)^2 - w^2}\left\{-\frac{\alpha+r}{\sqrt{2}}(v_+ + v_-) + \frac{w}{\sqrt{2}}(v_+ - v_-)\right\} \\
&= \frac{1}{(\alpha+r)^2 - w^2}\left\{-\frac{1}{\sqrt{2}}(\alpha+r-w)v_+ - \frac{1}{\sqrt{2}}(\alpha+r+w)v_-\right\} \\
&= -\frac{1}{\sqrt{2}\,(\alpha+r+w)}v_+ - \frac{1}{\sqrt{2}\,(\alpha+r-w)}v_-.
\end{aligned}
$$

Hence with this special choice of initial state $\exp(At)x(0)$ will exactly cancel out the part of the exponential response not proportional to $\exp(-\alpha t)$, including any possible divergent contribution.

To sum up what we have learned from our simple example, the simple exponential-response relation $y(t) = H_{yu}(s)\exp(st)$ holds in the limit of large t when s is not an eigenvalue of A and the eigenvalues of A all have real part more negative than the real part of s. It can also hold for arbitrary t when the special initial condition $x(0) = (sI - A)^{-1}B$ is chosen; this special choice cancels terms in the response to $u(t) = \exp(st)$ that are not proportional to $\exp(st)$. Arranging for this cancellation is especially important when A has some unstable eigenvalues. The response to $u(t) = \exp(s_p t)$ where s_p is an eigenvalue of A can be explicitly computed. Instability of an input-output system is not necessarily reflected in any divergences of its frequency response.

Closed-loop poles with feedback

Earlier we saw that the feedback system

$$
\begin{aligned}
y_1 &= H_1 u_1, \\
u_2 &= y_1, \quad y_2 = H_2 u_2, \\
u_1 &= r - y_2,
\end{aligned}
$$

has overall transfer function

$$
H_{y_1 r} = \frac{H_1}{1 + H_1 H_2}.
$$

Let's set $H_2 = 1$ (unit feedback) and look at the poles that result. At the beginning of this set of notes we did a fairly general calculation

$$
A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad C = \begin{bmatrix} c_1 & c_2 \end{bmatrix}, \quad D = 0,
$$

$$
H_{yu}(s) = \frac{s(-b_1c_1 - b_2c_2) + b_1c_1a_{22} - b_1c_2a_{21} - b_2c_1a_{12} + b_2c_2a_{11}}{s^2 - s(a_{11} + a_{22}) + a_{11}a_{22} - a_{12}a_{21}}.
$$

For simplicity let's set $a_{11} = a_{22} = r,\ a_{12} = a_{21} = w,\ b_1 = c_1 = 1,\ b_2 = c_2 = 0,$ (which corresponds to the example we just looked at) and call the result $H_1(s).$ Then

$$H_1(s) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} s-r & -w \\ -w & s-r \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \frac{1}{(s-r)^2 - w^2} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} s-r & w \\ w & s-r \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \frac{s-r}{s^2 - 2rs + r^2 - w^2}$$

$$= \frac{s-r}{(s-r+w)(s-r-w)},$$

and

$$H_{y_1 r} \to \frac{s-r}{s^2 - 2rs + r^2 - w^2\left(1 + \frac{s-r}{s^2-2ws+r^2-w^2}\right)}$$

$$= \frac{s-r}{s^2 - 2rs + r^2 - w^2 + s - r}$$

$$= \frac{s-r}{s^2 + s(1-2r) + r^2 - w^2 - r}.$$

The denominator has roots

$$s = \frac{(2r-1) \pm \sqrt{1 - 4r + 4r^2 - 4r^2 + 4w^2 + 4r}}{2}$$

$$= \frac{(2r-1) \pm \sqrt{4w^2 + 1}}{2}.$$

If we check the cases that we have previously considered,

$$(r,w) = (-2,1) \to s = \frac{-5 \pm \sqrt{5}}{2},$$

$$(r,w) = (1,-2) \to s = \frac{1 \pm \sqrt{17}}{2},$$

$$(r,w) = (2,1) \to s = \frac{3 \pm \sqrt{5}}{2}.$$

Clearly these poles are shifted relative to the case without feedback $(r \pm w)$. Which systems to you expect to exhibit unstable behavior?

Note that we could model the sort of feedback system considered above directly in the state-space model. Generalizing to proportional feedback with gain K,

$$\dot{x} = Ax + Bu = Ax + B(r - Ky) = Ax + B(r - KCx)$$

$$= (A - KBC)x + Br.$$

This has effective dynamical matrix

$$A' = \begin{bmatrix} r-K & w \\ w & r \end{bmatrix},$$

whose eigenvalues are

$$\lambda = \frac{(2r-K) \pm \sqrt{4w^2+K^2}}{2},$$

in agreement of course with the transfer-function calculation we performed with $K = 1$. It would appear that it is not possible simply by choice of K to turn a general unstable (r,w) instance into a stable closed-loop system. Note that the Matlab function **`rlocus`** provides a convenient tool for tracking the changes in the poles and zeros of a system under proportional feedback.

Note that the reachability and observability matrices here are

$$W_r = \begin{bmatrix} 1 & r \\ 0 & w \end{bmatrix}, \quad W_o = \begin{bmatrix} 1 & 0 \\ r & w \end{bmatrix},$$

which in general will both be full-rank, so it should be the case that we can use feedback to stabilize the system; we just can't do it via proportional feedback. From the general form of the H_1 transfer function,

$$H_1(s) = \frac{s-r}{(s-r+w)(s-r-w)},$$

we have the general closed-loop transfer function

$$H_{cl}(s) = \frac{s-r}{(s-r+w)(s-r-w)+(s-r)H_2(s)}.$$

If we try a simple form

$$H_2(s) = \frac{(s+a)(s+b)}{s-r},$$

we have the new denominator polynomial

$$\begin{aligned} d_{cl}(s) &= (s-r+w)(s-r-w)+(s+a)(s+b) \\ &= 2s^2+(a+b-2r)s+r^2-w^2+ab, \end{aligned}$$

with solutions

$$\begin{aligned} \lambda_{cl} &= \frac{(2r-a-b) \pm \sqrt{(a+b-2r)^2-8(r^2-w^2+ab)}}{4} \\ &= \frac{(2r-a-b) \pm \sqrt{(a+b-2r)^2-8(r^2-w^2+ab)}}{4}. \end{aligned}$$

If we try $a = r,\ b = r+1,$ then

$$\lambda_{cl} \to \frac{-1 \pm \sqrt{1-8(2r^2+r-w^2)}}{4}.$$

In the case where $(r,w) = (2,1)$ (both open-loop eigenvalues positive) we thus have $a = 2,\ b = 3$ and

$$\lambda_{cl} = \frac{-1 \pm \sqrt{1-8(8+2-1)}}{4} = \frac{-1 \pm i\sqrt{71}}{4},$$

which is stable. We should note however that we did one not-so-robust thing in our design, which was to add a RHP pole in the controller transfer function $H_2(s)$ that cancels the RHP zero in $H_1(s)$. While technically valid, in practice this is a very bad idea because if the pole of the controller as implemented varies just a little bit from the actual zero of the plant, the closed-loop system will be left with an uncancelled unstable pole. For example if we set

$$H_2(s) = \frac{(s+2)(s+3)}{s-2.1},$$

we end up with a closed-loop transfer function with a RHP pole at $s \approx 1.995$.

We could instead try a separated design. If we use pole-placement to design a Luenberger observer with eigenvalues $(-5,-6)$ we obtain

$$L = \begin{bmatrix} 15 \\ 57 \end{bmatrix},$$

working still with our $(r,w) = (2,1)$ instance, and if we try to design a state-feedback law to obtain final eigenvalues $(-1,-2)$ we obtain

$$K = \begin{bmatrix} 7 & 13 \end{bmatrix}.$$

Checking the latter design first we note that

$$\begin{aligned} A - BK &= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 7 & 13 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 13 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -5 & -12 \\ 1 & 2 \end{bmatrix}, \end{aligned}$$

which indeed has eigenvalues $(-1,-2)$ under state-feedback. As for the observer design,

$$\begin{aligned} A - LC &= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 15 \\ 57 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 15 & 0 \\ 57 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -13 & 1 \\ -56 & 2 \end{bmatrix}, \end{aligned}$$

which indeed has eigenvalues $(-5,-6)$. Can we put these into a transfer-function notation?

For the combined observer-controller system we will use $u = -K\hat{x},$ where $\hat{x}$ is the estimate produced by the observer. For the observer, we have

$$\frac{d}{dt}\hat{x} = A\hat{x} + Bu + L(y - C\hat{x}),$$

where y is the output of the plant and $u = -K\hat{x},$ so

$$\frac{d}{dt}\hat{x} = A\hat{x} - BK\hat{x} + L(y - C\hat{x})$$
$$= (A - BK - LC)\hat{x} + Ly.$$

For the observer transfer function we think of y as the input and we can set $K\hat{x}$ as the output to send directly back to the plant. Hence the separated controller is a state-space system with

$$\hat{A} = A - BK - LC \to \begin{bmatrix} -20 & -12 \\ -56 & 2 \end{bmatrix}, \quad \hat{B} = L, \quad \hat{C} = K,$$

which has transfer function $C(sI - A)^{-1}B$ equal to

$$H_2(s) = \frac{846s - 1098}{s^2 + 18s - 712} \approx \frac{846(s - 1.298)}{(s + 37.16)(s - 19.16)}.$$

Feedback interconnection with $H_1(s)$ leads to

$$H_{cl}(s) = \frac{s^3 + 16s^2 - 748s + 1424}{(s + 6)(s + 5)(s + 2)(s + 1)},$$

which is indeed stable. There does not appear to be any attempt to cancel the RHP zero of $H_1(s)$ in this design.