ENAE 301:

$$\xrightarrow{\text{force}} \boxed{\text{Dynamics}} \xrightarrow{\text{motion}}$$

More generally:

$$\xrightarrow{\text{input}} \boxed{\text{System}} \xrightarrow{\text{output}}$$

ENAE 301:

$$\xrightarrow{\text{force}} \boxed{\text{Dynamics}} \xrightarrow{\text{motion}}$$

More generally:

$$\xrightarrow[u(t)]{\text{input}} \boxed{\text{System}} \xrightarrow[y(t)]{\text{output}}$$

A (dynamic) system "transforms" inputs $u(t)$ into outputs $y(t)$.

We must first understand as completely as possible this "transformation".

Simple Hovercraft Example

$$\frac{d}{dt}(mv) = f \quad \text{(dynamics)}$$

$$\frac{d}{dt}(y) = v \quad \text{(Kinematics)}$$

Where:

m = mass (assume constant)
v = velocity
y = position
f = applied force

Thus:

$$\left.\begin{aligned} m\frac{dv}{dt} &= f \\ \frac{dy}{dt} &= v \end{aligned}\right\} \text{Governing DE}$$

Equivalently: $m\dfrac{d^2y}{dt^2} = f$

Driving Force

Force driving system is due to fan:

$$f \approx K_f \, \omega$$

where K_f constant, ω is rotation rate of fan

Similarly: $\omega \approx K_m V_m$

where K_m constant, V_m is voltage applied to motor

Then $\boxed{\ddot{y}(t) = K\,u(t)}$, $K = \left[\dfrac{K_f K_m}{m}\right]$

treating $V_m(t) = u(t)$ as the input to the system

$$\xrightarrow{u(t)} \boxed{\text{System}} \xrightarrow{y(t)}$$

Analysis:

Given:

$$\boxed{\begin{aligned}\dot{v}(t) &= K u(t) \\ \dot{y}(t) &= v(t)\end{aligned}}$$

($v(t)$ velocity)

Then:

$$v(t) = v_0 + K\int_{\phi}^{t} u(\tau)\,d\tau$$

$$y(t) = y_0 + \int_{\phi}^{t} v(\sigma)\,d\sigma$$

Take $v_0 = y_0 = \phi$ for simplicity now, then

$$y(t) = K\int_{\phi}^{t}\left[\int_{\phi}^{\sigma} u(\tau)\,d\tau\right]d\sigma$$

So: $$y(t) = K\int_{\phi}^{t}\int_{\phi}^{\sigma} u(\tau)\,d\tau\,d\sigma$$ (Double integral!)

Or: $$y(t) = K\int_{\phi}^{t}(t-\tau)\,u(\tau)\,d\tau$$ (How...?)

Example Control Problem

Find $u(t)$ so that, for a specified t_f, y_f

$$v(t_f) = \emptyset \implies \boxed{\emptyset = K\int_{\emptyset}^{t_f} u(\tau)\,d\tau}$$

$$y(t_f) = y_f \implies \boxed{y_f = K\int_{\emptyset}^{t_f} (t_f - \tau)\,u(\tau)\,d\tau}$$

Solve for $u(\tau)$

Here, we are assuming vehicle starts at rest ($v(\emptyset) = \emptyset$) on the "start line" ($y(\emptyset) = \emptyset$).

Want the vehicle to move to position y_f in t_f seconds and stop there.

Many sol'ns $u(t)$ possible! Typically would also constrain:

1.) $|u(t)| \leq u_{max}$

2.) Behavior of $y(t)$, $t \in [\emptyset, t_f]$

Issues

1.) m, K_f, K_m Not Known precisely:

Hovercraft will not stop exactly where we want.

2.) Requires an accurate clock:

Must use correct $u(t)$ at exactly right times t.

3.) Cannot handle an external ("disturbance") force:

Headwind or cross-breeze will drive hovercraft off the track.

Mathematically sound, but not practical!

Do you drive like that? I hope not!

Mathematically sound, but not practical!

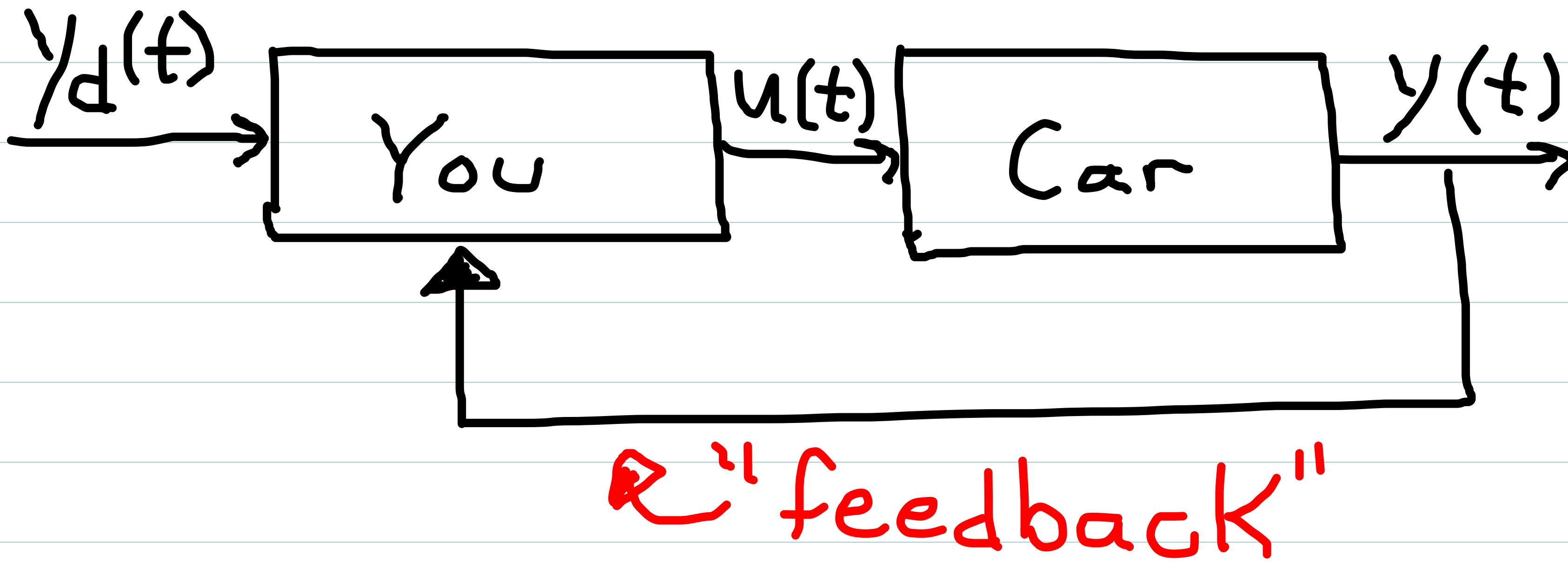
Do you drive like that? I hope not!

Instead you continually compare where you are ($y(t)$) with where you want to be ($y_d(t)$) and continually adjust actions ($u(t)$) based on difference.

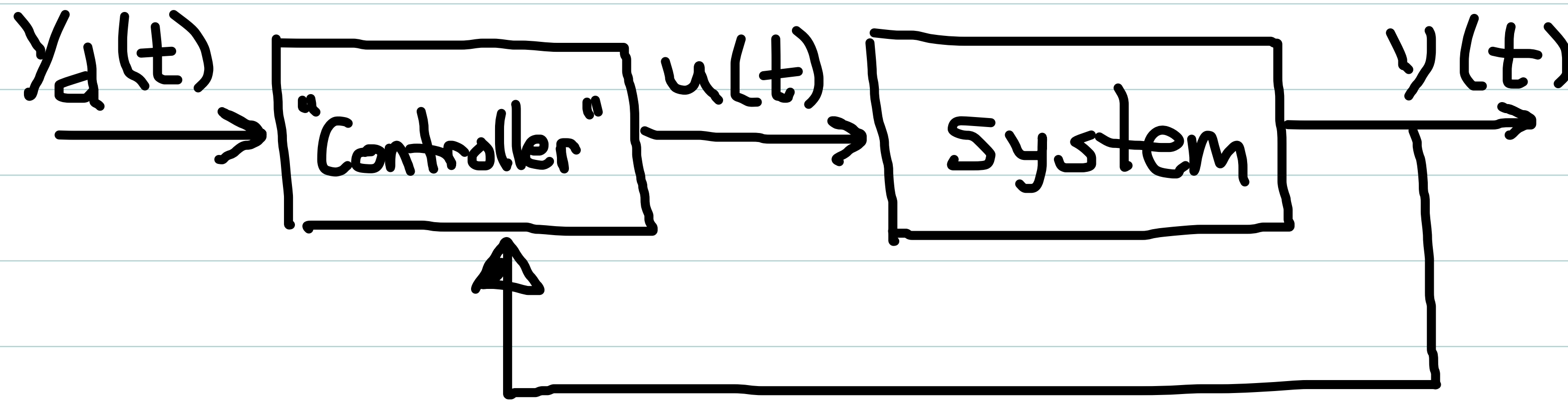
Mathematically sound, but not practical!

Do you drive like that? I hope not!

Instead you continually compare where you are ($y(t)$) with where you want to be ($y_d(t)$) and continually adjust actions ($u(t)$) based on difference.



Feedback Control



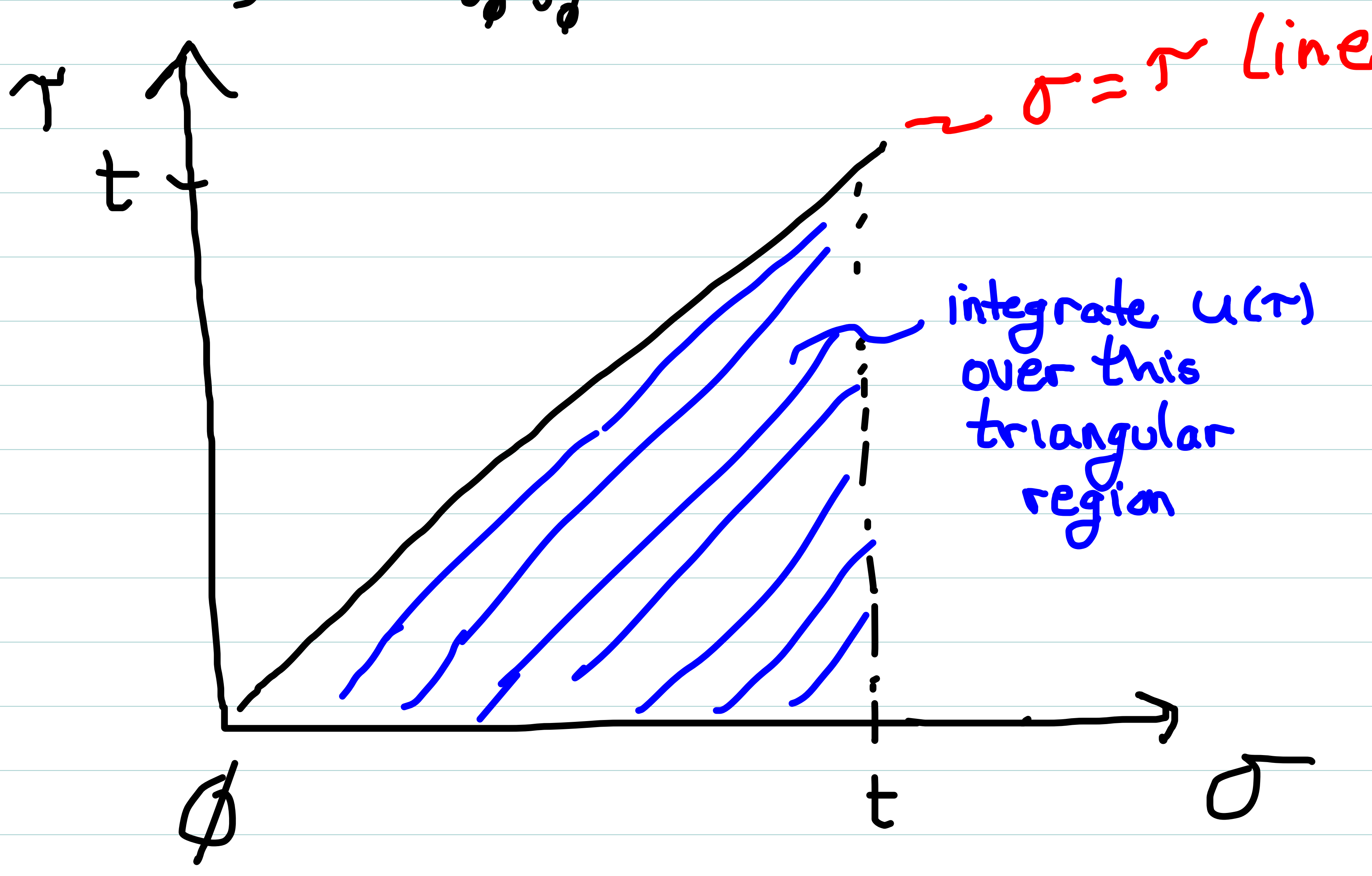
The controller is a mathematical algorithm (implemented as a computer program) which calculates required $u(t)$ from $y(t)$ and $y_d(t)$.

Addresses all 3 issues: uncertainty, disturbance, clocking

This course is about the derivation + implementation of suitable feedback control algorithms based on governing dynamics of system.

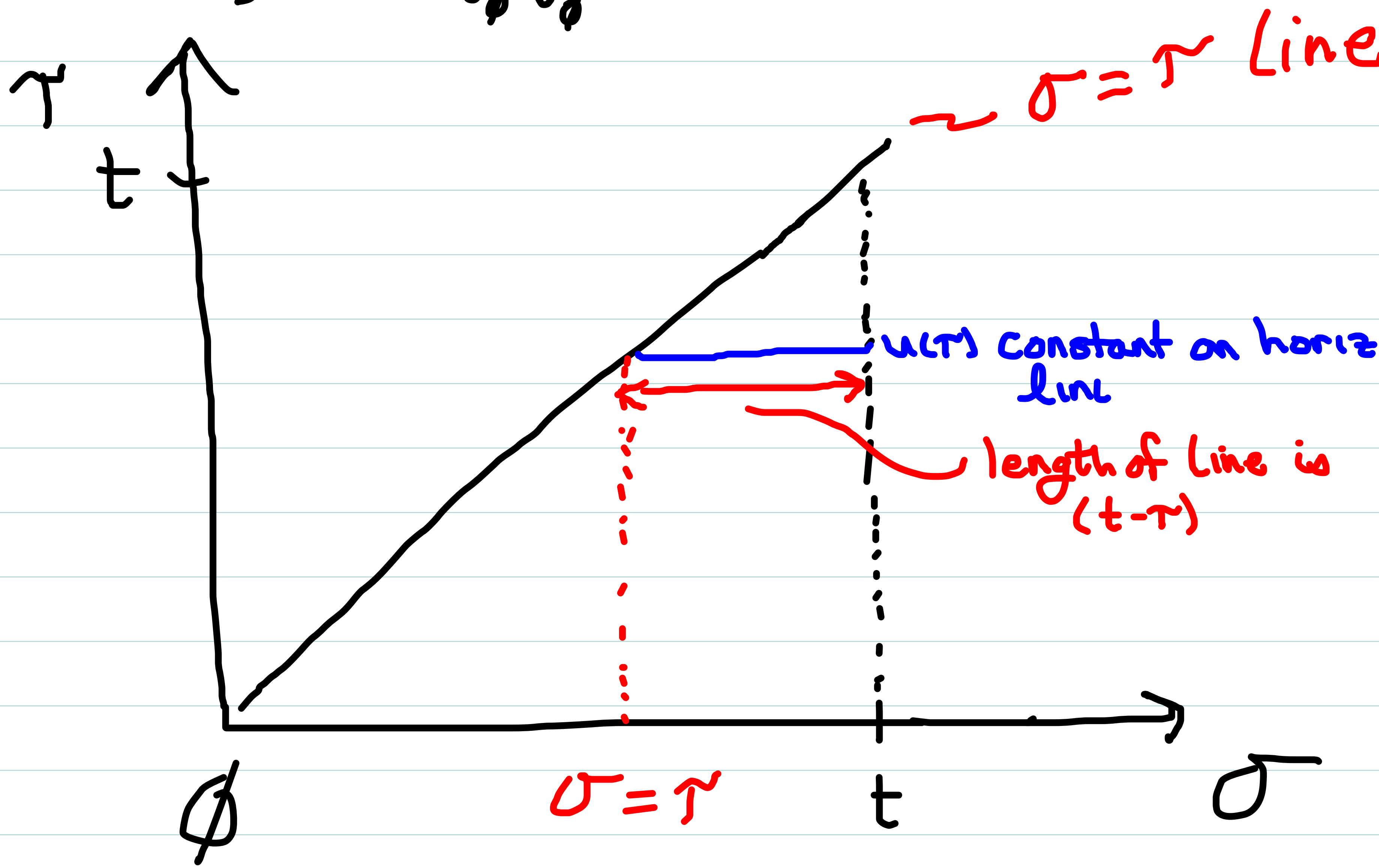
Reduction of the double integral

$$y(t) = K\int_0^t \int_0^\sigma u(\tau)\, d\tau\, d\sigma$$



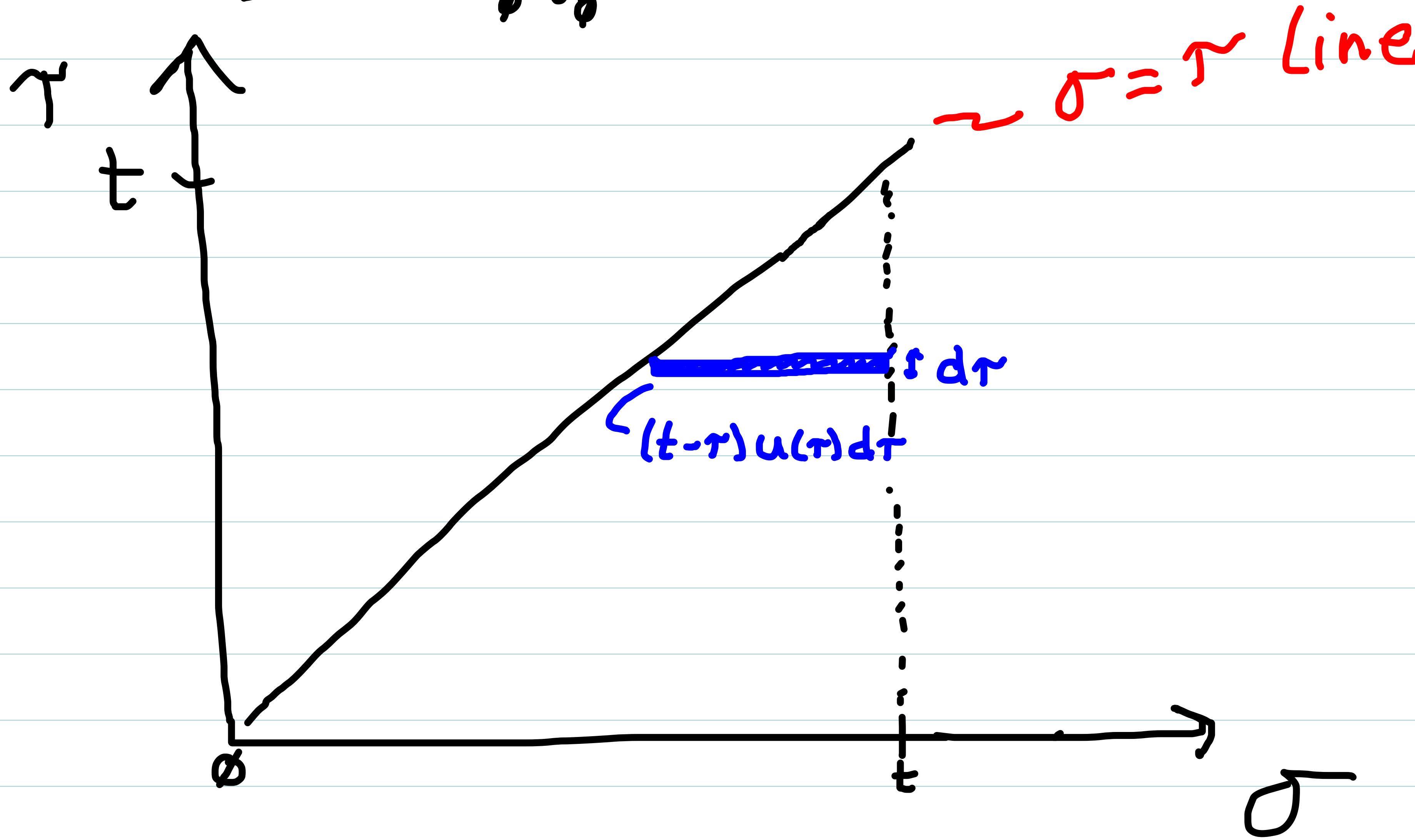
Reduction of this double integral

$$y(t) = K\int_{\phi}^{t}\int_{\phi}^{\sigma} u(\tau)\, d\tau\, d\sigma$$



Reduction of this double integral

$$y(t) = K\int_{\phi}^{t}\int_{\phi}^{\sigma} u(\tau)\, d\tau\, d\sigma$$



Integrate over all strips + multiply by K: $\Rightarrow y(t) = K\int_0^t (t-\tau)u(\tau)\,d\tau$

An alternate form

Our sol'n has the general form:

$$\boxed{y(t) = \int_{\phi}^{t} g(t-\tau)\, u(\tau)\, d\tau}$$

where here $g(t) = Kt$ [so $g(t-\tau) = K(t-\tau)$]

We will (indirectly) show that for <u>any</u> system, no matter how complex the dynamics, this relationship between $u(t)$ and $y(t)$ holds.

Different systems are characterizied by different functions $g(t)$.

The characteristic function $g(t)$ is called the

<u>Impulse response</u>

Implication

Suppose:

$$y_1(t) = \int_0^t g(t-\tau)\, u_1(\tau)\, d\tau$$

$$y_2(t) = \int_0^t g(t-\tau)\, u_2(\tau)\, d\tau$$

are two known input-output pairs.

Suppose that $u(t) = \alpha_1 u_1(t) + \alpha_2 u_2(t)$; α_1, α_2 constant

Then:

$$y(t) = \int_0^t g(t-\tau)\left[\alpha_1 u_1(\tau) + \alpha_2 u_2(\tau)\right] d\tau$$

$$= \alpha_1 \int_0^t g(t-\tau)\, u_1(\tau)\, d\tau + \alpha_2 \int_0^t g(t-\tau)\, u_2(\tau)\, d\tau$$

hence

$y(t) = \alpha_1 y_1(t) + \alpha_2 y_2(t)$ Principle of Linearity.

This suggests an approach:

1. Identify a "family" of functions $u_k(t)$ for which it is easy to calculate response $y_k(t)$:

$$u_k(t) \mapsto y_k(t) \quad \text{(easy)}$$

2. "Break down" an arbitrarily complicated $u(t)$ into a linear combination of the $u_k(t)$:

$$u(t) = \sum \alpha_k u_k(t) \qquad \text{(easy?)}$$

3. Use Linearity:

$$y(t) = \sum \alpha_k y_k(t) \qquad \text{(easy)}$$

Time varying complex numbers

$$z(t) = a(t) + b(t)j$$
$$= r(t)e^{j\theta(t)}$$

Important example:

$$z(t) = e^{st} \text{ with } s \in \mathbb{C}$$

"Complex exponential functions"

Let $s = \sigma + j\omega \quad \sigma, \omega \in \mathbb{R}$

So $\mathrm{Re}\{s\} = \sigma$, $\mathrm{Im}\{s\} = \omega$

① If $\omega = 0$, then

$$e^{st} = e^{\sigma t}$$ (<u>real</u> exponential)

② If $\sigma = 0$ then

$$e^{st} = e^{j\omega t} = \cos\omega t + j\sin\omega t$$

Note: $\mathrm{Im}\{s\}$ gives <u>frequency</u> of the oscillations

③ Most general case

$$e^{st} = e^{(\sigma + j\omega)t} = e^{\sigma t} e^{j\omega t}$$

$$= e^{\sigma t}\left[\cos\omega t + j\sin\omega t\right]$$

$$\left.\begin{aligned} \mathrm{Re}\{e^{st}\} &= e^{\sigma t}\cos(\omega t) \\ \mathrm{Im}\{e^{st}\} &= e^{\sigma t}\sin(\omega t) \end{aligned}\right]$$

$\sigma \to$ amplitude envelope

$\omega \to$ oscillation frequency

$s = \sigma + j\omega$ is the "Complex frequency"

Direct Solution

Our sol'n has the general pattern:

$$\boxed{y(t) = \int_{\phi}^{t} g(t-\tau)\, u(\tau)\, d\tau}$$

where here $g(t) = Kt$ $\left[\text{so } g(t-\tau) = K(t-\tau)\right]$

We will (indirectly) show that for <u>any</u> system, no matter how complex the dynamics, this relationship between $u(t)$ and $y(t)$ holds.

Different systems are characterized by different functions $g(t)$.

The characteristic function $g(t)$ is called the

<u>Impulse response</u>

Implication

Suppose:

$$y_1(t) = \int_{\phi}^{t} g(t-\tau)\, u_1(\tau)\, d\tau$$

$$y_2(t) = \int_{\phi}^{t} g(t-\tau)\, u_2(\tau)\, d\tau$$

are two Known input-output pairs.

Suppose that $u(t) = \alpha_1 u_1(t) + \alpha_2 u_2(t)$; α_1, α_2 constant

Then:

$$y(t) = \int_{\phi}^{t} g(t-\tau) \left[\alpha_1 u_1(\tau) + \alpha_2 u_2(\tau)\right] d\tau$$

$$= \alpha_1 \int_{\phi}^{t} g(t-\tau)\, u_1(\tau)\, d\tau + \alpha_2 \int_{\phi}^{t} g(t-\tau)\, u_2(\tau)\, d\tau$$

hence

$$y(t) = \alpha_1 y_1(t) + \alpha_2 y_2(t)$$ Principle of Linearity.

This suggests an approach:

1. Identify a "family" of functions $u_k(t)$ for which it is EASY to calculate response $y_k(t)$:

$$u_k(t) \mapsto y_k(t) \quad \text{(easy)}$$

2. "Break down" an arbitrarily complicated $u(t)$ into a linear combination of the $u_k(t)$:

$$u(t) = \sum \alpha_k u_k(t) \quad \text{(easy?)}$$

3. Use Linearity:

$$y(t) = \sum \alpha_k y_k(t) \quad \text{(easy)}$$

Time varying complex numbers

$$z(t) = a(t) + b(t)\,j$$
$$= r(t)\,e^{j\theta(t)}$$

Important example:

$$z(t) = e^{st} \text{ with } s \in \mathbb{C}$$

"complex exponential functions"

Let $s = \sigma + j\omega \quad \sigma, \omega \in \mathbb{R}$

So $\mathrm{Re}\{s\} = \sigma$, $\mathrm{Im}\{s\} = \omega$

① If $\omega = \emptyset$, then

$$e^{st} = e^{\sigma t} \text{ (\underline{real} exponential)}$$

② If $\sigma = \emptyset$ then

$$e^{st} = e^{j\omega t} = \cos\omega t + j\sin\omega t$$

Note: $\mathrm{Im}\{s\}$ gives *frequency* of the oscillations

③ Most general case

$$e^{st} = e^{(\sigma + j\omega)t} = e^{\sigma t} e^{j\omega t}$$

$$= e^{\sigma t}\left[\cos\omega t + j\sin\omega t\right]$$

$$\mathrm{Re}\{e^{st}\} = e^{\sigma t}\cos(\omega t)$$

$$\mathrm{Im}\{e^{st}\} = e^{\sigma t}\sin(\omega t)$$

$\sigma \rightarrow$ amplitude envelope

$\omega \rightarrow$ oscillation frequency

$s = \sigma + j\omega$ is the "Complex frequency"

Utility of e^{st}

For different values of s, e^{st} is:

- a constant
- a real exponential
- a pure sine/cosine wave
- an exponentially decaying or increasing sine/cosine

Covers 90% of cases needed to solve linear diff'l eq'ns

Complex Amplitudes

Now consider $z(t) = Ae^{st}$ with both $A, s \in \mathbb{C}$.

$s = \sigma + j\omega$, $A = re^{j\varphi}$ (polar)

$$Ae^{st} = (re^{j\varphi})(e^{(\sigma+j\omega)t})$$
$$= (re^{\sigma t})(e^{j(\omega t+\varphi)})$$
$$= re^{\sigma t}[\cos(\omega t+\varphi) + j\sin(\omega t+\varphi)]$$

So $$\boxed{\begin{aligned} \mathrm{Re}\{Ae^{st}\} &= re^{\sigma t}\cos(\omega t+\varphi) \\ \mathrm{Im}\{Ae^{st}\} &= re^{\sigma t}\sin(\omega t+\varphi) \end{aligned}}$$

$r = |A|$ is initial amplitude of oscillations

$\varphi = \measuredangle A$ is phase shift of oscillations

$\varphi > 0$ called "phase lead"

$\varphi < 0$ called "phase lag"

Property of e^{st}:

Let $f(t) = e^{st}$ for any $s \in \mathbb{C}$

Then $\dot{f}(t) = \frac{d}{dt} f(t) = \frac{d}{dt}(e^{st})$

$= s e^{st}$

or $\dot{f}(t) = s f(t)$

Similarly: $\ddot{f}(t) = s^2 f(t)$

$f^{(3)}(t) = s^3 f(t)$, etc

Linear constant coefficient (time invariant) Diff Eq'n

$$\alpha_n y^{(n)} + \alpha_{n-1} y^{(n-1)} + \cdots + \alpha_1 \dot{y} + \alpha_0 y$$
$$= \beta_m u^{(m)} + \cdots + \beta_1 \dot{u} + \beta_0 u$$

where $\alpha_n, \ldots \alpha_0$ and $\beta_m \ldots \beta_0$ are <u>real</u> and <u>constant</u>

Suppose $u(t) = U e^{st}$ with $s, U \in \mathbb{C}$

Is $y(t) = Y e^{st}$ a sol'n for some $Y \in \mathbb{C}$?

Substitute into DE

GIVES

$$r(s)Ye^{st} = q(s)Ue^{st}$$

With:

$$r(s) = \alpha_n s^n + \alpha_{n-1} s^{n-1} + \cdots + \alpha_1 s + \alpha_0$$

$$q(s) = \beta_m s^m + \cdots + \beta_1 s + \beta_0$$

So Assumption is consistent with

$$Y = \left[\frac{q(s)}{r(s)}\right] U = G(s) U$$

$$Ue^{st} = u(t) \longrightarrow \underset{\text{system}}{\boxed{G(s)}} \longrightarrow y(t) = Ye^{st}$$

$$Y = G(s)U$$

If $u(t) = Ue^{st}$ for some $U, s \in \mathbb{C}$ then $y(t) = Ye^{st}$, with $Y = G(s)U$

This is one possible sol'n of the DE, the forced sol'n, $\boxed{y_f(t)}$.

Other sol'ns are possible.

Other possible sol'ns

Now, suppose $u(t) = \emptyset$. Clearly here $y_f(t) = \emptyset$. But is $y(t) = \emptyset$ necessarily? Or can we still have sol'ns of the form $y(t) = Ce^{st}$? Substitute into DE:

$$r(s)Ce^{st} = \emptyset$$

which can be true for any s where

$$\boxed{r(s) = \emptyset}$$

$$r(s) = \alpha_n s^n + \cdots + \alpha_1 s + \alpha_0$$

There are n values of s for which $r(s) = 0$. We denote these

$$p_1, p_2, \ldots p_n$$

So $r(s)$ can be factored as

$$r(s) = \alpha_n (s - p_1)(s - p_2) \cdots (s - p_n)$$

$$= \alpha_n \prod_{k=1}^{n} (s - p_k)$$

For any p_K with $r(p_K) = \phi$,
$y(t) = e^{p_K t}$ is a sol'n of the DE
when $u(t) = \phi$. So is $y(t) = C_K e^{p_K t}$
for any constant C_K. So is any
<u>sum</u> of these terms:

$$y(t) = \sum_{K=1}^{n} C_K e^{p_K t} = y_h(t)$$

The "homogeneous" sol'n.

Proof!

Substitute $y(t) = \sum_{k=1}^{n} c_k e^{p_k t}$ into diff eq'n:

GIVES:

$$r(p_1)c_1 e^{p_1 t} + r(p_2)c_2 e^{p_2 t} + \cdots + r(p_n)c_n e^{p_n t} = \emptyset$$

Which is true if $r(p_1) = r(p_2) = \cdots = r(p_n) = \emptyset$

i.e. the p_k are zeros of polynomial $r(s)$

Since any $y_h(t)$ yields $\emptyset$ exactly when substituted into DE, we can add it to any other sol'n and still have a valid sol'n. Generally:

$$y(t) = y_h(t) + y_f(t)$$

where $y_h(t) = \sum_{k=1}^{n} C_k e^{p_k t}$

Both complex! (Generally)

and if $u(t) = U e^{st}$, then

$$y_f(t) = G(s) U e^{st}$$

But $y_f(t)$ is complex generally...?

$\Longrightarrow$ because $u(t)$ is complex here

Suppose $u(t) = B\sin(\omega t + \varphi)$ (real)

$= \mathrm{Im}\{U e^{st}\}$ Take matching Im part

with $U = B e^{j\varphi}$

and $s = j\omega$

Then $y_f(t) = \mathrm{Im}\{G(s) U e^{st}\}$

and similarly for cosine inputs, taking real part

What about $y_h(t)$?

Contains terms e^{pt}, where $r(p) = \emptyset$.

If p is complex, $p = \sigma + j\omega$, $\omega \neq \emptyset$

then e^{pt} is complex

<u>However</u>: in this case $r(p) = \emptyset \Longrightarrow r(\bar{p}) = \emptyset$

i.e. $\bar{p}$ is also a zero of $r(s)$.

$\Longrightarrow$ Complex roots of polynomials occur in "conjugate pairs".

Hence, with complex roots, $y_h(t)$ will contain

$$C_1 e^{pt} + C_2 e^{\bar{p}t}$$

<u>Fact:</u>

$$C_2 = \overline{C_1}$$

i.e. coef of $e^{\bar{p}t}$ will <u>always</u> be the conjugate of the coef of e^{pt}.

Thus, if $r(s)$ has a complex root p, $y_h(t)$ will contain

$$c e^{pt} + \bar{c} e^{\bar{p}t} = c e^{pt} + \overline{c e^{pt}}$$

Recap (DE Review)

Any constant coef linear diff'l eq'n has sol'n:

$$y(t) = y_h(t) + y_f(t)$$

where

$$y_h(t) = \sum_{k=1}^{n} c_k e^{p_k t}$$] homogeneous response

$$r(p_k) = 0, \quad k = 1, \ldots, n$$] p_k roots of char. poly $r(s)$

and $y_f(t)$ depends on specific forcing function (input)
("forced response")

For the specific case that $u(t) = Ue^{st}$, $U, s \in \mathbb{C}$

then

$$y_f(t) = G(s)u(t), \quad G(s) = \frac{q(s)}{r(s)}$$

"transfer function"

$$= Ye^{st}$$

with $Y = G(s)U$ (ordinary complex number multiplication!)

Complex math yields real sol'ns

Note that both $y_h(t)$ and $y_f(t)$ are complex-valued functions as we have written them

But physical systems will have only real-valued inputs and outputs.

For $y_f(t)$, note that we can express a real input as the real or imag part of a complex input:

$$u(t) = Ae^{\sigma t}\sin(\omega t+\varphi) = \mathrm{Im}\{\bar{U}e^{st}\}$$

with $\bar{U} = Ae^{j\varphi}$ and $s = \sigma + j\omega$

The corresponding real $y_f(t) = \mathrm{Im}\{G(s)\bar{U}e^{st}\}$

and similarly if input is cosinoidal we use the real part of the complex number calculation

Complex => real, cont

for $y_h(t)$: Sol'n contains terms e^{pt}, $r(p)=0$

This will be complex if root p is complex, i.e.

$p = \sigma + j\omega$, $\omega \neq 0$.

However, if this is true then $\bar{p} = \sigma - j\omega$ will

<u>also</u> be a root of $r(s)$, i.e. $r(\bar{p}) = r(p) = 0$

Complex roots of polynomials always occur in "conjugate pairs"

So sol'n for $y_h(t)$ will really look like

$$y_h(t) = C_1 e^{pt} + C_2 e^{\bar{p}t} + (\cdots \text{other terms})$$

Real-valued homogeneous response

So

$$y_h(t) = \underbrace{c_1 e^{pt} + c_2 e^{\bar{p}t}} + (\cdots \text{ other terms})$$

Fact: It is always the case that $c_1 = \bar{c_2}$,

i.e. the coefs. of conjugate terms are themselves conjugates

(This is b/c boundary cond'ns on DE are also real)

$$\Rightarrow \text{so, } y_h(t) = ce^{pt} + \bar{c}e^{\bar{p}t} + (\cdots)$$

$$= \underbrace{ce^{pt} + \overline{ce^{pt}}}_{\text{sum of conjugate terms}} + (\cdots)$$

$$= \underbrace{2\,\mathrm{Re}\{ce^{pt}\}} + (\cdots)$$

$\Rightarrow$ The two complex terms from conjugate roots $p, \bar{p}$ combine to form a real function of time!

Conclusion

$\Rightarrow$ When $p = \sigma + j\omega$, $\omega \neq 0$ is a root of char poly $r(s)$, the homog. sol'n will contain the real-valued function

$$2\,\mathrm{Re}\{c e^{pt}\} = A e^{\sigma t} \cos(\omega t + \varphi)$$

where $A = 2|c|$, $\varphi = \measuredangle c$ are determined by ICs.

$\Rightarrow$ {Complex roots of $r(s)$ correspond to real oscillations in homog. response.}

(Note, there may be several pairs of conjugate roots in $r(s)$, resulting in multiple oscillations w/ different frequencies + damping)

Recap (DE Review)

Any constant coef diff'l eq'n has sol'n

$$y(t) = y_h(t) + y_f(t)$$

where

$$y_h(t) = \sum_{k=1}^{n} c_k e^{p_k t}$$ homogeneous response

$$r(p_k) = 0, \quad k = 1, \ldots, n$$ p_k roots of char. poly $r(s)$

and $y_f(t)$ depends on specific forcing function (input)

("forced response")

For the specific case that $u(t) = U e^{st}$, $U, s \in \mathbb{C}$

then $y_f(t) = G(s) u(t)$, $G(s) = \frac{q(s)}{r(s)}$ "transfer function"

$$= Y e^{st}$$

with $Y = G(s) U$ (ordinary complex number multiplication!)

Boundary/initial conditions

Undetermined coefs c_K in $y_h(t)$ determined by boundary conditions $y(0), \dot{y}(0), \ldots, y^{(n-1)}(0)$

For simple problems, can often solve for c_K by substituting general form $y(t) = y_h(t) + y_f(t)$, differentiating, and matching stated B/Cs.

$\Rightarrow$ Results in a system of n equations in the n coefs c_K which can be solved (lin. algebra)

Warning: There are situations where this approach to compute c_K will not work.

Will cover this situation shortly.

Particularly if $u(t)$ is discontinuous at $t=0$ and one or more derivs of $u(t)$ appear in DE

Example

$$\ddot{y}(t) + 5\dot{y}(t) + 4y(t) = 2\dot{u}(t) + u(t)$$

where $y(0) = \dot{y}(0) = 0, \quad u(t) = 3\cos(2t - \pi/2)$

By inspection:

$$y(t) = \underbrace{c_1 e^{-t} + c_2 e^{-4t}}_{y_h(t)} + \underbrace{A\cos(2t + \varphi)}_{y_f(t)}$$

Only remaining problem is to calculate c_1, c_2, A, φ

Note: A, φ in $y_f(t)$ determined by $u(t)$, and are independent of c_1, c_2

With a little more calculation:

$$y(t) = \frac{1}{10}\left[7e^{-4t} - 4e^{-t}\right] + \underbrace{\left(\frac{3\sqrt{17}}{10}\right)}_{A} \cos\Big(2t \underbrace{- \pi/2 - \tan^{-1} 1/4}_{\varphi}\Big)$$

Forced response

(zero)

Here $u(t) = 3\cos(2t - \pi/2) = 3e^{\emptyset t}\cos(2t - \pi/2)$

$= \mathrm{Re}\{U e^{st}\}$ with $U = 3e^{-\pi/2\, j}$

$s = \emptyset + 2j$

and $y_f(t) = \mathrm{Re}\{G(s) U e^{st}\}$

$= \mathrm{Re}\{Y e^{st}\}$, $Y = G(s)U$

$= |Y| e^{\emptyset t} \cos(2t + \measuredangle Y)$ ($s = \emptyset + 2j$ from input)

All we need to do is the complex number multiplication $Y = G(s)U$ and convert to polar form $|Y|, \measuredangle Y$

$\Rightarrow$ We have U, still need transfer fn $G(s)$

Diff'l eq'n is

$$\dddot{y} + 5\dot{y} + 4y = 2\dot{u} + u$$

$$G(s) = \frac{q(s)}{r(s)}$$

$q(s) =$

$r(s) =$

so finally $G(s) =$

Note that $r(s) = (s+1)(s+4)$

which also gives us the general form for $y_h(t)$

$$c_1 e^{-t} + c_2 e^{-4t}$$

Evaluate $G(s)$ at complex freq of input, $s=2j$ here

$$G(2j) = \left.\frac{2s+1}{s^2+5s+4}\right|_{s=2j}$$

$$= \frac{1+4j}{10j}$$

$$= \frac{1}{10}(1-4j)$$

How...?

Here $u(t) = 3\cos(2t - \pi/2) = \mathrm{Re}\{U e^{st}\}$

with $s = 2j$ and $U = 3e^{-\pi/2 j}$

So $y_f(t) = \mathrm{Re}\{\underbrace{G(2j)(3e^{-\pi/2 j})}_{Y = G(2j)U}(e^{2jt})\}$

with here: $G(s) = \dfrac{2s+1}{s^2+5s+4}$

$$\Rightarrow G(2j) = \frac{1+4j}{(2j)^2+10j+4} = \frac{1}{10}(4-j) = \left(\frac{\sqrt{17}}{10} \measuredangle -\tan^{-1}\left(\tfrac{1}{4}\right)\right)$$

Hence:

$$y_f(t) = \overbrace{\frac{3\sqrt{17}}{10}}^{|Y|} \cos(2t \overbrace{- \pi/2 - \tan^{-1}(1/4)}^{\measuredangle Y})$$

So we know $y_f(t)$ exactly at this point

Homogeneous Sol'n

We have $r(s) = s^2 + 5s + 4$ (denom poly of $G(s)$)

Or: $r(s) = (s+1)(s+4)$

So $p_1 = -1$, $p_2 = -4$ and $y_h(t) = c_1 e^{-t} + c_2 e^{-4t}$

Then $y(t) = y_f(t) + y_h(t)$

$$= \frac{3\sqrt{17}}{10} \cos\left(2t - \pi/2 - \tan^{-1}(1/4)\right) + c_1 e^{-t} + c_2 e^{-4t}$$

So $y(\phi) = c_1 + c_2 - \frac{3}{10} = \phi$ (specified)

and $\dot{y}(\phi) = -c_1 - 4c_2 + 12/5 = \phi$ (specified)

Impose Boundary Cond'ns

Equivalently

$$\begin{bmatrix} 1 & 1 \\ -1 & -4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3/10 \\ -12/5 \end{bmatrix}$$

$\Rightarrow$ (linear algebra):

$$c_1 = -4/10, \quad c_2 = 7/10$$

So that:

$$y(t) = \frac{1}{10}\left[7e^{-4t} - 4e^{-t}\right] + \frac{3\sqrt{17}}{10}\cos\left(2t - \frac{\pi}{2} - \tan^{-1}\frac{1}{4}\right)$$

as claimed

final result

Recap

General sol'n of LTI DE is:

$$y(t) = y_h(t) + y_f(t)$$

Forced response $y_f(t)$ depends on $u(t)$

Homogeneous response is <u>independent</u> of $u(t)$:

$$\left. y_h(t) = \sum_{k=1}^{n} c_k e^{p_k t} \quad \text{where } r(p_k) = 0 \right\} \text{ for } \underline{\text{any}} \; u(t)$$

Specific coefs c_k depend on initial conditions and $u(t)$.

Repeated roots of $r(s)$

Above formula for $r(s)$ assumes the roots p_k are <u>non-repeated</u>

Suppose instead that there are repeated roots, for example:

$$r(s) = (s-p_1)^{\ell}(s-p_{\ell+1})\cdots(s-p_n)$$

i.e. p_1 is repeated ℓ times. Then:

$$y_h(t) = (c_1 + c_2 t + c_3 t^2 + \cdots + c_\ell t^{\ell-1})e^{p_1 t} + \sum_{k=\ell+1}^{n} c_k e^{p_k t}$$

(will prove later)

(Natural) Modes

$r(p_k) = 0$

$y_h(t)$ is a linear combination of $e^{p_k t}$ (or $t^i e^{p_k t}$). These describe solutions which are possible without any input

They are "natural" motions which are intrinsic to the dynamics of the system.

We call them the "MODES".

$\boxed{\text{MODES:}}$ Terms in sol'n for $y(t)$ of form e^{pt}, where $r(p) = \emptyset$

Two cases (non-repeated, to start)

① p real : e^{pt} is a real exponential function

"1st order mode"

(2) p complex: e^{pt} and $e^{\bar{p}t}$ both present in solution, and will combine to form the "2nd order mode"

$$Ae^{\sigma t}\cos(\omega t+\varphi)$$

where $\sigma = \mathrm{Re}\{p\}$, $\omega = \mathrm{Im}\{p\}$

and A, φ depend on the initial conditions

Stability

A mode e^{pt} is stable if

$$|e^{pt}| \to 0 \text{ as } t \to \infty$$

A system is stable if

$$|e^{p_k t}| \to 0 \text{ for all } k = 1, \ldots, n$$

i.e. if every mode is stable

Note: if true then $y_h(t) \to 0$ for any set of initial conditions.

Stability Condition

As usual, let $p = \sigma + j\omega$. Then:

$$|e^{pt}| = |e^{(\sigma + j\omega)t}|$$

$$= |e^{\sigma t} e^{j\omega t}| = |e^{\sigma t}| \, |e^{j\omega t}|$$

$$= |e^{\sigma t}|$$

So $|e^{pt}| \to 0$ only if $\sigma < 0$. Hence:

$$\boxed{\text{A mode is stable if } \sigma = \mathrm{Re}\{p\} < 0}$$

System Stability

The system is stable if:

$$\mathrm{Re}\{p_k\} < 0 \text{ for all } k = 1, \ldots, n$$

$\Rightarrow$ all roots of $r(s)$ have negative real parts

$\Rightarrow$ all roots of $r(s)$ lie to the left of imaginary axis in the complex plane.

$\Rightarrow$ all roots of $r(s)$ lie in left half of complex plane (LHP)

Instability

A mode e^{pt} is unstable if $Re\{p\} > 0$

$\Rightarrow$ root p lies to right of imag axis

$\Rightarrow$ p is in "right half plane" (RHP)

A system is unstable if:

$$Re\{p_k\} > 0 \text{ for any } k = 1, \ldots, n$$

i.e. if any roots of $r(s)$ are in RHP.

What about repeated modes?

Repeated real modes have terms like:

$t^i e^{pt}$ (powers of t multiplying e^{pt})

Fact: for any $i > 0$, if $Re\{p\} < \emptyset$ then

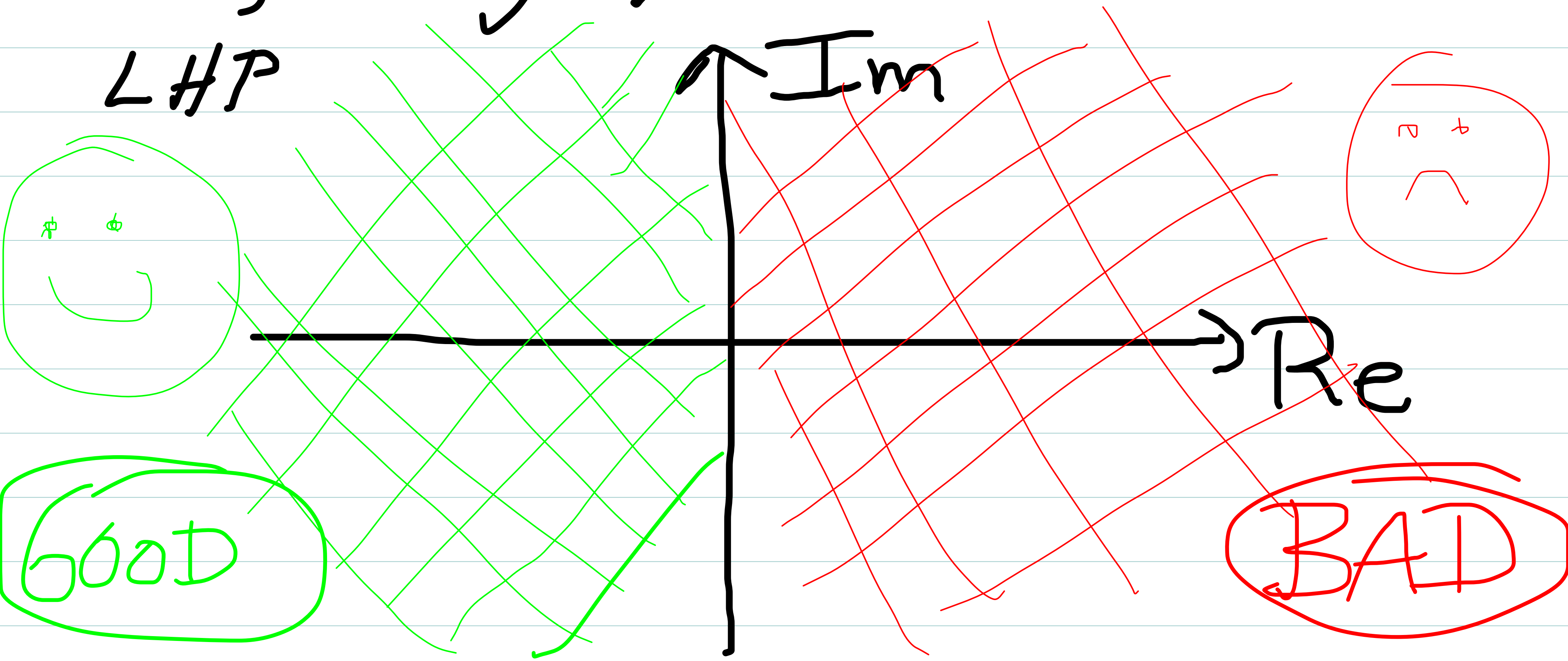
$$\lim_{t \to \infty} |t^i e^{pt}| \to \emptyset$$

Thus a repeated mode is stable as long as the repeated roots are in LHP.

Conversely, a repeated mode is unstable if repeated roots in RHP.

Hence:

A system is stable if all roots of $r(s)$, including repeated roots, lie in LHP



Recap

Stable mode: $|e^{pt}| \to 0$ as $t \to \infty$

$$\Longleftrightarrow \text{Re}\{p\} < 0$$

Unstable mode: $|e^{pt}| \to \infty$ as $t \to \infty$

$$\Longleftrightarrow \text{Re}\{p\} > 0$$

Stable **system**: $\text{Re}\{p_k\} < 0$ for **all** $k = 1, \ldots, n$

Unstable **system**: $\text{Re}\{p_k\} > 0$ for **any** $k = 1, \ldots, n$

What happens if $\text{Re}\{p\} = 0$?

Marginally stable MODES

$\mathrm{Re}\{p\} = \emptyset \Rightarrow |e^{pt}| = |e^{j\omega t}| = 1 \quad \forall t \geq 0$

i.e. the magnitude is constant

$\Rightarrow$ neither increasing nor decreasing with time

$\Rightarrow$ neither stable nor unstable

"Marginally stable"

Repeated MODES with $\mathrm{Re}\{p\} = \emptyset$ will increase in magnitude polynomially in t

$\Rightarrow$ Not as "bad" as exponential growth

An alternate decomposition of $y(t)$

$$y(t) = y_h(t) + y_f(t)$$

$$= y_{tr}(t) + y_{ss}(t) \quad \} \text{ (regroup terms)}$$

$y_{tr}(t)$ is the "transient response", which satisfies:

$$\lim_{t \to \infty} |y_{tr}(t)| \to 0$$

$y_{ss}(t)$ is the "steady-state" response, which is all remaining terms in $y(t)$.

Notes:

1. If system is stable, $y_{tr}(t)$ contains $y_h(t)$ but $y_{tr}(t)$ would also contain decaying terms in $y_f(t)$ (if any).

2. Conversely, marginally stable terms in $y_h(t)$ (if any) would be part of $y_{ss}(t)$.

3. "steady-state" is not a useful concept if system is unstable.

Example: Stable system with constant input

$$u(t) = U_0 \longrightarrow \boxed{G(s)} \longrightarrow y(t) = ?$$

(U_0 real)

constant!

$$y(t) = y_h(t) + y_f(t) = y_h(t) + \boxed{G(0)U_0}$$

Since system is stable, $y_h(t) \to 0$ as $t \to \infty$

So here:

$$y_{tr}(t) = y_h(t)$$

$$y_{ss}(t) = G(0)\,U_0 \quad \text{(constant)}$$

Very common and important case!

A Different Example

$$G(s) = \frac{s+2}{s(s+1)}, \quad u(t) = e^{-3t}$$

$$y_h(t) = c_1 + c_2 e^{-t}$$

$$y_f(t) = G(-3)e^{-3t} = -\tfrac{1}{6}e^{-3t}$$

$$y_{tr}(t) = c_2 e^{-t} - \tfrac{1}{6}e^{-3t}$$

$$y_{ss}(t) = c_1$$

Note: system is <u>not</u> stable here.

Convergence metrics

Useful to quantify how quickly stable modes decay to $\emptyset$.

"<u>2% criterion</u>": Defines the <u>settling time</u> t_s for a mode to be such that

$$|e^{pt}| \leq .02 \quad \forall t \geq t_s$$

For a 1^{st} order mode ($p = \sigma$, real)

$$t_s = \frac{\ln(.02)}{\sigma} \approx \frac{4}{|\sigma|} = \frac{4}{|\mathrm{Re}\{p\}|}$$

2^{nd} order settling time

For a 2^{nd} order mode $e^{pt}, e^{\bar{p}t}$ with $p = \sigma + j\omega$, $\omega \neq 0$, the calculation is more complicated due to the oscillations.

However:

$$t_s \approx \frac{4}{|\sigma|} = \frac{4}{|Re\{p\}|}$$

is still a useful approximation in these cases also.

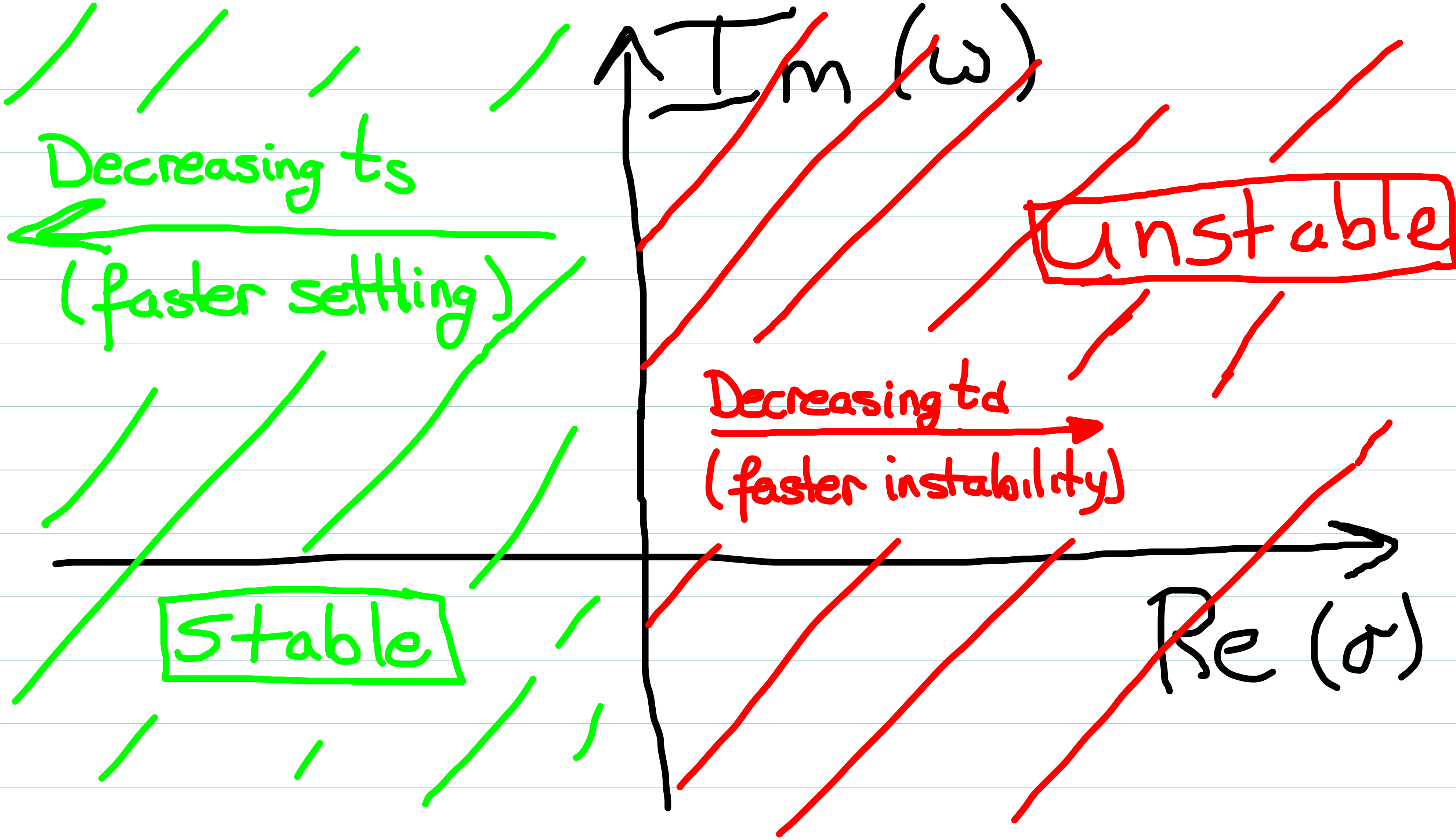
"Doubling time" of unstable modes

When $\sigma > 0$, the doubling time t_d is such that

$$|e^{\sigma t_d}| = 2 \quad \Rightarrow \quad \boxed{t_d \approx \frac{0.7}{\sigma}}$$

Smaller $t_d \iff$ "more unstable" system

$\Rightarrow$ Faster rate of increase for amplitude



Settling times decrease the further left of the imag axis the root p is.

To a first approximation, the settling time of a system is the settling time of its slowest mode

$\Rightarrow$ Mode closest to imag Axis determines settling time

$\Rightarrow$ Called the "dominant mode"

$\Rightarrow$ Only a "first cut." Will refine later

2^{nd} Order "Damping ratio"

For 2^{nd} order modes we also define the

damping ratio

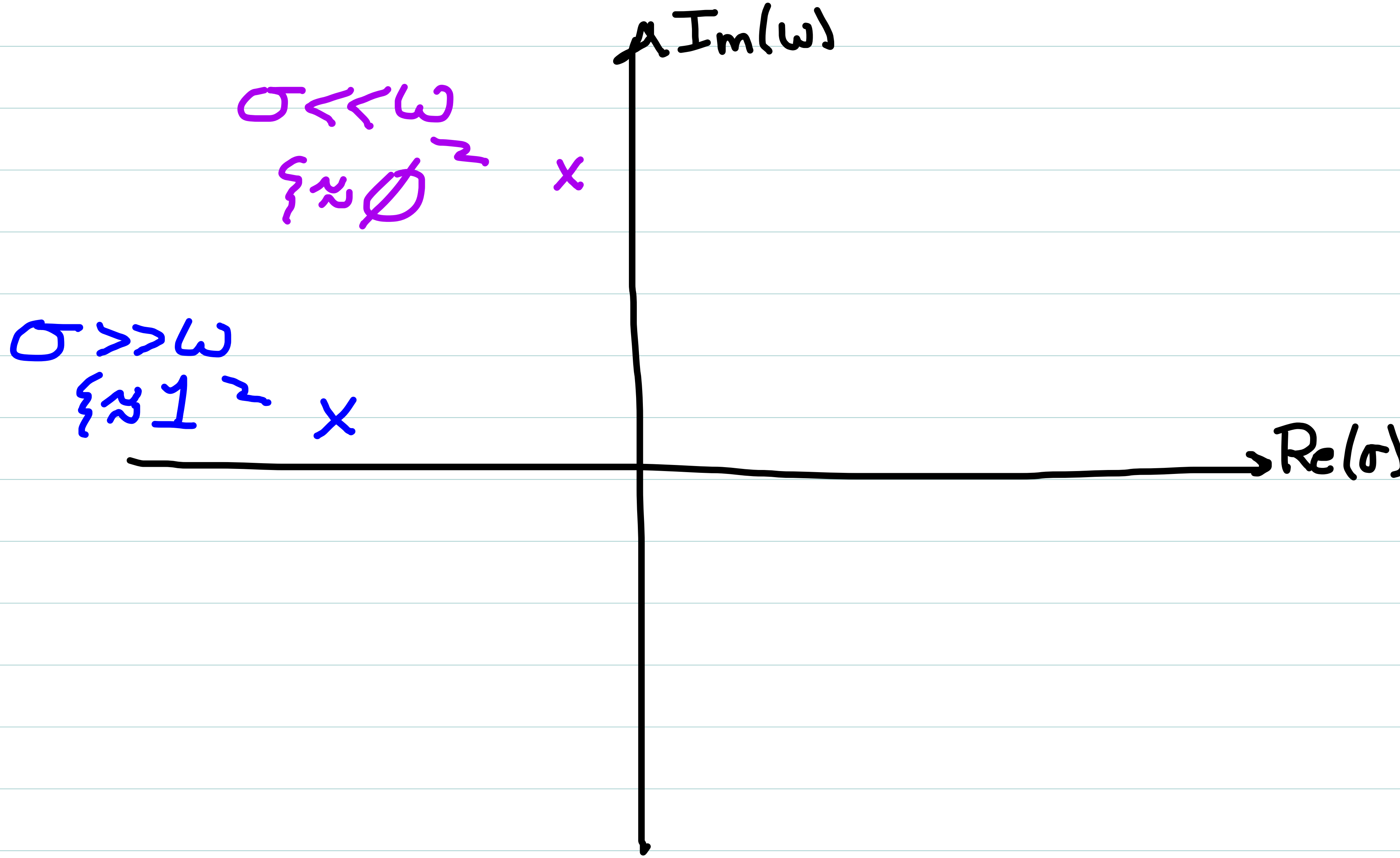
$$\boxed{\zeta = \left|\frac{\sigma}{p}\right| = \frac{|\sigma|}{\sqrt{\sigma^2+\omega^2}}}$$

A non-dimensional comparison of convergence rate to oscillation frequency

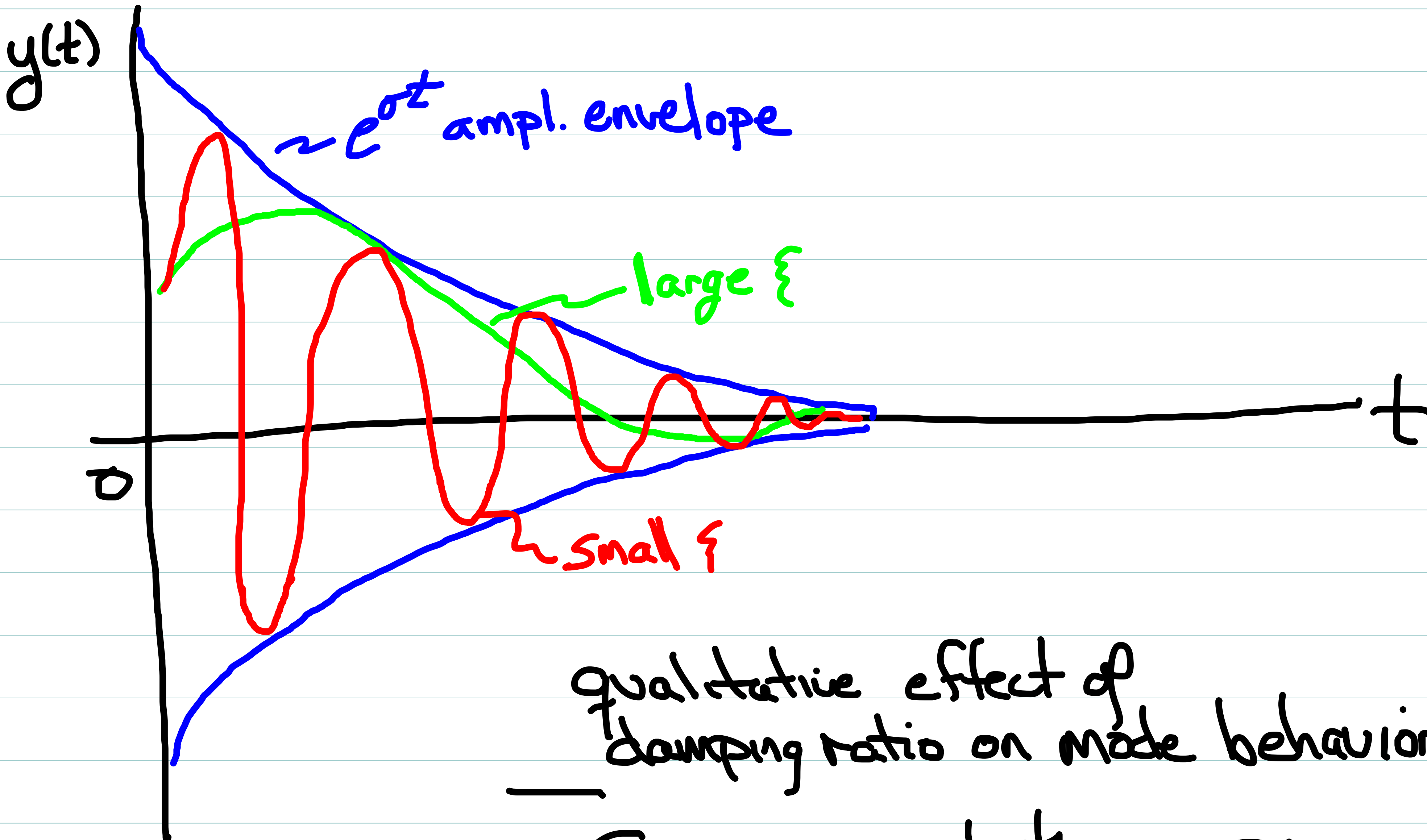
$0 \leq \zeta \leq 1$ for a stable mode

$\zeta \approx 0 \Longleftrightarrow$ many oscillations before 2% criterion reached

$\zeta \approx 1 \Longleftrightarrow$ less than one complete oscillation before 2% criterion reached.



Will explore in greater detail later



qualitative effect of damping ratio on mode behavior

Same σ in both CASES, different ω

Transfer functions

$$G(s) = \frac{q(s)}{r(s)}$$

Compactly gives us all information we need to predict major features of system response

- $y_h(t)$, modes, stability : all from $r(s)$ the denominator polynomial of $G(s)$

$$r(s) = \alpha_n \prod_{k=1}^{n} (s - p_k)$$

- forced response : Evaluate $G(s)$ at specific complex values of s.

Numerator Terms

Can also factor $q(s)$:

$$q(s) = \beta_m (s - z_1)(s - z_2) \cdots (s - z_m)$$

where $q(z_i) = \emptyset$ for $i = 1, \ldots, m$

The values z_i are called the <u>zeros</u> of $G(s)$

$$\text{since } G(z_i) = \frac{\cancel{q(z_i)} \xrightarrow{} \emptyset}{r(z_i)} = \emptyset$$

The values p_k are called the <u>poles</u> of $G(s)$

$$\text{since } G(p_k) = \frac{q(p_k)}{\emptyset \xleftarrow{} \cancel{r(p_k)}} = \infty$$

Zero/Pole/Gain (ZPK) form

$$G(s) = K\left[\frac{\prod_{i=1}^{m}(s-z_i)}{\prod_{k=1}^{n}(s-p_k)}\right]$$

Poles p_k satisfy $r(p_k) = \emptyset$

Zeros z_i satisfy $q(z_i) = \emptyset$

Gain: $K = \frac{\beta_m}{\alpha_n}$ (always real)

Alternate ZPK form:

When $G(s)$ has complex poles and/or zeros, we commonly combine the conjugate roots of $r(s)$ or $q(s)$ into 2^{nd} order polynomials.

For example, if $p = \sigma + j\omega$ and $\bar{p} = \sigma - j\omega$ are complex roots of $r(s)$:

$$(s-p)(s-\bar{p}) = s^2 - 2\sigma s + (\sigma^2 + \omega^2)$$

replace with in $G(s)$

Transfer functions

$$G(s) = \frac{q(s)}{r(s)}$$

Compactly gives us all information we need to predict major features of system response

- $y_h(t)$, modes, stability : all from $r(s)$ the denominator polynomial of $G(s)$

$$r(s) = \alpha_n \prod_{k=1}^{n} (s - p_k)$$

- forced response : Evaluate $G(s)$ at specific complex values of s.

Numerator Terms

Can also factor $q(s)$:

$$q(s) = \beta_m (s - z_1)(s - z_2) \cdots (s - z_m)$$

where $q(z_i) = \emptyset$ for $i = 1, \ldots, m$

The values z_i are called the **zeros** of $G(s)$

since
$$G(z_i) = \frac{\cancelto{\emptyset}{q(z_i)}}{r(z_i)} = \emptyset$$

The values p_k are called the **poles** of $G(s)$

since
$$G(p_k) = \frac{q(p_k)}{\cancelto{\emptyset}{r(p_k)}} = \infty$$

Zero/Pole/Gain (ZPK) form

$$G(s) = K\left[\frac{\prod_{i=1}^{m}(s-z_i)}{\prod_{k=1}^{n}(s-p_k)}\right]$$

Poles p_k satisfy $r(p_k) = \emptyset$

Zeros z_i satisfy $q(z_i) = \emptyset$

Gain: $K = \frac{\beta_m}{\alpha_n}$ (always real)

Alternate ZPK form:

When G(s) has complex poles and/or zeros, we commonly combine the conjugate roots of r(s) or q(s) into 2nd order polynomials.

For example, if $p = \sigma + j\omega$ and $\bar{p} = \sigma - j\omega$ are complex roots of r(s):

$$\underbrace{(s-p)(s-\bar{p})} = \underline{s^2 - 2\sigma s + (\sigma^2 + \omega^2)}$$

replace with in G(s)

Stability and $G(s)$

→ $G(s)$ is stable if all its poles are in LHP.

→ $G(s)$ is unstable if any of its poles are in RHP.

→ What role do zeros of $G(s)$ have in stability?

⟹ ABSOLUTELY NONE!

→ OK, so what role do zeros play?

Effect of zeros in $G(s)$

→ Certainly zeros influence the coefficients C_K of homogeneous response.

→ They also influence calculation of $y_f(t)$.

→ <u>Special example</u>: Suppose $u(t) = e^{z_i t}$

then:

$$y_f(t) = G(z_i) e^{z_i t} = \varnothing$$

The forced response is exactly zero here!

"Input absorbing" property of zeros

More complicated $u(t)$

$$u(t) = U e^{st} \Rightarrow y_f(t) = G(s) U e^{st}$$

Suppose $u(t) = U_1 e^{s_1 t} + U_2 e^{s_2 t}$

Substitute into DE, can show

$$y_f(t) = G(s_1) U_1 e^{s_1 t} + G(s_2) U_2 e^{s_2 t}$$

More complicated $u(t)$

$$u(t) = U e^{st} \implies y_f(t) = G(s) U e^{st}$$

Suppose $u(t) = \boxed{U_1 e^{s_1 t}} + \boxed{U_2 e^{s_2 t}}$

Substitute into DE, can show

$$y_f(t) = \boxed{G(s_1) U_1 e^{s_1 t}} + \boxed{G(s_2) U_2 e^{s_2 t}}$$

The sum of the responses to the individual parts of the input.

Linearity of Systems

If $y_1(t)$ is a possible sol'n of DE with input $u_1(t)$

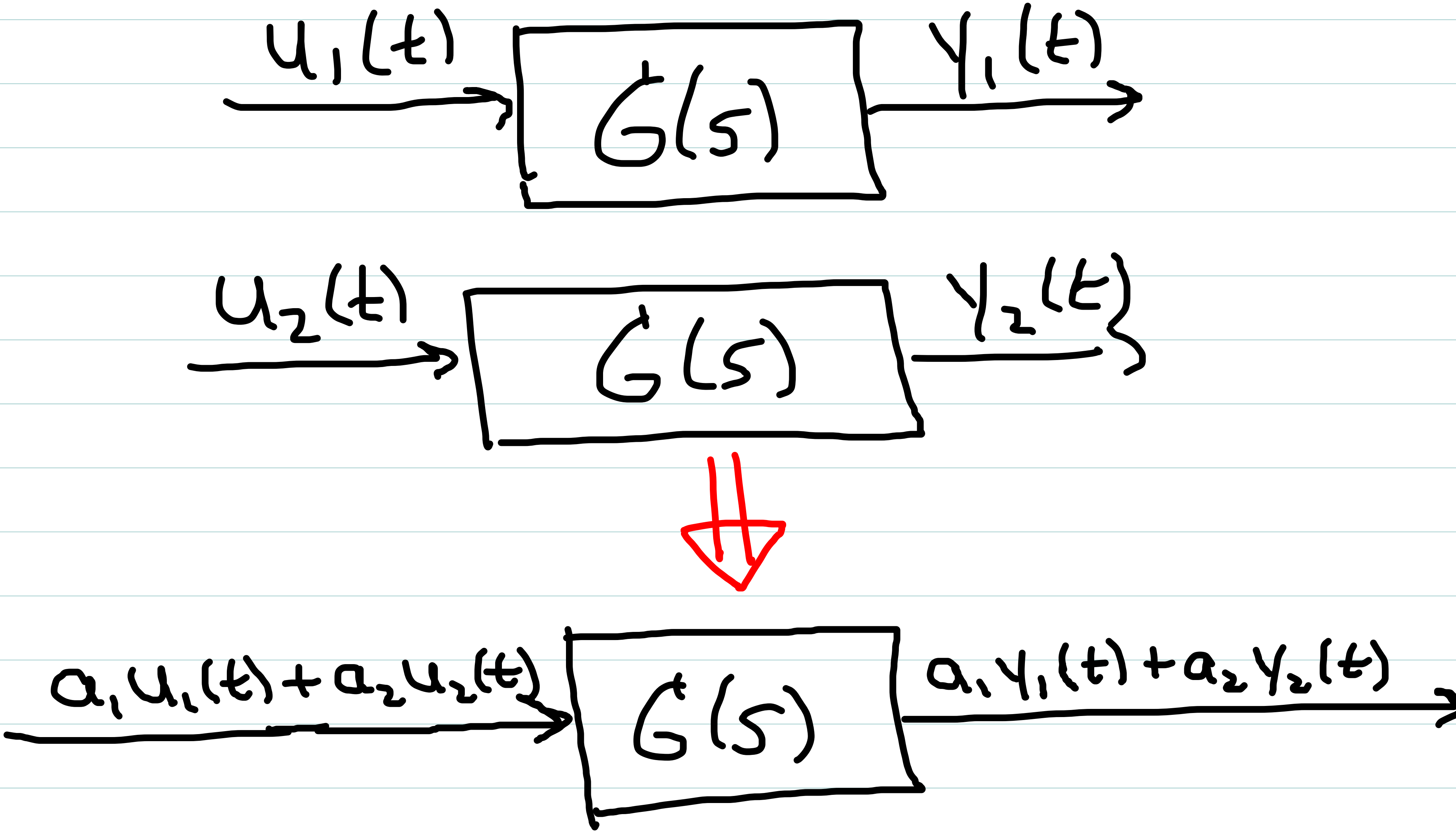
and similarly $y_2(t)$ is a sol'n for input $u_2(t)$

Then:

$$y(t) = a_1 y_1(t) + a_2 y_2(t)$$

is a sol'n for input $u(t) = a_1 u_1(t) + a_2 u_2(t)$

For any constants a_1, a_2 and any inputs $u_1(t), u_2(t)$


$u_1(t)$
$G(s)$
$y_1(t)$
$u_2(t)$
$G(s)$
$y_2(t)$
$a_1 u_1(t) + a_2 u_2(t)$
$G(s)$
$a_1 y_1(t) + a_2 y_2(t)$

We've already seen an example

$$u_1(t) = e^{st} \longrightarrow y_1(t) = G(s)e^{st}$$

$$u_2(t) = \emptyset \longrightarrow y_2(t) = y_h(t) = \sum_{k=1}^{n} c_k e^{p_k t}$$

$$u(t) = U e^{st} = U e^{st} + \emptyset$$

$$= U u_1(t) + u_2(t)$$

$$\Longrightarrow \quad y(t) = U y_1(t) + y_2(t)$$

$$= U G(s) e^{st} + y_h(t)$$

$$= y_f(t) + y_h(t)$$

Linearity can be used multiple times

$$u(t) = \sum_{i=1}^{N} a_i u_i(t) \implies y(t) = \sum_{i=1}^{N} a_i y_i(t)$$

$y_i(t)$ sol'n for $u_i(t)$

$\implies$ Holds for any number N

In particular,

$$u(t) = \sum_{i=1}^{N} U_i e^{s_i t} \implies y(t) = \sum_{i=1}^{N} G(s_i) U_i e^{s_i t}$$

Even for infinite sum, $N = \infty$.

Is this enough to make **any** $u(t)$?

Not quite, need to go to differential limit

$$\sum_{i=1}^{N} U_i e^{s_i t} \longrightarrow \int U(s) e^{st}\, ds$$

integral over all complex freqs

i.e. $u(t) = \int U(s) e^{st}\, ds$

$U(s)$ is the "amount" (complex amplitude) of e^{st} present in $u(t)$, for **each** $s \in \mathbb{C}$.

Similarly $y(t) = \sum_{i=1}^{N} G(s_i) U_i e^{s_i t} \to \int G(s) U(s) e^{st}\, ds$

OR: $y(t) = \int Y(s) e^{st}\, ds$ with $\boxed{Y(s) = G(s) U(s)}$

Laplace Transform

More formally, for any $f(t)$ define:

(1) $$f(t) = \frac{1}{2\pi j} \int F(s) e^{st} \, ds$$

Normalizing constant

Where:

(2) $$F(s) = \int_0^\infty f(t) e^{-st} \, dt$$

Notation: $F(s) = \mathcal{L}\{f(t)\}$ (transform)

$f(t) = \mathcal{L}^{-1}\{F(s)\}$ (inverse transform)

Limitations of Laplace Transform

Only defined for $f(t)$ where the integral (2) converges.

Requires: $$e^{-\sigma_0 t}|f(t)| \to \emptyset$$

for some finite $\sigma_0 \in \mathbb{R}$

The transform $F(s)$ is then defined for any

$s = \sigma + j\omega$ with $\sigma \geq \sigma_0$

and the integral (1) is over all values of s which satisfy this condition.] "region of convergence"

Examples

$f(t) = e^{pt}$ can be transformed for any finite $p \in \mathbb{C}$

However, $f(t) = e^{t^2}$ cannot be transformed since $e^{-\sigma_0 t} f(t) = e^{(t^2 - \sigma_0 t)} \to \infty$

for any finite σ_0.

Note:

When working with Laplace transforms we assume we are using values of s in the region of convergence. (ROC)

By above def'n of ROC,

$$\lim_{t \to \infty} e^{-st} f(t) = \emptyset$$

for these values of s.

Laplace Transform

More formally, for any $f(t)$ define:

$$(1) \qquad f(t) = \frac{1}{2\pi j} \int F(s) e^{st} \, ds$$

Normalizing constant

Where:

$$(2) \qquad F(s) = \int_0^{\infty} f(t) e^{-st} \, dt$$

Notation: $F(s) = \mathcal{L}\{f(t)\}$ (transform)

$f(t) = \mathcal{L}^{-1}\{F(s)\}$ (inverse transform)

Limitations of Laplace Transform

Only defined for $f(t)$ where the integral (2) converges.

Requires: $$e^{-\sigma_0 t}|f(t)| \to \emptyset$$

for some finite $\sigma_0 \in \mathbb{R}$

The transform $F(s)$ is then defined for any

$s = \sigma + j\omega$ with $\sigma \geq \sigma_0$

and the integral (1) is over all values of s which satisfy this condition. "region of convergence"

Examples

$f(t) = e^{pt}$ can be transformed for any finite $p \in \mathbb{C}$

However, $f(t) = e^{t^2}$ cannot be transformed

since $e^{-\sigma_0 t} f(t) = e^{(t^2 - \sigma_0 t)} \to \infty$

for any finite σ_0.

Note:

When working with Laplace transforms we assume we are using values of s in the region of convergence. (ROC)

By above def'n of ROC,

$$\lim_{t \to \infty} e^{-st} f(t) = \emptyset$$

for these values of s.

Fundamental Transform

(only one you need!)

$$\boxed{\mathcal{L}\{e^{pt}\} = \frac{1}{s-p}} \qquad \forall p \in \mathbb{C}$$

$$\mathcal{L}\{e^{pt}\} = \int_0^\infty e^{pt} e^{-st}\,dt = \int_0^\infty e^{(p-s)t}\,dt$$

$$= \left[\left(\frac{1}{p-s}\right) e^{(p-s)t}\right]_{t=0}^{t=\infty}$$

$$= \left(\frac{1}{p-s}\right)\left[\cancel{e^{(p-s)\infty}} - 1\right]$$

$$\emptyset$$

for any s in ROC

Property #1: Linearity

$$\boxed{\mathcal{L}\{a_1 f_1(t) + a_2 f_2(t)\} = a_1 F_1(s) + a_2 F_2(s)}$$

for any transformable functions $f_1(t), f_2(t)$
any (complex) constants a_1, a_2

$$\int_0^\infty \{a_1 f_1(t) + a_2 f_2(t)\} e^{-st}\,dt$$

$$= a_1 \underbrace{\boxed{\int_0^\infty f_1(t) e^{-st}\,dt}}_{F_1(s)} + a_2 \underbrace{\boxed{\int_0^\infty f_2(t) e^{-st}\,dt}}_{F_2(s)}$$

And generally:

$$\mathcal{L}\left\{\sum_{i=1}^{N} a_i f_i(t)\right\} = \sum_{i=1}^{N} a_i F_i(s)$$

Linearity lets us build more complex transforms:

Consider: $f(t) = Ae^{at}\cos(bt+\psi)$

$$= c e^{pt} + \bar{c} e^{\bar{p}t}$$

with $p = a + bj$, $c = \left(\frac{A}{2}\right)e^{j\psi}$ (polar form)

Then by linearity

$$\boxed{\mathcal{L}\{Ae^{at}\cos(bt+\psi)\} = \frac{c}{s-p} + \frac{\bar{c}}{(s-\bar{p})}}$$

We can combine the two terms:

$$\mathcal{L}\{Ae^{at}\cos(bt+\psi)\} = \frac{c}{s-p} + \frac{\bar{c}}{s-\bar{p}}$$

$$= \frac{A[(s-a)\cos\psi - b\sin\psi]}{s^2 - 2as + (a^2+b^2)}$$

so

$$\boxed{\mathcal{L}\{Ae^{at}\cos(bt+\psi)\} = \frac{A[(s-a)\cos\psi - b\sin\psi]}{(s-a)^2 + b^2}}$$

But we will see it is often easier to keep the two terms separate when solving problems.

Fundamental Transforms

$$\mathcal{L}\{e^{pt}\} = \frac{1}{s-p} \quad \text{for any } p \in \mathbb{C}$$

$$\mathcal{L}\{Ae^{at}\cos(bt+\varphi)\} = \frac{c}{s-p} + \frac{\bar{c}}{s-\bar{p}}$$

with $p = a + bj$ and $c = \left(\frac{A}{2}\right)e^{j\varphi}$

What is $\mathcal{L}\{c\}$ for an arbitrary constant c?

i.e. $\mathcal{L}\{f(t)\}$ with $f(t) = c$ for all $t \geq 0$

Fundamental Transforms

$$\mathcal{L}\{e^{pt}\} = \frac{1}{s-p} \quad \text{for any } p \in \mathbb{C}$$

$$\mathcal{L}\{Ae^{at}\cos(bt+\psi)\} = \frac{c}{s-p} + \frac{\bar{c}}{s-\bar{p}}$$

with $p = a + bj$ and $c = \left(\frac{A}{2}\right)e^{j\psi}$

What is $\mathcal{L}\{c\}$ for an arbitrary constant c?

i.e. $\mathcal{L}\{f(t)\}$ with $f(t) = c$ for all $t \geq 0$

$$f(t) = c = c\,e^{0t} \implies F(s) = \left.\frac{c}{s-p}\right|_{p=0}$$

Fundamental Transforms

$$\mathcal{L}\{e^{pt}\} = \frac{1}{s-p} \quad \text{for any } p \in \mathbb{C}$$

$$\mathcal{L}\{Ae^{at}\cos(bt+\varphi)\} = \frac{c}{s-p} + \frac{\bar{c}}{s-\bar{p}}$$

with $p = a + bj$ and $c = \left(\frac{A}{2}\right)e^{j\varphi}$

What is $\mathcal{L}\{c\}$ for an arbitrary constant c?

i.e. $\mathcal{L}\{f(t)\}$ with $f(t) = c$ for all $t \geq 0$

$$f(t) = c = ce^{0t} \implies F(s) = \left.\frac{c}{s-p}\right]_{p=0}$$

Hence $\boxed{\mathcal{L}\{c\} = \frac{c}{s}}$ for any $c \in \mathbb{C}$

Common Mistakes

$\mathcal{L}\{c\} \neq c$ $\quad$ $\left(\mathcal{L}\{c\} = \frac{c}{s}\right)$

$\mathcal{L}\{f_1(t) f_2(t)\} \neq F_1(s) F_2(s)$

$\mathcal{L}\{f_1(t) f_2(t)\} = \langle \text{unspeakably ugly} \rangle$

Property #2: Diff'n rule

$$\mathcal{L}\{\dot{f}(t)\} = sF(s) - f(0)$$

$$= \int_0^\infty \frac{df}{dt} e^{-st} dt$$

$$= \int_0^\infty e^{-st} df$$

$$\text{(by parts)} \quad = \left[e^{-st} f(t)\right]_{t=0}^{t=\infty} + s \underbrace{\boxed{\int_0^\infty f(t) e^{-st} dt}}_{F(s)}$$

$$= sF(s) - f(0)$$

Higher Derivatives

$$\mathcal{L}\{\ddot{f}(t)\} = \mathcal{L}\{\dot{f}_1(t)\} \quad \text{with } f_1(t) = \dot{f}(t)$$

$$= sF_1(s) - f_1(0)$$

but $F_1(s) = \mathcal{L}\{\dot{f}(t)\} = sF(s) - f(0)$

So $\boxed{\mathcal{L}\{\ddot{f}(t)\} = s^2F(s) - \dot{f}(0) - sf(0)}$

and generally

$$\boxed{\mathcal{L}\{f^{(k)}(t)\} = s^kF(s) - f^{(k-1)}(0) - sf^{(k-2)}(0) - \cdots - s^{k-1}f(0)}$$

__Note:__ Laplace will allow us to __directly__ account for IC effects (No Linear algebra!)

Property #3: "t-mult" rule

$$\boxed{\mathcal{L}\{tf(t)\} = -\frac{d}{ds}F(s)}$$

$$\begin{aligned}
\mathcal{L}\{tf(t)\} &= \int_0^\infty tf(t)e^{-st}\,dt \\
&= \int_0^\infty f(t)\left[te^{-st}\right]dt \\
&= \int_0^\infty f(t)\left[-\frac{d}{ds}e^{-st}\right]dt \\
&= \int_0^\infty -\frac{d}{ds}\left[f(t)e^{-st}\right]dt \\
&= -\frac{d}{ds}\int_0^\infty f(t)e^{-st}\,dt
\end{aligned}$$

Use of t-mult rule

$$\mathcal{L}\{te^{pt}\} = -\frac{d}{ds}\left[\frac{1}{s-p}\right]$$

$$= \frac{1}{(s-p)^2}$$

Similarly $\mathcal{L}\{t^2e^{pt}\} = \mathcal{L}\{t f_1(t)\}$, $f_1(t) = te^{pt}$

$$= -\frac{d}{ds}F_1(s) = -\frac{d}{ds}\left[\frac{1}{(s-p)^2}\right]$$

So $\mathcal{L}\{t^2e^{pt}\} = \dfrac{2}{(s-p)^3}$

Generally:

$$\boxed{\mathcal{L}\{t^k e^{pt}\} = \frac{k!}{(s-p)^{k+1}}}$$

Recap: Laplace Transform

$$f(t) = \mathcal{L}^{-1}\{F(s)\} = \frac{1}{2\pi j}\int F(s)e^{st}\,ds$$

$$F(s) = \mathcal{L}\{f(t)\} = \int_0^\infty f(t)e^{-st}\,dt$$

Properties:

1.) **Linearity**: $\mathcal{L}\left\{\sum_{i=1}^{N} a_i f_i(t)\right\} = \sum_{i=1}^{N} a_i F_i(s)$

2.) **Diff. rule**: $\mathcal{L}\{\dot{f}(t)\} = sF(s) - f(0)$

$$\mathcal{L}\{\ddot{f}(t)\} = s^2F(s) - \dot{f}(0) - sf(0)$$

$$\vdots$$

$$\mathcal{L}\{f^{(k)}(t)\} = s^kF(s) - f^{(k-1)}(0) - \cdots - s^{k-1}f(0)$$

Use of $\mathcal{L}T$ for Diff'l Eqn

$$\mathcal{L}\{\alpha_n y^{(n)}(t) + \cdots + \alpha_1 \dot{y}(t) + \alpha_0 y(t)\}$$

$$= \mathcal{L}\{\beta_m u^{(m)} + \cdots + \beta_1 \dot{u}(t) + \beta_0 u(t)\}$$

Gives:

$$\alpha_n\left[s^n Y(s) - y^{(n-1)}(0) - s y^{(n-2)}(0) - \cdots - s^{n-1} y(0)\right]$$

$$+ \cdots + \alpha_1\left[sY(s) - y(0)\right] + \alpha_0 Y(s)$$

$$= \beta_m\left[s^m U(s) - u^{(m-1)}(0) - s u^{(m-2)}(0) - \cdots - s^{m-1} u(0)\right]$$

$$+ \cdots + \beta_1\left[sU(s) - u(0)\right] + \beta_0 U(s)$$

Collect Terms

$$\boxed{r(s)Y(s) - c(s) = q(s)U(s) - b(s)}$$

$r(s) = \alpha_n s^n + \cdots + \alpha_1 s + \alpha_0$

$q(s) = \beta_m s^m + \cdots + \beta_1 s + \beta_0$

$c(s)$ = $n-1$ order polynomial in s from IC terms on $y(t)$

$b(s)$ = $m-1$ order polynomial in s from IC terms on $u(t)$.

Re-arrange for $Y(s)$

IC terms

$$Y(s) = \left[\frac{q(s)}{r(s)}\right]U(s) + \left[\frac{c(s)-b(s)}{r(s)}\right]$$

Or:

$$Y(s) = G(s)U(s) + \left[\frac{c(s)-b(s)}{r(s)}\right]$$

Alternate def'n of TF:

$$G(s) = \left[\frac{Y(s)}{U(s)}\right]_{ICs=0} = \left[\frac{\mathcal{L}\{y(t)\}}{\mathcal{L}\{u(t)\}}\right]_{\substack{\text{zero}\\ \text{ICs}}}$$

Example

$$2y^{(3)} + 8\ddot{y} + 14\dot{y} + 10y = 3\ddot{u} + 15\dot{u} + 18u$$

$$2\left[s^3Y(s) - \ddot{y}_0 - s\dot{y}_0 - s^2y_0\right] + 8\left[s^2Y(s) - \dot{y}_0 - sy_0\right]$$

$$+ 14\left[sY(s) - y_0\right] + 10Y(s)$$

$$= 3\left[s^2U(s) - \dot{u}_0 - su_0\right] + 15\left[sU(s) - u_0\right] + 18U(s)$$

OR:

$$(2s^3 + 8s^2 + 14s + 10)Y(s)$$
$$- \left[2s^2y_0 + s(2\dot{y}_0 + 8y_0) + (2\ddot{y}_0 + 8\dot{y}_0 + 14y_0)\right]$$
$$= (3s^2 + 15s + 18)U(s) - \left[3su_0 + (3\dot{u}_0 + 15u_0)\right]$$

Example

$$2y^{(3)} + 8\ddot{y} + 14\dot{y} + 10y = 3\ddot{u} + 15\dot{u} + 18u$$

$$2[s^3Y(s) - \ddot{y}_0 - s\dot{y}_0 - s^2y_0] + 8[s^2Y(s) - \dot{y}_0 - sy_0]$$

$$+ 14[sY(s) - y_0] + 10Y(s)$$

$$= 3[s^2U(s) - \dot{u}_0 - su_0] + 15[sU(s) - u_0] + 18U(s)$$

OR:

$$\overbrace{(2s^3 + 8s^2 + 14s + 10)}^{r(s)}Y(s) - [2s^2y_0 + s(2\dot{y}_0 + 8y_0) + (2\ddot{y}_0 + 8\dot{y}_0 + 14y_0)]$$

$$= \underbrace{(3s^2 + 15s + 18)}_{q(s)}U(s) - [3su_0 + (3\dot{u}_0 + 15u_0)]$$

Example

$$2y^{(3)} + 8\ddot{y} + 14\dot{y} + 10y = 3\ddot{u} + 15\dot{u} + 18u$$

$$2\left[s^3Y(s) - \ddot{y}_0 - s\dot{y}_0 - s^2y_0\right] + 8\left[s^2Y(s) - \dot{y}_0 - sy_0\right]$$

$$+ 14\left[sY(s) - y_0\right] + 10Y(s)$$

$$= 3\left[s^2U(s) - \dot{u}_0 - su_0\right] + 15\left[sU(s) - u_0\right] + 18U(s)$$

OR:

$$(2s^3 + 8s^2 + 14s + 10)Y(s)$$

$$- \underbrace{\left[2s^2y_0 + s(2\dot{y}_0 + 8y_0) + (2\ddot{y}_0 + 8\dot{y}_0 + 14y_0)\right]}_{c(s)}$$

$$= (3s^2 + 15s + 18)U(s) - \underbrace{\left[3su_0 + (3\dot{u}_0 + 15u_0)\right]}_{b(s)}$$

Thus:

$$Y(s) = \overbrace{\left[\frac{3s^2+15s+18}{2s^3+8s^2+14s+10}\right]}^{G(s)} U(s)$$

$$+\left[\frac{2s^2y_0+s(2\dot{y}_0+8y_0-3u_0)+(2\ddot{y}_0+8\dot{y}_0+14y_0-3\dot{u}_0-15u_0)}{2s^3+8s^2+14s+10}\right]$$

$\Rightarrow$ We assume all ICs on $y(t)$ known; and $u(t)$ known so $U(s)$ can be computed and ICs on $u(t)$

$\Rightarrow$ All terms on RHS are known, so we know $Y(s)$

$\Rightarrow$ "simply" invert transform to get $y(t)$

$$y(t) = \mathcal{L}^{-1}\{Y(s)\}$$

Inverse Transform

$$y(t) = \mathcal{L}^{-1}\{Y(s)\}$$
$$= \frac{1}{2\pi j}\int Y(s)e^{st}\,ds$$

$\Rightarrow$ contour integral over ROC in complex plane

$\Rightarrow$ ugly! Math 463

$\Rightarrow$ We can sidestep this in many cases

General form of $Y(s)$

$$Y(s) = \left[\frac{q(s)}{r(s)}\right] U(s) + \left[\frac{c(s) - b(s)}{r(s)}\right]$$

all polynomials

Suppose $U(s)$ is <u>rational</u> in s
(ratio of polynomials)

i.e. $U(s) = \dfrac{a(s)}{h(s)}$, $a(s)$ $h(s)$ polys

<u>Note:</u>
1. Not true for every $u(t)$
2. True for <u>many</u> "useful" $u(t)$

Then...

$$U(s)$$

$$Y(s) = \left[\frac{q(s)}{r(s)}\right]\left(\frac{a(s)}{h(s)}\right) + \frac{c(s)-b(s)}{r(s)}$$

$$= \frac{q(s)a(s) + h(s)\left[c(s)-b(s)\right]}{r(s)h(s)}$$

or

$$Y(s) = \frac{N(s)}{D(s)}$$

where both $N(s)$ and $D(s)$ are polynomials

(i.e. $Y(s)$ is rational)

$$Y(s) = \frac{N(s)}{D(s)}$$

Suppose $\deg\{N(s)\} < \deg\{D(s)\} = L$

Let d_ℓ be the roots of $D(s)$: $D(d_\ell) = \phi$

Then:

$$Y(s) = \frac{A_1}{s-d_1} + \frac{A_2}{s-d_2} + \cdots + \frac{A_L}{s-d_L}$$

$$= \sum_{\ell=1}^{L} \frac{A_\ell}{s-d_\ell}$$

"Partial fraction expansion"

and

$$y(t) = \sum_{\ell=1}^{L} A_\ell e^{d_\ell t}$$

How to find expansion coefficients

"Residue formula":

$$A_\ell = \left[(s-d_\ell)\,Y(s)\right]_{s=d_\ell}$$

(also called "cover up" rule).

Example:

$$Y(s) = \frac{2s+3}{(s+2)(s+3)}$$

$$Y(s) = \frac{A_1}{s+2} + \frac{A_2}{s+3}$$

$$A_1 = \left[\frac{2s+3}{s+3}\right]_{s=-2} = -1, \quad A_2 = \left[\frac{2s+3}{s+2}\right]_{s=-3} = 3$$

$$y(t) = 3e^{-3t} - e^{-2t}$$

Complex d_ℓ

Note if d_ℓ is a complex root of $D(s)$, then its conjugate $\bar{d}_\ell$ will also be a root.

The residue formula then tells us that

for d_ℓ: $\quad A_\ell = \left[(s-d_\ell)Y(s)\right]_{s=d_\ell}$

and for $\bar{d}_\ell$ we instead have

$$\left[(s-\bar{d}_\ell)Y(s)\right]_{s=\bar{d}_\ell} = \bar{A}_\ell$$

i.e. the PFE coefficients are also conjugates

Complex d_ℓ (cont)

Thus, the expression for $y(t)$ will contain

$$A_\ell e^{d_\ell t} + \bar{A}_\ell e^{\bar{d}_\ell t}$$

$$= 2|A_\ell| e^{\sigma t} \cos(\omega t + \measuredangle A_\ell)$$

where $\sigma = \mathrm{Re}\{d_\ell\}$ $\quad \omega = \mathrm{Im}\{d_\ell\}$

Example:

$$Y(s) = \frac{4(s^2+2s+6)}{(s+1)(s^2+4s+13)}$$

$$d_1 = -1; \quad d_2 = -2+3j; \quad d_3 = -2-3j = \bar{d}_2$$

Then:

$$A_1 = \left[(s+1)Y(s)\right]_{s=-1} = 2$$

$$A_2 = \left[(s+2-3j)Y(s)\right]_{s=-2+3j} = 1+j = \boxed{\sqrt{2}\angle^{\pi/4} = A_2}$$

$$A_3 = \left[(s+2+3j)Y(s)\right]_{s=-2-3j} = 1-j = \bar{A}_2$$

Hence:

$$y(t) = 2e^{-t} + \boxed{(1+j)e^{(-2+3j)t} + (1-j)e^{(-2-3j)t}}$$

or:

$$y(t) = 2e^{-t} + \boxed{2\sqrt{2}\, e^{-2t}\cos(3t+\pi/4)}$$

Recap

$$Y(s) = \underbrace{\left[\frac{q(s)}{r(s)}\right]}_{G(s)} U(s) + \underbrace{\left[\frac{c(s)-b(s)}{r(s)}\right]}_{\text{IC terms}}$$

If $U(s)$ rational, $U(s) = \dfrac{a(s)}{h(s)}$

Then $\quad Y(s) = \dfrac{N(s)}{D(s)} \quad$ (also rational)

$$= \sum_{\ell=1}^{L} \frac{A_\ell}{(s-d_\ell)} \quad \text{where} \quad D(d_\ell) = \emptyset$$

and $\quad A_\ell = \left[(s-d_\ell)Y(s)\right]_{s=d_\ell}$

Inverse transform:

$$y(t) = \sum_{\ell=1}^{L} A_\ell e^{d_\ell t}$$

Assumptions

Above assumes:

(1.) $\deg\{N(s)\} < \deg\{D(s)\}$

(2.) No repeated roots of $D(s)$

Simplest, most common case

Both can be relaxed:

(1.) Suppose $\mathrm{Deg}\{N(s)\} = \mathrm{Deg}\{D(s)\}$

Then do polynomial long division:

$$Y(s) = \frac{N(s)}{D(s)} = A_0 + \frac{N_1(s)}{D(s)}, \quad \mathrm{Deg}\{N_1(s)\} < \mathrm{Deg}\{D(s)\}$$

and $\dfrac{N_1(s)}{D(s)}$ can be expanded using above

So:

$$Y(s) = \frac{N(s)}{D(s)} = A_0 + \frac{N_1(s)}{D(s)}$$

$$= A_0 + \sum_{\ell=1}^{L} \frac{A_\ell}{(s-d_\ell)}$$

PFE

Where:

$$A_\ell = \left[(s-d_\ell)\frac{N_1(s)}{D(s)}\right]_{s=d_\ell}$$

<u>Inverse transforming</u>:

$$y(t) = \mathcal{L}^{-1}\{A_0\} + \sum_{\ell=1}^{L} A_\ell e^{d_\ell t}$$

What is this?? We'll see later...

<u>Note:</u> $\text{Deg}\{N(s)\} > \text{Deg}\{D(s)\}$
nonphysical + won't be seen

Repeated Roots

Now suppose:

$$D(s) = (s-d_1)^K (s-d_{K+1}) \cdots (s-d_L)$$

i.e. d_1 is repeated K times, then:

$$Y(s) = \sum_{\ell=1}^{K} \frac{A_\ell}{(s-d_1)^\ell} + \sum_{\ell=K+1}^{L} \frac{A_\ell}{(s-d_\ell)}$$

<u>for $\ell = K+1, \ldots, L$</u> :

$$A_\ell = \left[(s-d_\ell)Y(s)\right]_{s=d_\ell}$$ (unchanged)

<u>for $\ell = 1, \ldots, K$</u> :

(ugh!)

$$A_\ell = \frac{1}{(K-\ell)!} \left\{ \frac{d^{K-\ell}}{ds^{K-\ell}} \left[(s-d_1)^K Y(s)\right] \right\}_{s=d_1}$$

Inverse Transform (Repeated Roots)

$$Y(s) = \sum_{\ell=1}^{K} \frac{A_\ell}{(s-d_1)^\ell} + \sum_{\ell=K+1}^{L} \frac{A_\ell}{(s-d_\ell)}$$

$$\Rightarrow \quad y(t) = \sum_{\ell=1}^{K} \left(\frac{A_\ell t^{\ell-1}}{(\ell-1)!} \right) e^{d_1 t} + \sum_{\ell=K+1}^{L} A_\ell e^{d_\ell t}$$

Example:

$$Y(s) = \frac{2s+1}{(s+1)^3(s+2)} \qquad d_1 = -1, \; K=3$$
$$d_4 = -2$$

$$\Rightarrow y(t) = \left[A_1 + A_2 t + \frac{A_3}{2} t^2 \right] e^{-t} + A_4 e^{-2t}$$

$$A_3 = \left[(s+1)^3 Y(s) \right]_{s=-1} = -1$$

$$A_2 = \left(\frac{1}{1}\right) \left\{ \frac{d}{ds} \left[(s+1)^3 Y(s) \right] \right\}_{s=-1} = \left[\frac{3}{(s+2)^2} \right]_{s=-1} = 3$$

$$A_1 = \left(\frac{1}{2}\right)\left\{\frac{d^2}{ds^2}\left[(s+1)^3 Y(s)\right]\right\}_{s=-1}$$

$$= \left(\frac{1}{2}\right)\left\{\frac{d}{ds}\left[\frac{3}{(s+2)^2}\right]\right\}_{s=-1} = -3$$

And

$$A_4 = \left[(s+2)Y(s)\right]_{s=-2} = 3$$

So finally:

$$y(t) = \left[-3 + 3t - \frac{1}{2}t^2\right]e^{-t} + 3e^{-2t}$$

<u>Note</u>: You aren't responsible for repeated root residue formula. <u>However</u> you should know the general pattern for repeated root solutions.

Alternate System Models

- A dynamical analysis does not always result in a high-order DE Directly connecting $u(t)$ and $y(t)$
- Sometimes the analysis (initially) results in a system of 1^{st} order DEs describing the evolution of the Dynamics
- Each first order equation describes the rate of change of a single physical variable (like airspeed, pitch angle, and angle of attack)
- Generically label these $x_K(t)$ $(K=1\ldots n)$ Known as the $\boxed{\text{state variables}}$ for the system.

"State variable" form of Dynamics

System of 1st order DEs describing how rate of change in each state depends on other states and forcing input

rate of change of each state — Linear combination of states — effect of input

$$\left\{\begin{aligned}
\dot{x}_1(t) &= \overbrace{a_{11}x_1(t) + a_{12}x_2(t) + \cdots + a_{1n}x_n(t)}^{\text{Linear combination of states}} + \overbrace{b_1 u(t)}^{\text{effect of input}} \\
\dot{x}_2(t) &= a_{21}x_1(t) + a_{22}x_2(t) + \cdots + a_{2n}x_n(t) + b_2 u(t) \\
&\ \ \vdots \\
\dot{x}_n(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + \cdots + a_{nn}x_n(t) + b_n u(t)
\end{aligned}\right.$$

$\rightarrow$ **n 1st order DEs**

$\Rightarrow$ easier to represent in Matrix/Vector form

"State-space" Dynamical model

$$\dot{\underset{\sim}{x}}(t) = A\underset{\sim}{x}(t) + B\,u(t)$$ ← "state equation"

with:

$$\underset{\sim}{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} \qquad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad (n \times 1)$$

"state vector"

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad (n \times n)$$

What about output?

Output $y(t)$ can be any 1 of the states, or any weighted combination of states (and input) as appropriate.

i.e. $y(t) = C_1 x_1(t) + C_2 x_2(t) + \cdots + C_n x_n(t) + D u(t)$

or $y(t) = C \underline{x}(t) + D u(t)$ "output equation"

where $C = [C_1 \ C_2 \ \ldots \ C_n]$ $(1 \times n)$

So complete model is

Standard "state-space" model of dynamics

$$\boxed{\begin{aligned} \dot{\underline{x}}(t) &= A \underline{x}(t) + B u(t) \\ y(t) &= C \underline{x}(t) + D u(t) \end{aligned}}$$

Example from HW#1

fan: $I\dot{\omega} = K_m i_m - D\omega$

motor: $L\dfrac{di_m}{dt} = V_m - Ri_m - E\omega$

velocity

vehicle: $m\ddot{y} = K_f\omega \Rightarrow m\dot{v} = K_f\omega, \quad \dot{y} = \underline{v}$

So:

$$\frac{d}{dt}\underbrace{\begin{bmatrix}\omega\\ i_m\\ v\\ y\end{bmatrix}}_{\underline{x}(t)} = \underbrace{\begin{bmatrix}-D/I & K_m/I & 0 & 0\\ -E/L & -R/L & 0 & 0\\ K_f/m & 0 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}}_{A}\begin{bmatrix}\omega\\ i_m\\ v\\ y\end{bmatrix} + \underbrace{\begin{bmatrix}0\\ 1/L\\ 0\\ 0\end{bmatrix}}_{B}\underbrace{V_m}_{u}$$

$$y(t) = \underbrace{\begin{bmatrix}0 & 0 & 0 & 1\end{bmatrix}}_{C}\underline{x}(t) + \underbrace{0}_{D}\,u(t)$$

Where is $G(s)$ for this model?

Not as easy to see transfer function by inspection.

But, we can still use laplace.

Laplace can be applied to vectors too, just apply it to each component of the vector

$$\mathcal{L}\{\underline{x}(t)\} = \underline{X}(s) = \begin{bmatrix} \mathcal{L}\{x_1(t)\} \\ \vdots \\ \mathcal{L}\{x_n(t)\} \end{bmatrix} = \begin{bmatrix} X_1(s) \\ \vdots \\ X_n(s) \end{bmatrix}$$

Lineararity: $\mathcal{L}\{A\underline{x}_1(t) + B\underline{x}_2(t)\} = A\underline{x}_1(s) + B\underline{x}_2(s)$

Derivative rule

$$\mathcal{L}\{\dot{\underline{x}}(t)\} = \begin{bmatrix} sX_1(s) - X_1(0) \\ sX_2(s) - X_2(0) \\ \vdots \\ sX_n(s) - X_n(0) \end{bmatrix} = s\underline{X}(s) - \underline{X}_0$$

Apply Laplace to State space model

$$\dot{\underline{x}}(t) = A\underline{x}(t) + B u(t)$$
$$y(t) = C\underline{x}(t) + D u(t)$$

$$\underline{x}_0 = \begin{bmatrix} x_1(0) \\ x_2(0) \\ \vdots \\ x_n(0) \end{bmatrix}$$

Initial state values

$$\Rightarrow \quad s\underline{x}(s) - \underline{x}_0 = A\underline{x}(s) + B u(s)$$
$$y(s) = C\underline{x}(s) + D u(s)$$

1^{st} eq'n is equivalent to:

$$(s\mathbb{I} - A)\underline{x}(s) = \underline{x}_0 + B u(s) \qquad (\mathbb{I} = n \times n \text{ identity})$$

$$\Rightarrow \quad \underline{x}(s) = [s\mathbb{I} - A]^{-1}[\underline{x}_0 + B u(s)]$$

Substitute into 2^{nd} eq'n:

$$\boxed{y(s) = C[s\mathbb{I} - A]^{-1}\underline{x}_0 + [C(s\mathbb{I} - A)^{-1}B + D]U(s)}$$

$$y(s) = \overbrace{\underbrace{C[sI-A]^{-1}\underline{x}_0}_{\text{effect of ICs}}}^{} + \overbrace{\underbrace{[C(sI-A)^{-1}B+D]U(s)}_{\text{effect of input}}}^{}$$

Recall: TF derived assuming ICs $=0 \Rightarrow \underline{x}_0 = \underline{0}$

Then

$$y(s) = \overbrace{\boxed{[C(sI-A)^{-1}B+D]}}^{G(s)}U(s)$$

Hence, for any (A,B,C,D) state space representation the corresponding transfer function is:

$$G(s) = [C\overbrace{(sI-A)^{-1}}^{n\times n \text{ matrix inverse}}B+D]$$

Now recall for arbitrary matrix M

$$M^{-1} = \frac{\mathrm{Adj}(M)}{\mathrm{Det}(M)}$$

Adj = $n\times n$ Matrix of cofactors

Det = <u>Scalar</u> Determinant

Thus

$$(s\mathbb{I}-A)^{-1} = \frac{Q(s)}{r(s)}$$

where

$$\begin{cases} Q(s) = \mathrm{Adj}(s\mathbb{I}-A) & (n\times n \text{ matrix.}) \\ r(s) = \mathrm{Det}(s\mathbb{I}-A) & \underline{\text{polynomial}} \text{ in } s. \end{cases}$$

and

$$G(s) = \frac{CQ(s)B}{r(s)} + D = \frac{CQ(s)B + Dr(s)}{r(s)}$$

where both $CQ(s)B$ and $r(s)$ are polynomials

$\Rightarrow$ zeros where $\boxed{CQ(s)B + Dr(s)} = 0$ $\quad q(s)$

$\Rightarrow$ poles where $r(s) = 0$.

So the poles of $G(s)$ will satisfy

$$r(s) = 0 = \mathrm{Det}(sI - A)$$

$\Rightarrow$ $(sI - A)$ is <u>singular</u>, at these values of s, i.e. there exists <u>nonzero</u> $\underline{v}$

so that

$$(sI - A)\underline{v} = \underline{0}$$

(singular matrices have nontrivial nullspace)

or:

$$A\underline{v} = s\underline{v} \quad \text{for any } s \text{ with } r(s) = 0$$

$$\boxed{\Rightarrow \text{poles of } G(s) \text{ are } \boxed{\underline{\text{eigenvalues}}} \text{ of } A!!}$$

Converting from $G(s)$ to state space

- Sometimes it is useful to reverse process described above, i.e.

  Given $G(s)$, find $[A, B, C, D]$

- In fact, for a given $G(s)$ there are infinitely many equivalent $[A, B, C, D]$

  $\Rightarrow$ many more DoF in $A(n^2)$, $B(n)$, $C(n)$ than in polynomials $r(s)$ $(n+1)$ and $q(s)$ $(m \leq n)$

- One "canonical" conversion is easy to obtain where the coefs. of polys $r(s)$ and $q(s)$ appear as rows and/or cols of $[A, B, C]$

- Known as "companion forms"

Companion form (one possibility)

Given $$G(s) = \frac{q(s)}{r(s)} = \frac{\beta_{n-1} s^{n-1} + \cdots + \beta_1 s + \beta_0}{s^n + \alpha_{n-1} s^{n-1} + \cdots + \alpha_1 s + \alpha_0}$$

Take
$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 \cdots & 0 \\ & & & & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & \cdots & & -\alpha_{n-1} \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$$C = \begin{bmatrix} \beta_0 & \beta_1 & \cdots & \beta_{n-1} \end{bmatrix} \qquad D = 0$$

(Note: $A_1 = A^T$, $B_1 = C^T$, $C_1 = B^T$ works too!)

Example

$$G(s) = \frac{3s^2 - 4s + 5}{s^3 + 2s^2 - s + 7}$$

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -7 & 1 & -2 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 5 & -4 & 3 \end{bmatrix} \qquad D = 0$$

One possible state space model for this TF

Philosophical Question: What is $t=0$?

=> The instant we start acting on the system with external input.

=> In control theory, we assume these inputs are completely "off" for $t<0$.

=> $u(t)$, $\dot{u}(t)$, $\ddot{u}(t)$, etc all <u>zero</u> for $t<0$

=> <u>Discontinuities</u> exist when $u(0) \neq 0$

$$u(t) = \begin{cases} e^{-t}, & t \geq 0 \\ 0 & t<0 \end{cases}$$

$$u(t) = \begin{cases} e^{pt}, & t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\Rightarrow \boxed{u(t) = e^{pt}\,\mathbb{1}(t)}$$

where

$$\boxed{\mathbb{1}(t) = \begin{cases} 1 & t \geq 0 \\ 0 & \text{otherwise} \end{cases}}$$

"Unit step function" (Very important!)

Now, Laplace is concerned about behavior of functions only for $t \geq 0$.

For all intents and purposes, functions in Laplace are considered 0 for $t < 0$

Implication

Formally:

$$\mathcal{L}^{-1}\left\{\frac{1}{s-p}\right\} = e^{pt}\,\mathbb{1}(t)$$

$$= \begin{cases} e^{pt}, & t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Now generally, our diff'l eq'ns will involve <u>derivatives</u> of these discontinuous functions

$\Rightarrow$ creates <u>singularities</u> in analysis at $t=0$

$$\frac{d}{dt}\mathbb{1}(t) = \begin{cases} 0 & t \neq 0 \\ \end{cases}$$

Implication

Formally:

$$\mathcal{L}^{-1}\left\{\frac{1}{s-p}\right\} = e^{pt}\,\mathbb{1}(t)$$

$$= \begin{cases} e^{pt}, & t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Now generally, our diff'l eq'ns will involve <u>derivatives</u> of these discontinuous functions

$\Rightarrow$ creates <u>singularities</u> in analysis at $t=0$

$$\frac{d}{dt}\mathbb{1}(t) = \begin{cases} 0 & t \neq 0 \\ \infty & t = 0 \ (???) \end{cases}$$

Theoretical problems in integrals when discontinuities or singularities at one of the end points.

$$F(s) = \int_0^{\infty} f(t) e^{-st} dt$$

possible problem here

Resolve these by taking lower limit at $t = \phi^-$ (the instant before $t = \phi$).

$\Rightarrow$ integral "sees" effect of singularities at $t = \phi$.

Starting the integral at 0^- instead of 0

- Avoids singularities at endpoints
- Causes transform to "see" singularities and discontinuities at $t=0$, so their effects will be reflected in the solutions for $y(t)$.

Hence:

$$F(s) = \int_{0^-}^{\infty} f(t) e^{-st} \, dt$$

<u>Implications</u>:

Assumed ICs for $y(t)$: just <u>before</u> $t=0$

$$\mathcal{L}\{\dot{y}(t)\} = sY(s) - y(0^-)$$

$$\mathcal{L}\{\ddot{y}(t)\} = s^2Y(s) - \dot{y}(0^-) - sy(0^-)$$

etc.

$$\mathcal{L}\{\dot{u}(t)\} = sU(s) - u(0^-)$$

$$\mathcal{L}\{\ddot{u}(t)\} = s^2U(s) - \dot{u}(0^-) - su(0^-)$$

etc.

Always $= 0$ in our analysis!

Thus:

$$Y(s) = \left[\frac{q(s)}{r(s)}\right]U(s) + \left[\frac{c(s) - \cancel{b(s)}^{\,\emptyset}}{r(s)}\right]$$

$$= \left[\frac{q(s)}{r(s)}\right]U(s) + \left[\frac{c(s)}{r(s)}\right]$$

$\Rightarrow$ IC polynomial $b(s)$ for input vanishes

$\Rightarrow$ specific to controls convention for $u(t)$

$\Rightarrow$ Not a common assumption in regular math classes.

$\Rightarrow$ In controls, want to know effect of discontinuities

Common, discontinuous "test functions"

$u(t) = \mathbb{1}(t)$ (unit step) $\mathcal{L}\{\mathbb{1}(t)\} = \frac{1}{s}$

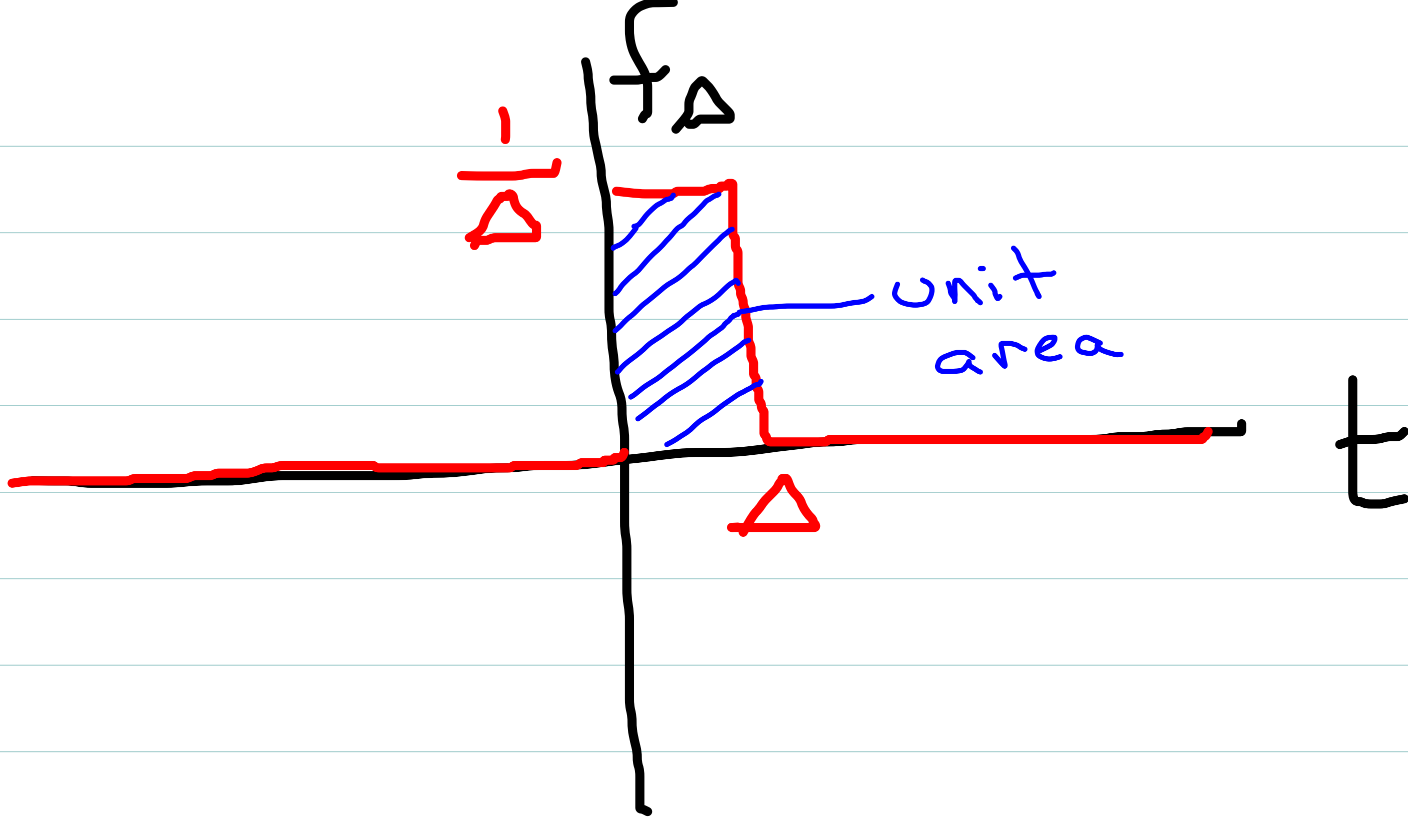
$$u(t) = \cos(\omega t)\,\mathbb{1}(t)$$

$$= \begin{cases} \cos(\omega t) & t \geq \phi \\ \phi & t < \phi \end{cases}$$

$$\Longrightarrow \quad u(t) = f_\Delta(t)$$

$$= \begin{cases} \frac{1}{\Delta} & \phi \leq t \leq \Delta \\ \phi & \text{otherwise} \end{cases}$$

"Unit <u>pulse</u> function"



Note: for <u>any</u> $\Delta > 0$

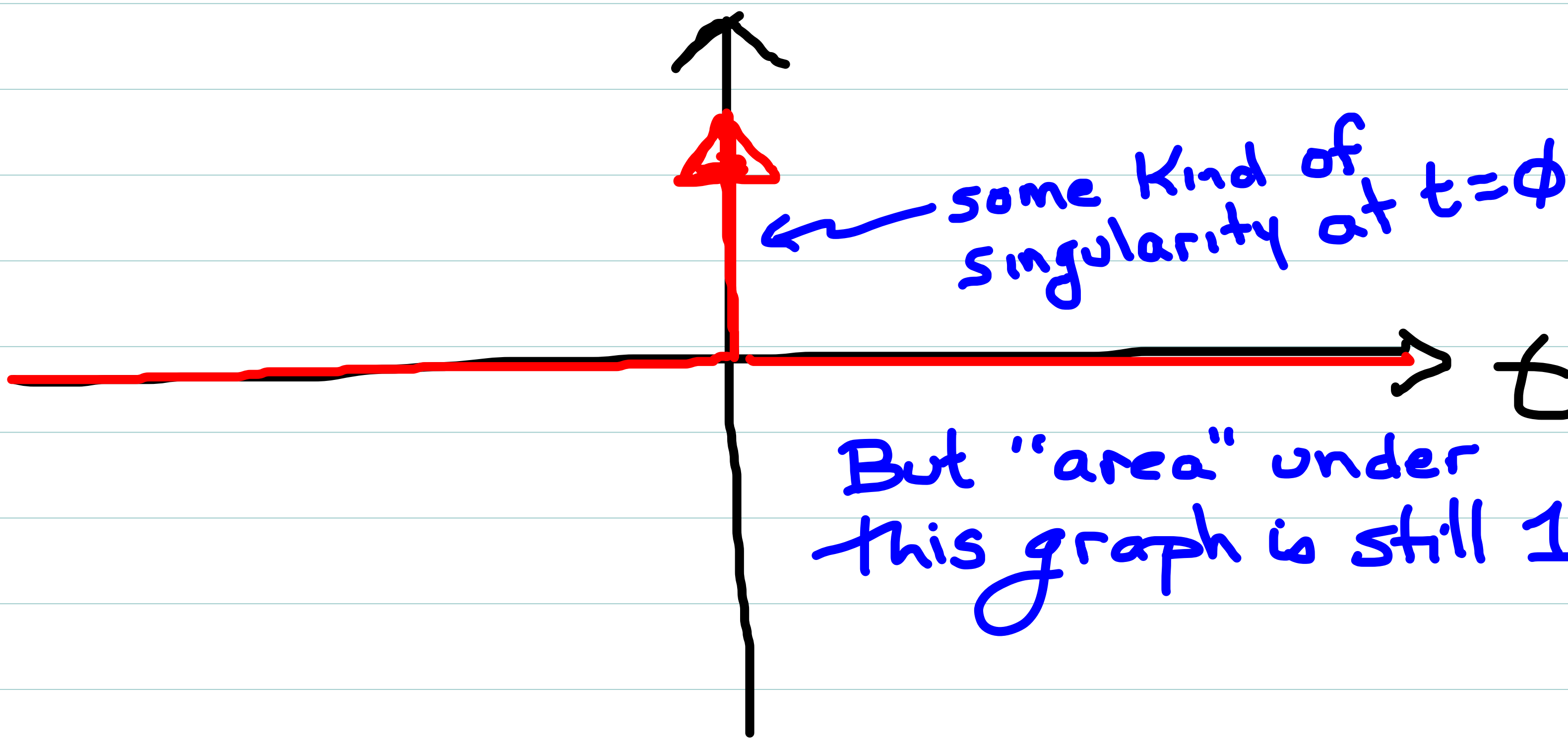
$$\int_{0^-}^{\infty} f_\Delta(t)\,dt = \int_{0^-}^{\Delta} \left(\frac{1}{\Delta}\right) dt$$

$$= 1$$

What is $\lim_{\Delta \to 0} f_\Delta(t)$?

$$= \lim_{\Delta \to 0} \begin{cases} \frac{1}{\Delta} & \phi \le t \le \Delta \\ \phi & \text{otherwise} \end{cases}$$

$$= \begin{cases} \infty & t = \phi \\ \phi & \text{otherwise} \end{cases}$$



But "area" under this graph is still 1...?

Define:

$$\boxed{\delta(t) = \lim_{\Delta \to 0} f_\Delta(t)}$$

"ideal impulse": models delivering a unit of input energy over negligibly small time.

(sharp "kick")

Alternate names:

"delta function"
"impulse function"
"Dirac delta"

Note: Not really a meaningful function at all!

More formally, belongs to a class of mathematical objects called

"distributions" or "generalized functions"

Suppose $\delta(t)$ appears in an integral

$$\int_{-\infty}^{\infty} \delta(t)\, h(t)\, dt \quad , \quad h(t) \text{ arbitrary f'n}$$

$$= \int_{-\infty}^{\infty} \left[\lim_{\Delta \to 0} f_\Delta(t)\right] h(t)\, dt$$

$$= \lim_{\Delta \to 0} \int_{-\infty}^{\infty} f_\Delta(t)\, h(t)\, dt$$

$$= \lim_{\Delta \to 0} \left\{ \frac{1}{\Delta} \int_{0^-}^{\Delta} h(t)\, dt \right\}$$

$$\cong \left(\frac{1}{\Delta}\right)\left(\Delta\, h(\emptyset)\right)$$

$$= h(\emptyset)$$

Note: with $h(t) = 1$ for all t, we get

$$\int_{0^-}^{\infty} \delta(t)\, dt = 1$$

Defining Property of $\delta(t)$

$$\int_a^b \delta(t)h(t)\,dt = \begin{cases} h(\phi) & \text{if } \phi \in (a,b) \\ \phi & \text{otherwise} \end{cases}$$

"sifting property"

⟹ $\delta(t)$ is defined by what it does in an integral

Not as an ordinary function

Now we can compute:

$$\mathcal{L}\{\delta(t)\} = \int_{0^-}^{\infty} \delta(t)e^{-st}\,dt$$

$$= \left[e^{-st}\right]_{t=\phi} = 1$$

Thus:

$$\boxed{\mathcal{L}\{\delta(t)\} = 1}$$

and by Linearity:

$$\mathcal{L}\{c\,\delta(t)\} = c \quad \text{for any constant } c.$$

Now recall $\frac{d}{dt}\mathbb{1}(t) = \begin{cases} \infty & t=0 \\ 0 & \text{otherwise} \end{cases}$

which looks like $\frac{d}{dt}\mathbb{1}(t) = \delta(t)$.

Is this formally true?

$$\mathcal{L}\left\{\frac{d}{dt}\mathbb{1}(t)\right\} = s\,\mathcal{L}\{\mathbb{1}(t)\} - \mathbb{1}(0^-)$$

Thus:

$$\mathcal{L}\{\delta(t)\} = 1$$

and by Linearity:

$$\mathcal{L}\{c\,\delta(t)\} = c \quad \text{for any constant } c.$$

Now recall $\displaystyle \frac{d}{dt}\mathbb{1}(t) = \begin{cases} \infty & t=0 \\ 0 & \text{otherwise} \end{cases}$

which looks like $\displaystyle \frac{d}{dt}\mathbb{1}(t) = \delta(t)$.

Is this formally true?

$$\mathcal{L}\left\{\frac{d}{dt}\mathbb{1}(t)\right\} = s\,\overset{1/s}{\mathcal{L}\{\mathbb{1}(t)\}} - \cancel{\mathbb{1}(0^-)} \rightarrow 0$$

$$= 1 = \mathcal{L}\{\delta(t)\}$$

YES

Recap: Unit Impulse

$$\delta(t) = \lim_{\Delta \to 0} f_\Delta(t) = \begin{cases} 0 & t \neq 0 \\ \infty & t = 0 \end{cases}$$

"sifting property":

$$\int_a^b \delta(t) h(t)\, dt = \begin{cases} h(0) & \text{if } 0 \in [a,b] \\ 0 & \text{otherwise} \end{cases}$$

Laplace Transform:

$$\mathcal{L}\{\delta(t)\} = 1$$

Useful property:

$$\frac{d}{dt} \mathbb{1}(t) = \delta(t)$$

Impulse Response

The <u>impulse response</u> of a system is the output $y(t)$ when $u(t) = \delta(t)$ and all ICs on $y(t)$ are zero.

$$Y(s) = G(s)U(s) + \frac{[c(s) - b(s)]}{r(s)}$$

$\Rightarrow u(t) = \delta(t) \Rightarrow b(s) = \emptyset$ and $U(s) = 1$

$\Rightarrow$ all ICs on $y(t)$ zero $\Rightarrow c(s) = \emptyset$

So:

$$Y(s) = G(s)$$

and thus

$$y(t) = \mathcal{L}^{-1}\{G(s)\} \triangleq g(t)$$

$$\boxed{\text{The \underline{impulse response} } g(t) \text{ is the inverse transform of the transfer function } G(s)}$$

Conversely, Knowledge (or measurement) of $g(t)$ tells us what the transfer function is, and hence the governing diff'l eq'ns.

$\Longrightarrow$ Foundation of "system identification" theory.

Theoretical problems in integrals when discontinuities or singularities at one of the end points.

$$F(s) = \int_{0}^{\infty} f(t) e^{-st} dt$$

→ possible problem here

Resolve these by taking lower limit at $t = \emptyset^-$ (the instant before $t = \emptyset$).

⇒ integral "sees" effect of singularities at $t = \emptyset$.

Starting the integral at 0^- instead of 0

- Avoids singularities at endpoints
- Causes transform to "see" singularities and discontinuities at $t=0$, so their effects will be reflected in the solutions for $y(t)$.

Hence:

$$F(s) = \int_{0^-}^{\infty} f(t) e^{-st}\,dt$$

Implications:

Assumed ICs for $y(t)$: just before $t=0$

$$\mathcal{L}\{\dot{y}(t)\} = sY(s) - \boxed{y(0^-)}$$

$$\mathcal{L}\{\ddot{y}(t)\} = s^2Y(s) - \boxed{\dot{y}(0^-) - sy(0^-)}$$

etc.

$$\mathcal{L}\{\dot{u}(t)\} = sU(s) - \boxed{u(0^-)}$$

$$\mathcal{L}\{\ddot{u}(t)\} = s^2U(s) - \boxed{\dot{u}(0^-) - su(0^-)}$$

etc.

Always $= 0$ in our analysis!

$\Rightarrow b(s) = 0$ always

Thus:

$$Y(s) = \left[\frac{q(s)}{r(s)}\right] U(s) + \left[\frac{c(s) - \cancel{b(s)}^{\,\emptyset}}{r(s)}\right]$$

$$\boxed{= \left[\frac{q(s)}{r(s)}\right] U(s) + \left[\frac{c(s)}{r(s)}\right]}$$

=> IC polynomial $b(s)$ for input vanishes

=> specific to controls convention for $u(t)$

=> Not a common assumption in regular math classes.

=> In controls, want to know effect of discontinuities

Common, discontinuous "test functions"

$u(t) = \mathbb{1}(t)$ (unit step) $\mathcal{L}\{\mathbb{1}(t)\} = \frac{1}{s}$

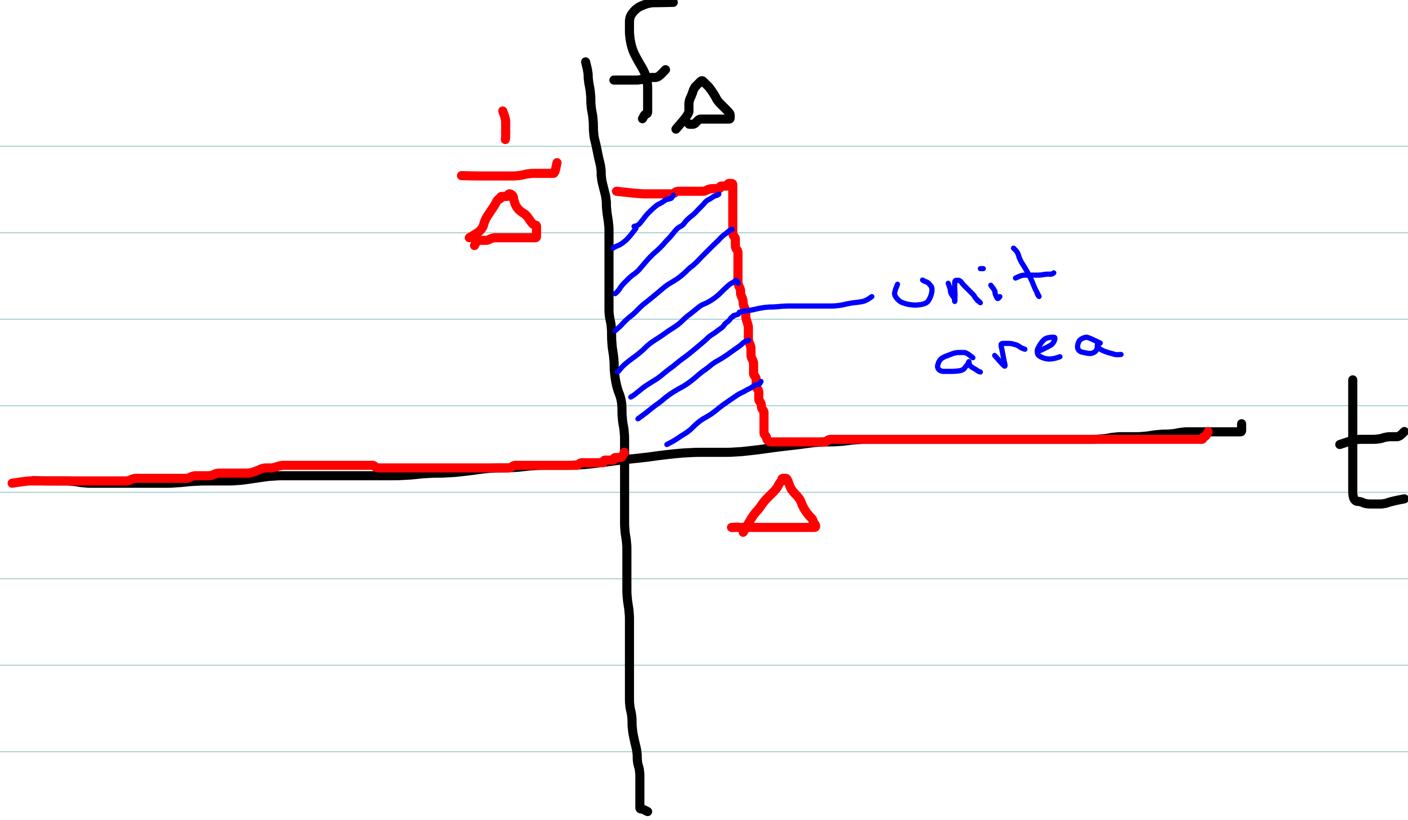
$$u(t) = \cos(\omega t)\,\mathbb{1}(t)$$

$$= \begin{cases} \cos(\omega t) & t \geq \emptyset \\ \emptyset & t < \emptyset \end{cases}$$

$$\Longrightarrow \quad u(t) = f_\Delta(t)$$

$$= \begin{cases} \frac{1}{\Delta} & \emptyset \leq t \leq \Delta \\ \emptyset & \text{otherwise} \end{cases}$$

"Unit <u>pulse</u> function"



Note: for <u>any</u> $\Delta > 0$

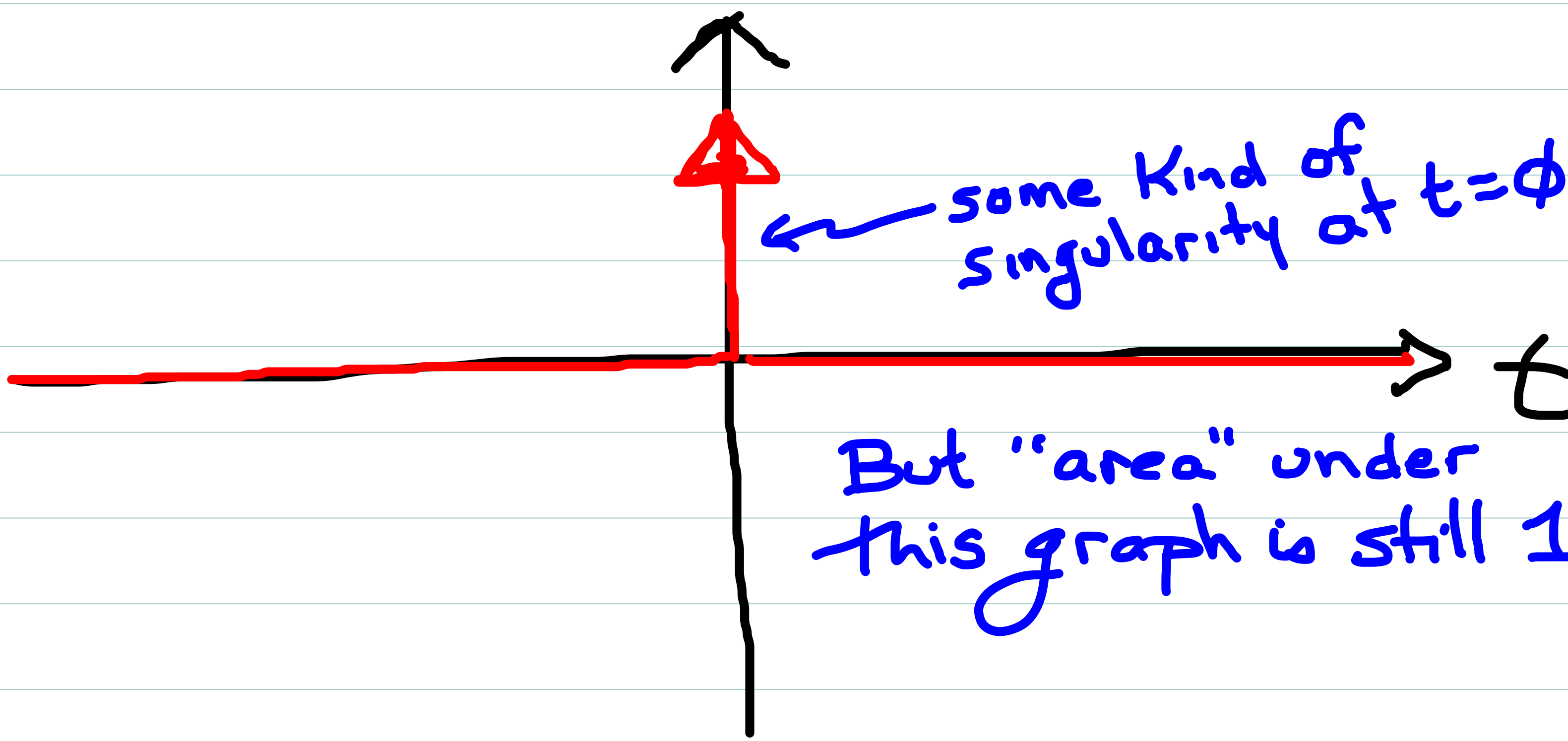
$$\int_{0^-}^{\infty} f_\Delta(t)\,dt = \int_{0^-}^{\Delta} \left(\frac{1}{\Delta}\right) dt$$

$$= 1$$

What is $\lim_{\Delta \to 0} f_\Delta(t)$?

$$= \lim_{\Delta \to 0} \begin{cases} \frac{1}{\Delta} & \phi \le t \le \Delta \\ \phi & \text{otherwise} \end{cases}$$

$$= \begin{cases} \infty & t = \phi \\ \phi & \text{otherwise} \end{cases}$$



But "area" under this graph is still 1...?

Define:

$$\boxed{\delta(t) = \lim_{\Delta \to 0} f_\Delta(t)}$$

"ideal impulse": models delivering a unit of input energy over negligibly small time.

(sharp "kick")

Alternate names:

"delta function"
"impulse function"
"Dirac delta"

Note: *Not really a meaningful function at all!*

More formally, belongs to a class of mathematical objects called "distributions" or "generalized functions"

Suppose $\delta(t)$ appears in an integral

$$\int_{-\infty}^{\infty} \delta(t)\, h(t)\, dt \quad , \quad h(t) \text{ arbitrary f'n}$$

$$= \int_{-\infty}^{\infty} \left[\lim_{\Delta \to 0} f_\Delta(t) \right] h(t)\, dt$$

$$= \lim_{\Delta \to 0} \int_{-\infty}^{\infty} f_\Delta(t)\, h(t)\, dt$$

$$= \lim_{\Delta \to 0} \left\{ \frac{1}{\Delta} \int_{0^-}^{\Delta} h(t)\, dt \right\}$$

$$\cong \left(\frac{1}{\Delta}\right) \left(\Delta\, h(\phi)\right)$$

$$= h(\phi)$$

Note: with $h(t) = 1$ for all t, we get

$$\int_{0^-}^{\infty} \delta(t)\, dt = 1$$

Defining Property of $\delta(t)$

$$\int_a^b \delta(t)h(t)\,dt = \begin{cases} h(0) & \text{if } 0 \in (a,b) \\ 0 & \text{otherwise} \end{cases}$$

"sifting property"

$\Rightarrow$ $\delta(t)$ is defined by what it does in an integral

Not as an ordinary function

Now we can compute:

$$\mathcal{L}\{\delta(t)\} = \int_{0^-}^{\infty} \delta(t)e^{-st}\,dt$$

$$= \left[e^{-st}\right]_{t=0} = 1$$

Thus:

$$\boxed{\mathcal{L}\{\delta(t)\} = 1}$$

$$\Rightarrow \mathcal{L}^{-1}\{c\} = c\,\delta(t)$$

and by Linearity:

$$\mathcal{L}\{c\,\delta(t)\} = c \quad \text{for any constant } c.$$

Now recall

$$\frac{d}{dt}\mathbb{1}(t) = \begin{cases} \infty & t=0 \\ 0 & \text{otherwise} \end{cases}$$

which looks like $\frac{d}{dt}\mathbb{1}(t) = \delta(t)$.

Is this formally true?

$$\mathcal{L}\left\{\frac{d}{dt}\mathbb{1}(t)\right\} = s\,\mathcal{L}\{\mathbb{1}(t)\} - \mathbb{1}(0^-)$$

Thus:

$$\mathcal{L}\{\delta(t)\} = 1$$

and by Linearity:

$$\mathcal{L}\{c\,\delta(t)\} = c \quad \text{for any constant } c.$$

Now recall $\displaystyle \frac{d}{dt}\mathbb{1}(t) = \begin{cases} \infty & t=0 \\ 0 & \text{otherwise} \end{cases}$

which looks like $\displaystyle \frac{d}{dt}\mathbb{1}(t) = \delta(t).$

Is this formally true?

$$\mathcal{L}\left\{\frac{d}{dt}\mathbb{1}(t)\right\} = s\,\underbrace{\mathcal{L}\{\mathbb{1}(t)\}}_{1/s} - \cancel{\mathbb{1}(0^-)} \rightarrow 0$$

$$= \underline{1} = \mathcal{L}\{\delta(t)\}$$

YES

Recap: Unit Impulse

$$\delta(t) = \lim_{\Delta \to 0} f_\Delta(t) = \begin{cases} 0 & t \neq 0 \\ \infty & t = 0 \end{cases}$$

"sifting property":

$$\int_a^b \delta(t) h(t)\, dt = \begin{cases} h(0) & \text{if } 0 \in [a,b] \\ 0 & \text{otherwise} \end{cases}$$

Laplace Transform:

$$\mathcal{L}\{\delta(t)\} = 1$$

Useful property:

$$\frac{d}{dt} \mathbb{1}(t) = \delta(t)$$

Impulse Response

The *impulse response* of a system is the output $y(t)$ when $u(t) = \delta(t)$ and all ICs on $y(t)$ are zero.

$$Y(s) = G(s)U(s) + \frac{[c(s) - b(s)]}{r(s)}$$

$\Rightarrow u(t) = \delta(t) \Rightarrow b(s) = \emptyset$ and $U(s) = 1$

$\Rightarrow$ all ICs on $y(t)$ zero $\Rightarrow c(s) = \emptyset$

So:

$$Y(s) = G(s)$$

and thus

$$y(t) = \mathcal{L}^{-1}\{G(s)\} \triangleq g(t)$$

$\boxed{\text{The } \underline{\text{impulse response}}\ g(t) \text{ is the inverse transform of the transfer function } G(s)}$

Conversely, knowledge (or <u>measurement</u>) of $g(t)$ tells us what the transfer function is, and hence the governing diff'l eq'ns.

$\Rightarrow$ Foundation of "system identification" theory.

Additional Laplace Property

For any two functions $f_1(t)$, $f_2(t)$ with transforms $F_1(s)$, $F_2(s)$

$$\mathcal{L}^{-1}\{F_1(s)F_2(s)\} = \int_{0^-}^{\infty} f_1(t-\tau)f_2(\tau)\,d\tau$$

"convolution"

Implication: $\mathcal{L}^{-1}\{G(s)U(s)\} = \int_{0^-}^{\infty} g(t-\tau)u(\tau)\,d\tau$

proving generally what we showed specifically for the hovercraft problem.

There we had $\ddot{y}(t) = K u(t)$

$$\Rightarrow G(s) = \frac{K}{s^2} \Rightarrow g(t) = Kt \qquad \mathcal{L}^{-1}\left\{\frac{1}{s^2}\right\} = t$$

and thus $g(t-\tau) = K(t-\tau)$.

Note:

Laplace actually let's us "divide out" the effect of <u>any</u> Known input to recover the transfer function (impulse response)

$$Y(s) = G(s)U(s) \qquad (\text{assuming } \emptyset \text{ ICs})$$

$$\left[Y(s) = G(s)U(s)\right] \times \left(\frac{1}{U(s)}\right)$$

$$\left[\frac{Y(s)}{U(s)}\right] = G(s)\left[\frac{U(s)}{U(s)}\right]$$

$$= G(s) \cdot 1$$

response to ideal impulse.

Structure of Impulse Response

$$g(t) = \mathcal{L}^{-1}\{G(s)\} = \mathcal{L}^{-1}\left\{\frac{q(s)}{r(s)}\right\}$$

$$= \mathcal{L}^{-1}\left\{\sum_{k=1}^{n} \frac{\gamma_k}{(s-p_k)}\right\}$$ p_k poles of $G(s)$

or

$$\boxed{\begin{aligned} g(t) &= \sum_{k=1}^{n} \gamma_k e^{p_k t} \\ \gamma_k &= \left[(s-p_k)G(s)\right]_{s=p_k} \end{aligned}}$$

(assuming non-repeated modes for simplicity)

$$g(t) = \sum_{K=1}^{n} \gamma_K e^{P_K t}$$

<u>Note:</u>

$\Rightarrow$ $g(t)$ is a <u>specific</u> linear combination of the modes.

$\Rightarrow$ Like a special homogeneous response

Alternate characterization of system stability

$$\boxed{\lim_{t \to \infty} |g(t)| \to \emptyset}$$

(if system is stable)

Impulse response in state-space

For a state-space model recall: $G(s) = C(sI-A)^{-1}B + D$

Assume $D=0$ for simplicity (most common case)

then $$g(t) = \mathcal{L}^{-1}\{C(sI-A)^{-1}B\}$$
$$= C\,\mathcal{L}^{-1}\{(sI-A)^{-1}\}B \quad \text{(linearity)}$$

Let $\Phi(s) \triangleq (sI-A)^{-1}$ ($n \times n$ matrix)

and $\phi(t) \triangleq \mathcal{L}^{-1}\{\Phi(s)\}$ ($n \times n$ matrix)

Then
$$g(t) = C\phi(t)B$$

$\Rightarrow$ what is this matrix $\phi(t)$??

Matrix Exponential Function

Note for scalar a,

$$\mathcal{L}^{-1}\left\{\frac{1}{s-a}\right\} = \mathcal{L}\{(s-a)^{-1}\} = e^{at}$$

By analogy

$$\mathcal{L}^{-1}\{\Phi(s)\} = \mathcal{L}^{-1}\{(s\mathbb{I}-A)^{-1}\} \triangleq e^{At}$$

The "matrix exponential function"

How to calculate it?

Laplace (and its inverse) works on matrices just like it does on vectors:

Apply inverse transform to each entry of $\Phi(s) = (s\mathbb{I}-A)^{-1}$ (so n^2 inverse transforms ugh!)

General Observations

Recall that $\Phi(s) = (sI - A)^{-1} = \dfrac{Q(s)}{r(s)}$

$Q(s) = \text{Adj}(sI - A)$ nxn matrix of polynomials in s

$r(s) = \text{Det}(sI - A)$ ordinary polynomial in s

Hence:

1.) Each entry of $\Phi(s)$ is rational
$\Rightarrow$ can use PFE tricks for inverse xform

2.) Each entry of $\Phi(s)$ has same denom, hence same poles (roots of $r(s)$)

Recall: roots of $r(s)$ same as eigenvalues of A

Thus: Each entry of $e^{At} = \mathcal{L}^{-1}\{(sI - A)^{-1}\}$ will be a linear combination of $e^{\lambda_k t}$ where λ_k are eigenvalues of A

Example

$A = \begin{bmatrix} 0 & 1 \\ -6 & -5 \end{bmatrix}$ => companion form. By inspection

$r(s) = s^2 + 5s + 6 = (s+2)(s+3)$

=> each entry of e^{At} is a linear comb. of e^{-2t}, e^{-3t}

$(sI - A) = \begin{bmatrix} s & -1 \\ 6 & s+5 \end{bmatrix}$ $Q(s) = \mathrm{Adj}(sI - A) = \begin{bmatrix} s+5 & 1 \\ -6 & s \end{bmatrix}$

$$(sI - A)^{-1} = \begin{bmatrix} \dfrac{s+5}{(s+2)(s+3)} & \dfrac{1}{(s+2)(s+3)} \\[2ex] \dfrac{-6}{(s+2)(s+3)} & \dfrac{s}{(s+2)(s+3)} \end{bmatrix}$$

inverse xform each entry separately using PFE

Example, cont

(1,1) $\dfrac{s+5}{(s+2)(s+3)} = \dfrac{A_1}{s+2} + \dfrac{A_2}{s+3}$ $\quad A_1 = \dfrac{-2+5}{1} = 3$

$A_2 = \dfrac{-3+5}{-1} = -2$

(1,2) $\dfrac{1}{(s+2)(s+3)} = \dfrac{A_1}{s+2} + \dfrac{A_2}{s+3}$ $\quad A_1 = \dfrac{1}{1} = 1$

$A_2 = \dfrac{1}{-1} = -1$

(2,1) $\dfrac{-6}{(s+2)(s+3)}$ is just $-6 \times \left(\dfrac{1}{(s+2)(s+3)}\right)$

(2,2) $\dfrac{s}{(s+2)(s+3)} = \dfrac{A_1}{s+2} + \dfrac{A_2}{s+3}$ $\quad A_1 = \dfrac{-2}{1} = -2$

$A_2 = \dfrac{-3}{-1} = 3$

Thus here:

$$e^{At} = \mathcal{L}^{-1}\{(sI-A)^{-1}\} = \begin{bmatrix} 3e^{-2t}-2e^{-3t} & e^{-2t}-e^{-3t} \\ 6(e^{-3t}-e^{-2t}) & 3e^{-3t}-2e^{-2t} \end{bmatrix}$$

Recap

So with a state-space model we equivalently have

$$g(t) = Ce^{At}B$$

where:

- e^{At} is an $n \times n$ matrix
- each entry of e^{At} is a linear combination of modes
- modes determined by poles (p_k) $\iff$ eigenvalues (λ_k) of A

and thus:

$$g(t) = Ce^{At}B = \sum_{k=1}^{n} \gamma_k e^{p_k t}$$

a linear combination of modes, just like before.

Step Responses

The (unit) step response of a system is the output $y(t)$ when $u(t)=\mathbb{1}(t)$ and all ICs on $y(t)$ are zero.

$$Y(s)=G(s)U(s)+\frac{[\cancelto{\emptyset}{c(s)}-\cancelto{\emptyset}{b(s)}]}{r(s)}$$

$U(s)=\frac{1}{s}$ here, so

$$Y(s)=\left(\frac{1}{s}\right)G(s)=\frac{q(s)}{s\,r(s)}$$

Intermediate Case 3 Situations

If $1.1 < \dfrac{|P_2|}{|P_1|} < 5$ (or 8 or 10)

$$\frac{4}{|P_1|} < t_s < \frac{6}{|P_1|}$$

Unfortunately, there is no simple formula for interpolating between the two limits based on the exact ratio.

Step Responses

The (unit) step response of a system is the output $y(t)$ when $u(t) = \mathbb{1}(t)$ and all ICs on $y(t)$ are zero.

$$Y(s) = G(s)U(s) + \frac{[\cancel{c}(s) - \cancel{b}(s)]}{r(s)}$$

with $\cancel{c}(s) \to \emptyset$ and $\cancel{b}(s) \to \emptyset$

$U(s) = \frac{1}{s}$ here, so

$$Y(s) = \left(\frac{1}{s}\right) G(s) = \frac{q(s)}{s\, r(s)}$$

General Thoughts about step responses

① Every system has a unit step response:

$$Y(s) = \left[\left(\frac{1}{s}\right) G(s)\right]$$

$$y(t) = \mathcal{L}^{-1}\left\{\frac{1}{s} G(s)\right\} \triangleq y_{us}(t)$$

Find $y_{us}(t)$ as usual by partial fraction expansion and inverse transform of each term

However, we want to be able to predict main features of $y_{us}(t)$ by inspection for 1^{st} and 2^{nd} order systems

⇒ Very common special cases

⇒ "Building blocks" for more complex systems

(2.) (Use of linearity, I)

$$u(t) = c\,\mathbb{1}(t) \implies y(t) = c\,y_{us}(t)$$

All y(t) values are the unit step values multiplied by c.

Equivalent to "rescaling" vertical axis on plot of y(t), however horizontal (time) axis is unaffected

$\Rightarrow$ Characteristic times (t_s, t_r, t_p) are unaffected

We'll encounter these shortly.

$\Rightarrow$ Corresponding y(t) values scaled by c:

$$y_{ss} = c\,G(0), \quad y_p = c\,G(0)[1 + M_p]$$

$\Rightarrow$ True for any c, positive or negative

(3.) (Use of Linearity, II)

By definition, unit step response assumes all ICs are zero.

However, can easily "Add on" effects of nonzero ICs.

Nonzero now

$$Y(s) = \left[\frac{1}{s}G(s)\right] + \left[\frac{C(s)}{r(s)}\right]$$

$$y(t) = \mathcal{L}^{-1}\{Y(s)\} = \mathcal{L}^{-1}\left\{\left(\frac{1}{s}\right)G(s)\right\} + \mathcal{L}^{-1}\left\{\frac{C(s)}{r(s)}\right\}$$

$$= y_{us}(t) + \mathcal{L}^{-1}\left\{\frac{C(s)}{r(s)}\right\}$$

Added terms from ICs

Solve for last term by PFE

Effect of added terms on t_s, t_p, y_p etc depends on specific ICs. No simple formulae to quantify their effects.

"1^{st} Order" Responses

$$\dot{y}(t) + \alpha_0 y(t) = \beta_0 u(t) \implies G(s) = \frac{\beta_0}{s+\alpha_0}$$

Single real pole at $p_1 = -\alpha_0$ (stable if $\alpha_0 > 0$)

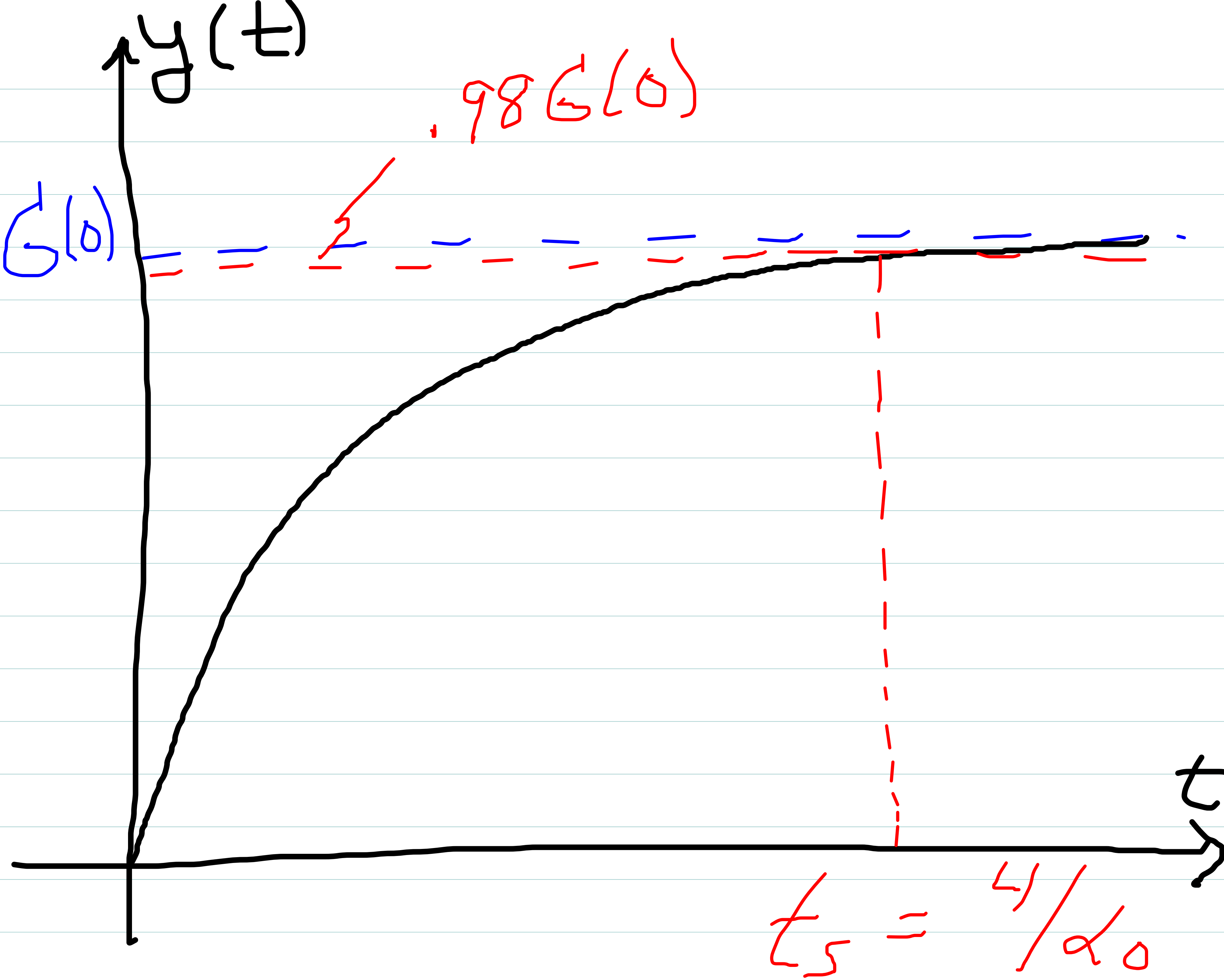
$$Y(s) = \frac{\beta_0}{s(s+\alpha_0)} = \frac{A_1}{s} + \frac{A_2}{s+\alpha_0}$$

$$A_1 = [sY(s)]_{s=0} = \frac{\beta_0}{\alpha_0} = G(0)$$

$$A_2 = [(s+\alpha_0)Y(s)]_{s=-\alpha_0} = \frac{-\beta_0}{\alpha_0} = -G(0)$$

Thus:

$$\boxed{y(t) = G(0)\left[1 - e^{-\alpha_0 t}\right]}$$



y(t)
.98G(0)
G(0)
t
$t_s = 4/\alpha_0$

Notes

1. Response asymptotically approaches steady-state

$$y_{ss}(t) = G(0)$$ (as expected)

2. Response never crosses its steady-state

3. Response settles within 2% of its steady-state in

$$t_s = \frac{4}{|Re\{p\}|} = \frac{4}{\alpha_0}$$

4. "Shape" of graph is same for any 1st order system

Responses only differ by:

- steady-state level, $G(0)$
- settling time, t_s

"2^{nd} Order" Step Responses

$$\ddot{y}(t) + \alpha_1 \dot{y}(t) + \alpha_0 y(t) = \beta_0 u(t) \Rightarrow G(s) = \frac{\beta_0}{s^2 + \alpha_1 s + \alpha_0}$$

2 poles, both stable if $\alpha_1 > 0, \alpha_0 > 0$.

3 possibilities for poles:

1. $\alpha_1^2 < 4\alpha_0 \Rightarrow p_1, p_2$ complex conjugates
2. $\alpha_1^2 = 4\alpha_0 \Rightarrow p_1 = p_2$ repeated real
3. $\alpha_1^2 > 4\alpha_0 \Rightarrow p_1, p_2$ real, non-repeated

Case 1. is most interesting (and complicated) tackle this after the other two

2nd order response, Case 2

$$G(s) = \frac{\beta_0}{s^2+\alpha_1 s+\alpha_0} \qquad \alpha_1^2 = 4\alpha_0 \quad (\zeta = 1)$$

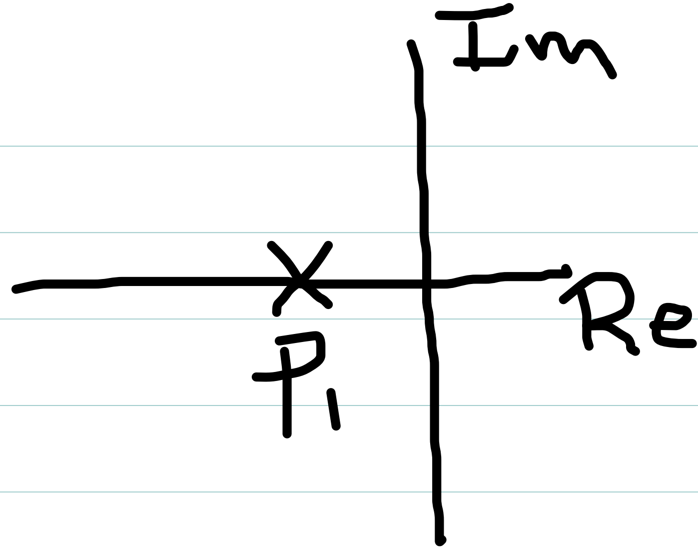
$$= \frac{\beta_0}{(s-p_1)^2} \qquad \text{repeated real pole}$$

$$Y(s) = \left(\frac{1}{s}\right) G(s) = \frac{A_1}{s} + \frac{A_2}{(s-p_1)} + \frac{A_3}{(s-p_1)^2}$$

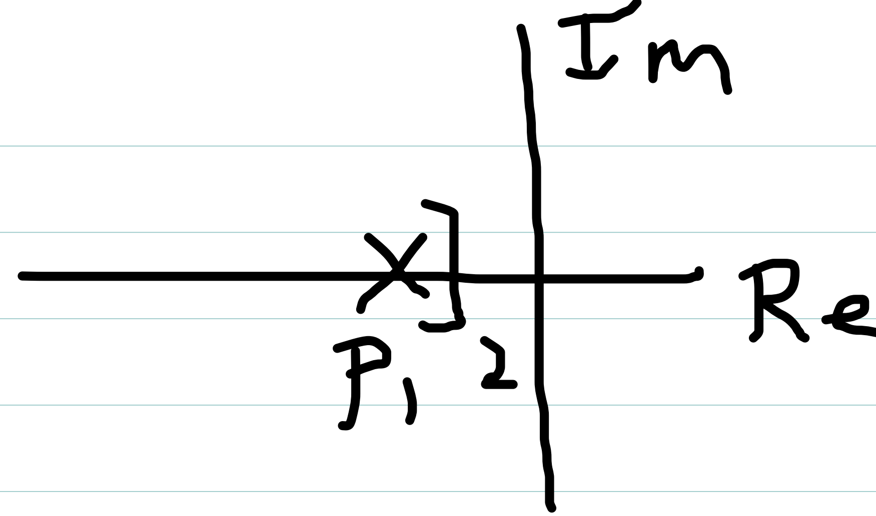
$$y(t) = G(\phi) + [A_2 + A_3 t]\, e^{p_1 t}$$

Non-oscillatory, since poles are real

features resemble 1st order response
(No overshoot, $y_{ss} = G(0)$ approached asymptotically from below), but $\boxed{t_s \text{ 50\% longer } \left(\frac{6}{|p_1|}\right)}$

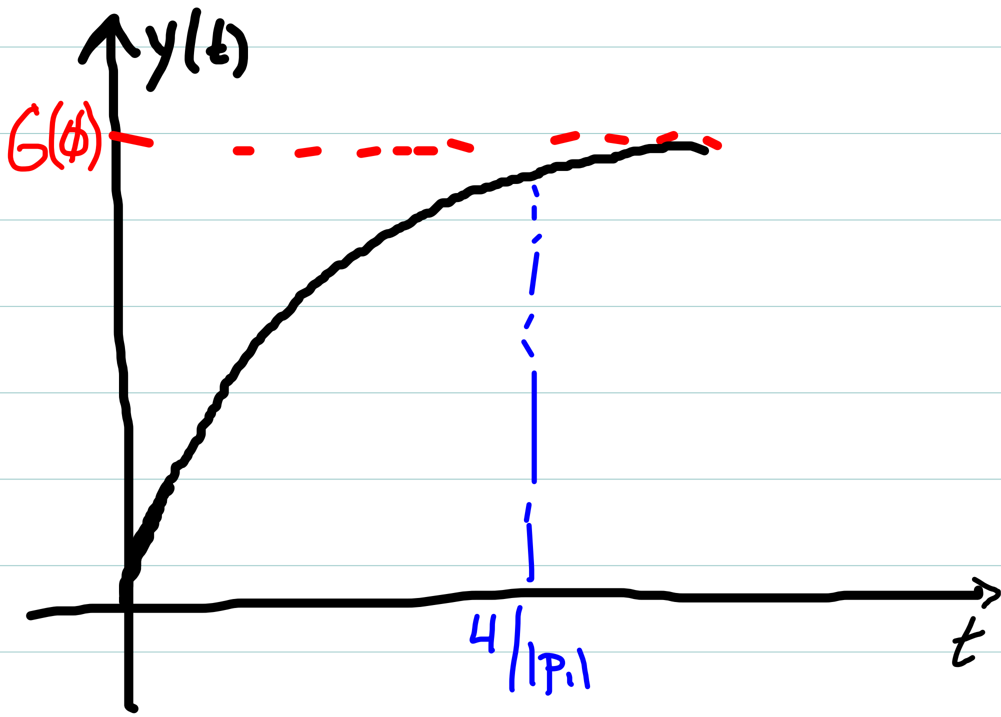


1^{st} order



2^{nd} order, repeated real

Add'l $te^{p_1 t}$ term
"slows down" response.



2^{nd} order Response, Case 3

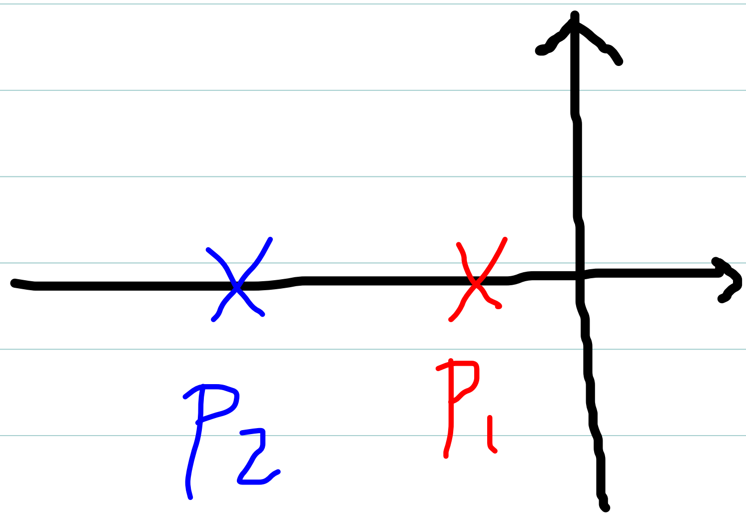
$\alpha_1^2 > 4\alpha_0$

$$Y(s) = \frac{\beta_0}{s(s-p_1)(s-p_2)} \qquad p_1 \neq p_2 .$$

$$\Rightarrow y(t) = G(0) + A_1 e^{p_1 t} + A_2 e^{p_2 t}$$

Assume for notation sake that poles are numbered so that

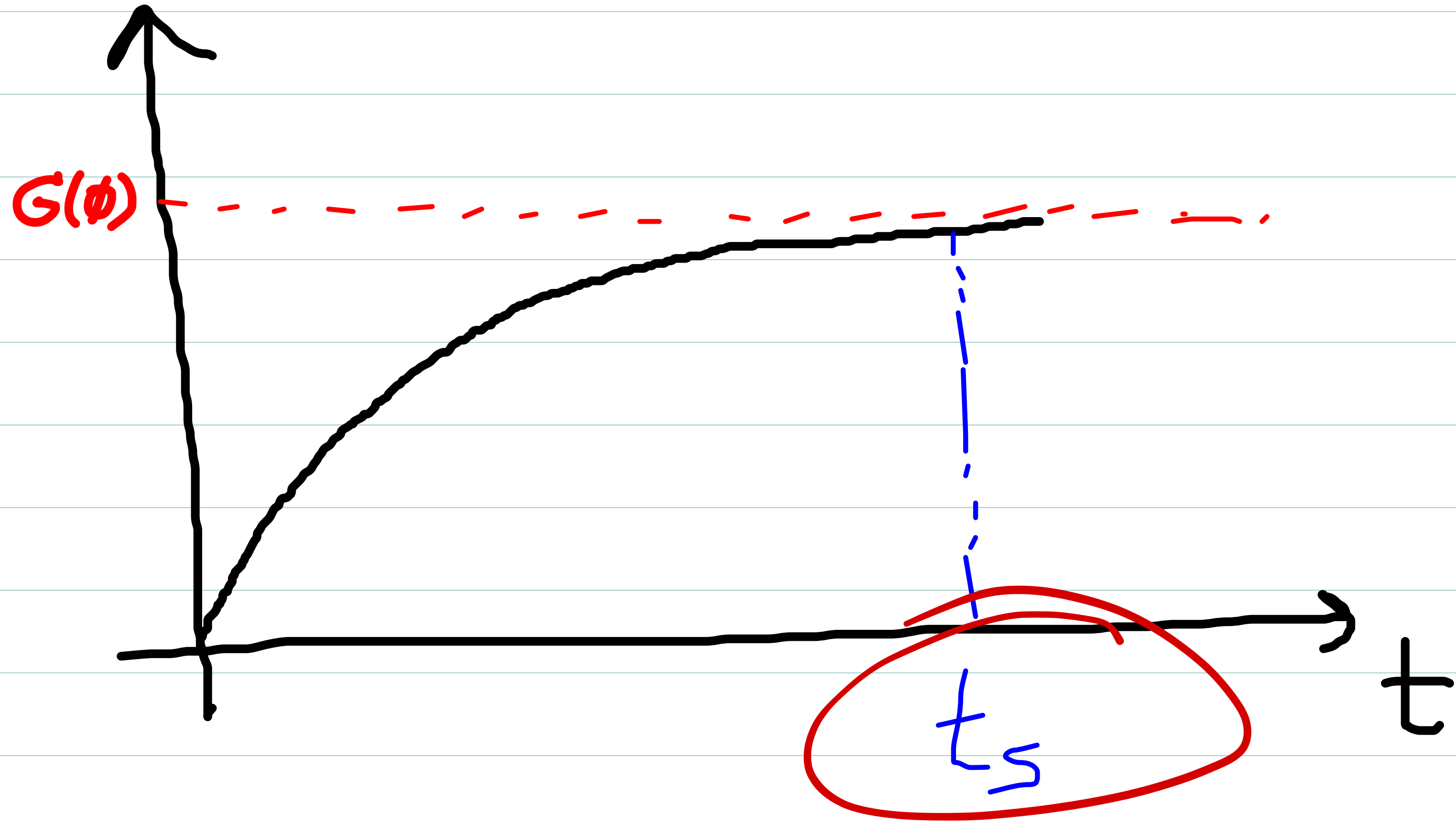
$$p_2 < p_1 \qquad (\Rightarrow |p_2| > |p_1| \text{ since } p_1, p_2 \text{ assumed negative})$$



p_1 is the "slow pole"

p_2 is the "fast pole"

General sol'n again resembles 1^{st} order response



t_s difficult to quantify precisely for arbitrary p_1, p_2

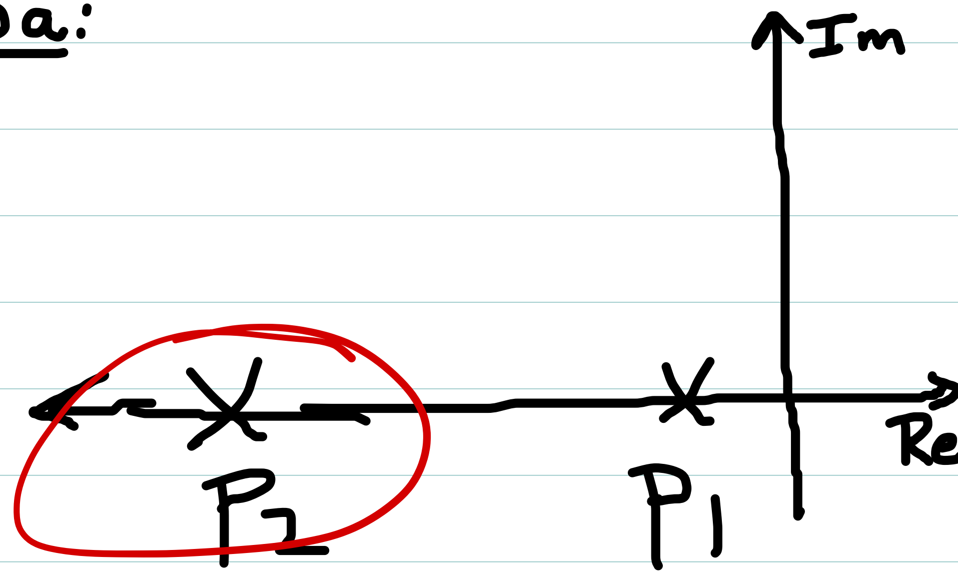
Two Limiting cases:

<u>Case 3a:</u> $|p_2| \gg |p_1|$

<u>Case 3b:</u> $|p_2| \approx |p_1|$

Case 3a:

$$y(t) = G(\phi) + A_1 e^{p_1 t} + A_2 e^{p_2 t}$$



$|p_2| >> |p_1|$

$\Rightarrow p_2$ much further into LHP than p_1

$\Rightarrow e^{p_2 t} \to \phi$ much faster than $e^{p_1 t}$

$\Rightarrow e^{p_1 t}$ controls settling time ("slow pole")

So $t_s \approx \dfrac{4}{|p_1|}$ in this case

$\Rightarrow$ Corresponds with previous "1st cut" of approximating system settling time with settling time of slowest mode.

Dominant Modes

When $|p_2| \gg |p_1|$ we say that mode $e^{p_1 t}$ "dominates" transient response, or that $e^{p_1 t}$ ("slow mode") is the

Dominant mode

What is a sufficient separation for a mode to be dominant

Generally, if $\boxed{|p_2| > 5|p_1| \quad \text{or} \quad |p_2| > 10|p_1|}$

i.e. if p_2 is 5-10 times further into LHP

$\Rightarrow$ setting time of $e^{p_2 t}$ 5-10 times faster than that of $e^{p_1 t}$

(5 is usually sufficient. Some authors use 8 or even 10)

Case 3b

$|p_2| \approx |p_1| \Rightarrow p_2 \approx p_1$ Poles are "nearly" repeated

Here it is best to approximate the settling time as though the poles were actually repeated

$$t_s \approx \frac{6}{|p_1|}$$

Simple rule of thumb for this:

$$1 \leq \frac{|p_2|}{|p_1|} \leq 1.1$$

Intermediate Case 3 Situations

If $1.1 < \frac{|P_2|}{|P_1|} < 5$ (or 8 or 10)

$$\frac{4}{|P_1|} < t_s < \frac{6}{|P_1|}$$

Unfortunately, there is no simple formula for interpolating between the two limits based on the exact ratio.

"2^{nd} Order" Step Responses

$$\ddot{y}(t)+\alpha_1\dot{y}(t)+\alpha_0 y(t)=\beta_0 u(t) \implies G(s)=\frac{\beta_0}{s^2+\alpha_1 s+\alpha_0}$$

2 poles, both stable if $\alpha_1>0$, $\alpha_0>0$.

3 possibilities for poles:

1. $\alpha_1^2<4\alpha_0 \implies p_1, p_2$ complex conjugates
2. $\alpha_1^2=4\alpha_0 \implies p_1=p_2$ repeated real
3. $\alpha_1^2>4\alpha_0 \implies p_1, p_2$ real, non-repeated

Case 1. is most interesting (and complicated) tackle this after the other two

Useful Observation (Case 1)

$p_1 = \sigma + j\omega_d \qquad \omega_d = \mathrm{Im}\{p_1\}$

Note slight change of notation! $\omega \to \omega_d$

$$s^2 + \alpha_1 s + \alpha_0 = (s - p_1)(s - \overline{p}_1)$$

$$= s^2 - (p_1 + \overline{p}_1)s + p_1\overline{p}_1$$

$$= s^2 - 2\,\mathrm{Re}\{p_1\}s + |p_1|^2$$

$$= s^2 - 2\sigma s + (\sigma^2 + \omega_d^2)$$

Hence:

$$\boxed{\begin{aligned} \alpha_1 &= -2\sigma = -2\,\mathrm{Re}\{p_1\} \\ \alpha_0 &= \sigma^2 + \omega_d^2 = |p_1|^2 \end{aligned}}$$

Rapidly identify pole location from coefs.

2nd Order Response, Case 1:

$$Y(s) = \frac{\beta_0}{s(s-p_1)(s-\bar{p}_1)} = \frac{A_1}{s} + \frac{A_2}{(s-p_1)} + \frac{\bar{A}_2}{(s-\bar{p}_1)}$$

$$A_1 = [sY(s)]_{s=0} = \frac{\beta_0}{p_1\bar{p}_1} = \frac{\beta_0}{\alpha_0} = G(0)$$

$$A_2 = [(s-p_1)Y(s)]_{s=p_1} = \frac{\beta_0}{p_1(p_1-\bar{p}_1)} = \frac{\beta_0}{(\sigma+j\omega_d)(2j\omega_d)}$$

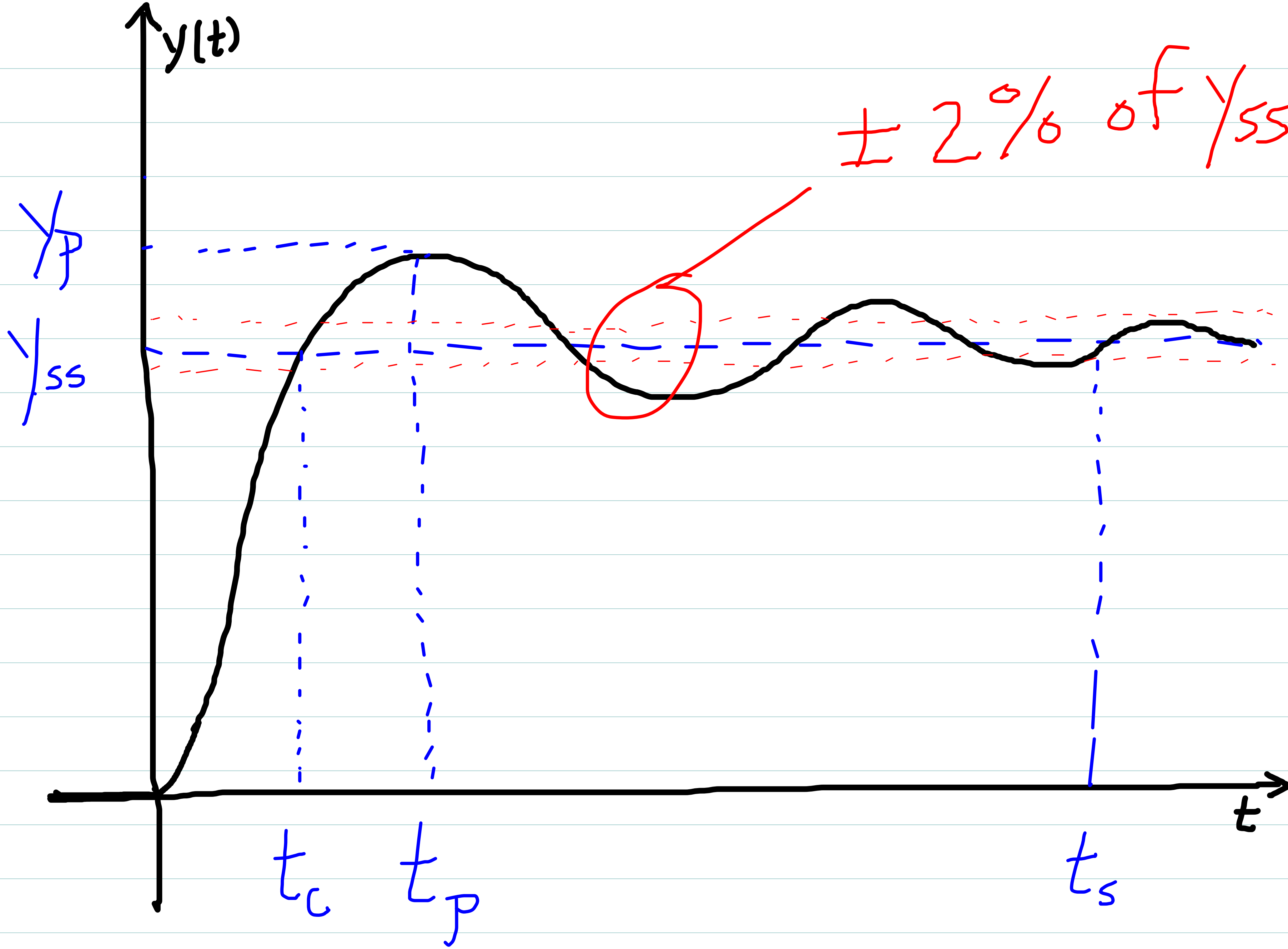
$$= \underbrace{\left(\frac{\beta_0}{2\alpha_0}\right)}_{\frac{1}{2}G(0)} \underbrace{\left(\frac{\alpha_0}{(\sigma+j\omega_d)(j\omega_d)}\right)}_{B}$$

SO:

$$y(t) = G(0) + 2|A_2|e^{\sigma t}\cos(\omega_d t + \measuredangle A_2)$$

OR:

$$\boxed{y(t) = G(0)\left[1 + |B|e^{\sigma t}\cos(\omega_d t + \measuredangle B)\right]}$$



y(t)
± 2% of y_ss
y_p
y_ss
t
t_c
t_p
t_s

General Observations

1. $y(t)$ continually oscillates about its steady-state value $y_{ss} = G(\phi)$
2. t_c = time steady-state is first crossed
3. 1st oscillation is largest, and creates an initial overshoot past the steady-state.
4. This initial overshoot has peak value y_p, and occurs at time t_p
5. Settling time t_s defined where response enters ±2% tolerance band and remains within it for times thereafter

Must learn to rapidly quantify these!!

$$y(t) = G(0)\left[1 + |B| e^{\sigma t} \cos(\omega_d t + \measuredangle B)\right]$$

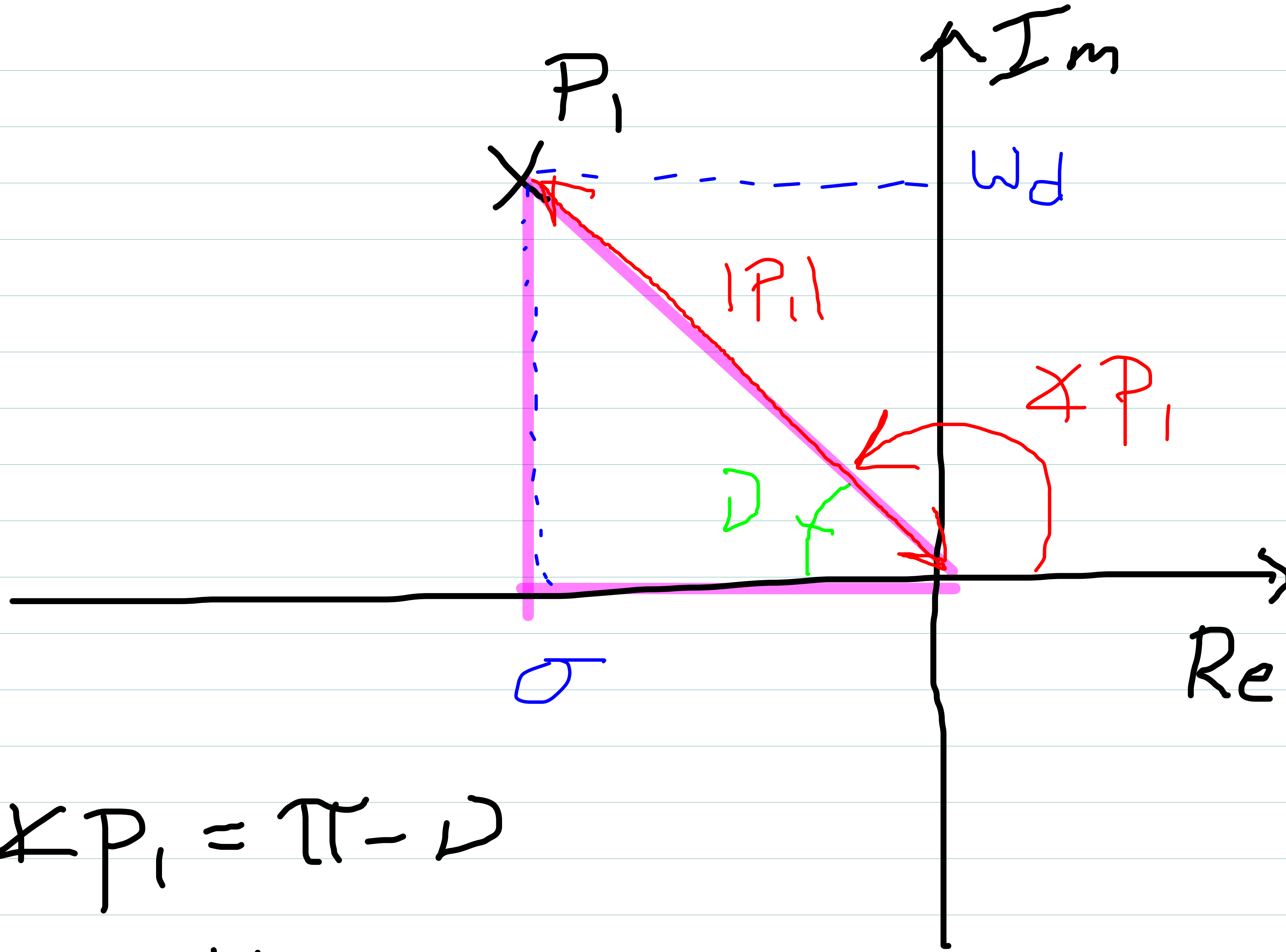
where: $$B = \frac{\alpha_0}{(j\omega_d)(\sigma + j\omega_d)} = \frac{|p_1|^2}{(j\omega_d)\, p_1}$$

$\Rightarrow$ Transient features completely determined by location of pole $p_1 = \sigma + j\omega_d$ in complex plane

$$|B| = \frac{|p_1|^2}{|j\omega_d| \cdot |p_1|} = \frac{|p_1|}{\omega_d}$$

$$\measuredangle B = \cancelto{0}{\measuredangle |p_1|^2} - \left(\measuredangle (j\omega_d) + \measuredangle p_1\right)$$

$$= -\left(\frac{\pi}{2} + \measuredangle p_1\right)$$ must quantify this!



$$\measuredangle P_1 = \pi - \nu$$

<u>Note</u>: $\nu > \not\alpha$ is supplement of $\measuredangle P_1$

So:

$$\measuredangle B = -\left(\frac{\pi}{2} + \measuredangle p_1\right) = -\left(\frac{\pi}{2} + (\pi - \nu)\right)$$

$$= -\frac{3\pi}{2} + \nu$$

and thus:

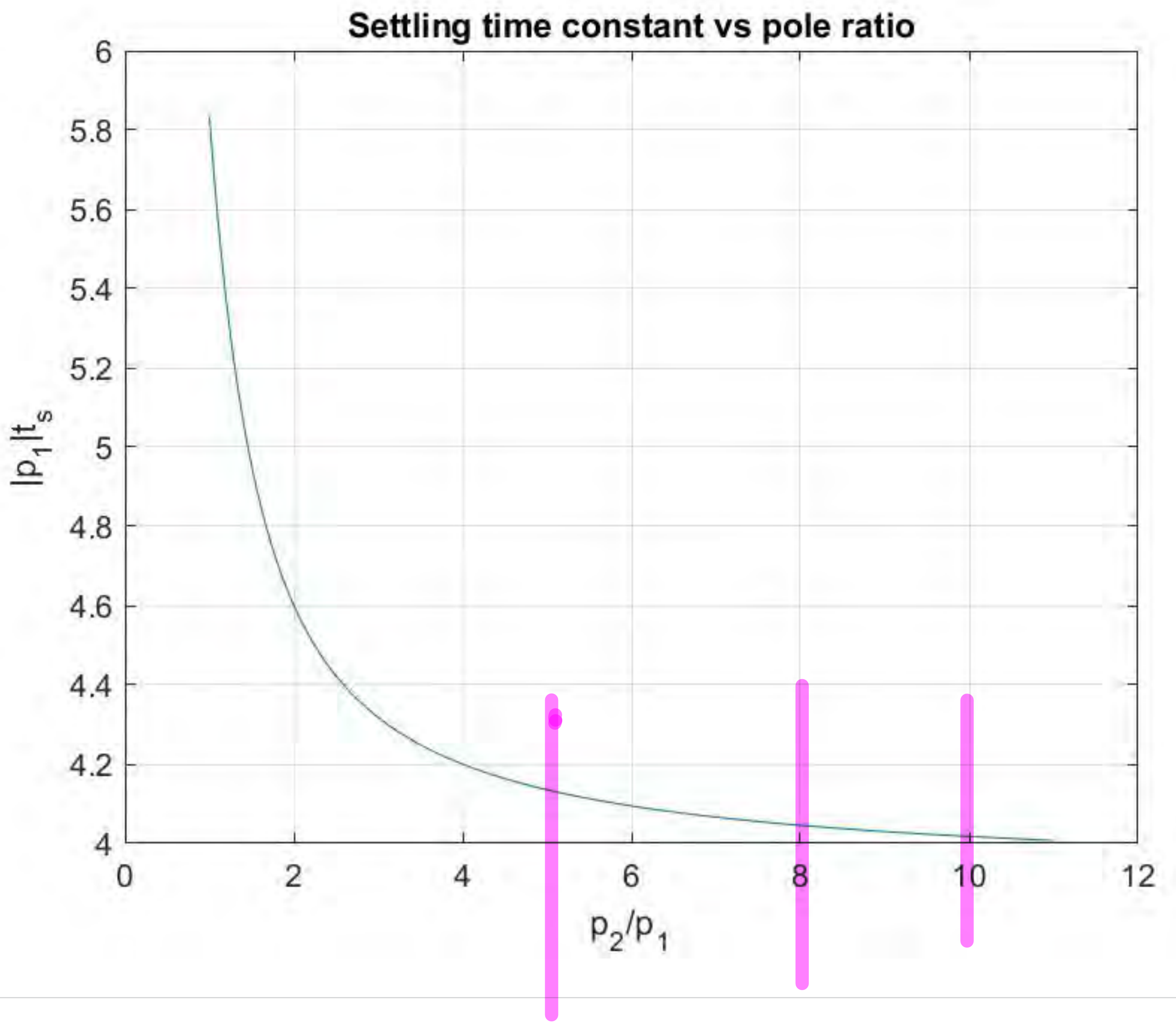
$$y(t) = G(0)\left[1 + \left(\frac{|p_1|}{\omega_d}\right) e^{\sigma t} \cos\left(\omega_d t - \frac{3\pi}{2} + \nu\right)\right]$$

so:

$$\boxed{y(t) = G(0)\left[1 - \left(\frac{|p_1|}{\omega_d}\right) e^{\sigma t} \sin(\omega_d t + \nu)\right]}$$

Need to understand how ν depends on p_1

2 real poles, $C\left(\frac{p_2}{p_1}\right)$ $\quad t_s = \frac{C(p_2/p_1)}{|p_1|}$



"2nd Order" Step Responses

$$\ddot{y}(t) + \alpha_1 \dot{y}(t) + \alpha_0 y(t) = \beta_0 u(t) \implies G(s) = \frac{\beta_0}{s^2 + \alpha_1 s + \alpha_0}$$

2 poles, both stable if $\alpha_1 > 0, \alpha_0 > 0$.

3 possibilities for poles:

1. $\alpha_1^2 < 4\alpha_0 \implies p_1, p_2$ complex conjugates
2. $\alpha_1^2 = 4\alpha_0 \implies p_1 = p_2$ repeated real
3. $\alpha_1^2 > 4\alpha_0 \implies p_1, p_2$ real, non-repeated

Case 1. is most interesting (and complicated)
tackle this after the other two

Useful Observation (Case 1)

$p_1 = \sigma + j\omega_d \qquad \omega_d = \mathrm{Im}\{p_1\}$

Note slight change of notation! $\omega \rightarrow \omega_d$

$$s^2 + \alpha_1 s + \alpha_0 = (s - p_1)(s - \bar{p}_1)$$

$$= s^2 - (p_1 + \bar{p}_1)s + p_1\bar{p}_1$$

$$= s^2 - 2\,\mathrm{Re}\{p_1\}s + |p_1|^2$$

$$= s^2 - 2\sigma s + (\sigma^2 + \omega_d^2)$$

Hence:

$$\boxed{\begin{aligned} \alpha_1 &= -2\sigma = -2\,\mathrm{Re}\{p_1\} \\ \alpha_0 &= \sigma^2 + \omega_d^2 = |p_1|^2 \end{aligned}}$$

Rapidly identify pole location from coefs.

2^{nd} Order Response, Case 1:

$$Y(s) = \frac{\beta_0}{s(s-p_1)(s-\bar{p}_1)} = \frac{A_1}{s} + \frac{A_2}{(s-p_1)} + \frac{\bar{A}_2}{(s-\bar{p}_1)}$$

$$A_1 = [sY(s)]_{s=0} = \frac{\beta_0}{p_1\bar{p}_1} = \frac{\beta_0}{\alpha_0} = G(0)$$

$$A_2 = [(s-p_1)Y(s)]_{s=p_1} = \frac{\beta_0}{p_1(p_1-\bar{p}_1)} = \frac{\beta_0}{(\sigma+j\omega_d)(2j\omega_d)}$$

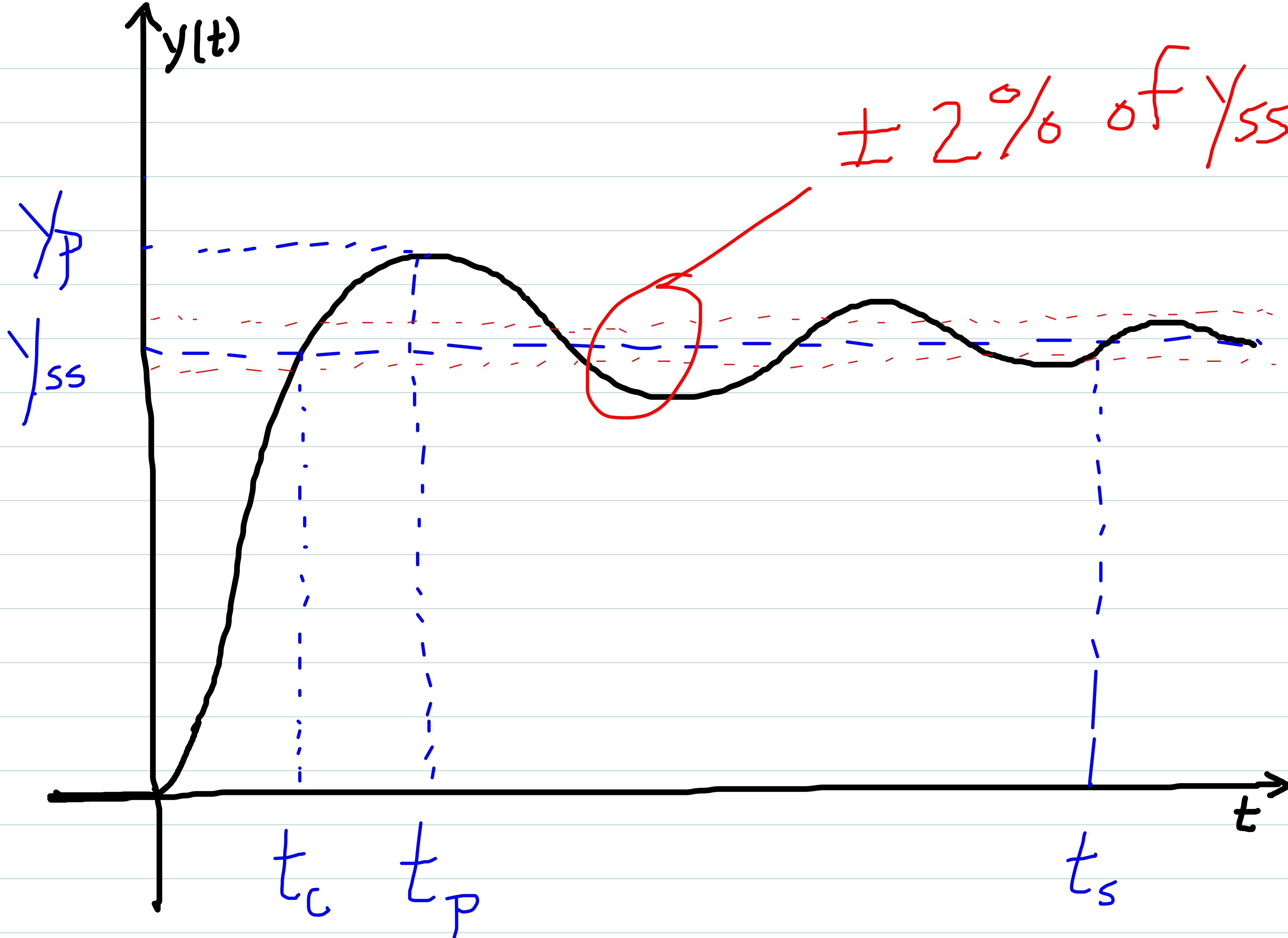
$$= \underbrace{\left(\frac{\beta_0}{2\alpha_0}\right)}_{\frac{1}{2}G(0)} \underbrace{\left(\frac{\alpha_0}{(\sigma+j\omega_d)(j\omega_d)}\right)}_{B}$$

SO:

$$y(t) = G(0) + 2|A_2|e^{\sigma t}\cos(\omega_d t + \measuredangle A_2)$$

OR:

$$\boxed{y(t) = G(0)\left[1 + |B|e^{\sigma t}\cos(\omega_d t + \measuredangle B)\right]}$$



y(t)
y_p
y_{ss}
± 2% of y_{ss}
t_c
t_p
t_s
t

General Observations

① $y(t)$ continually oscillates about its steady-state value

$y_{ss} = G(\phi)$

② t_c = time steady-state is first crossed

③ 1^{st} oscillation is largest, and creates an initial overshoot past the steady-state.

④ This initial overshoot has peak value y_p, and occurs at time t_p

⑤ Settling time t_s defined where response enters ±2% tolerance band and remains within it for times thereafter

Must learn to rapidly quantify these!!

$$y(t) = G(0)\left[1 + |B| e^{\sigma t} \cos(\omega_d t + \measuredangle B)\right]$$

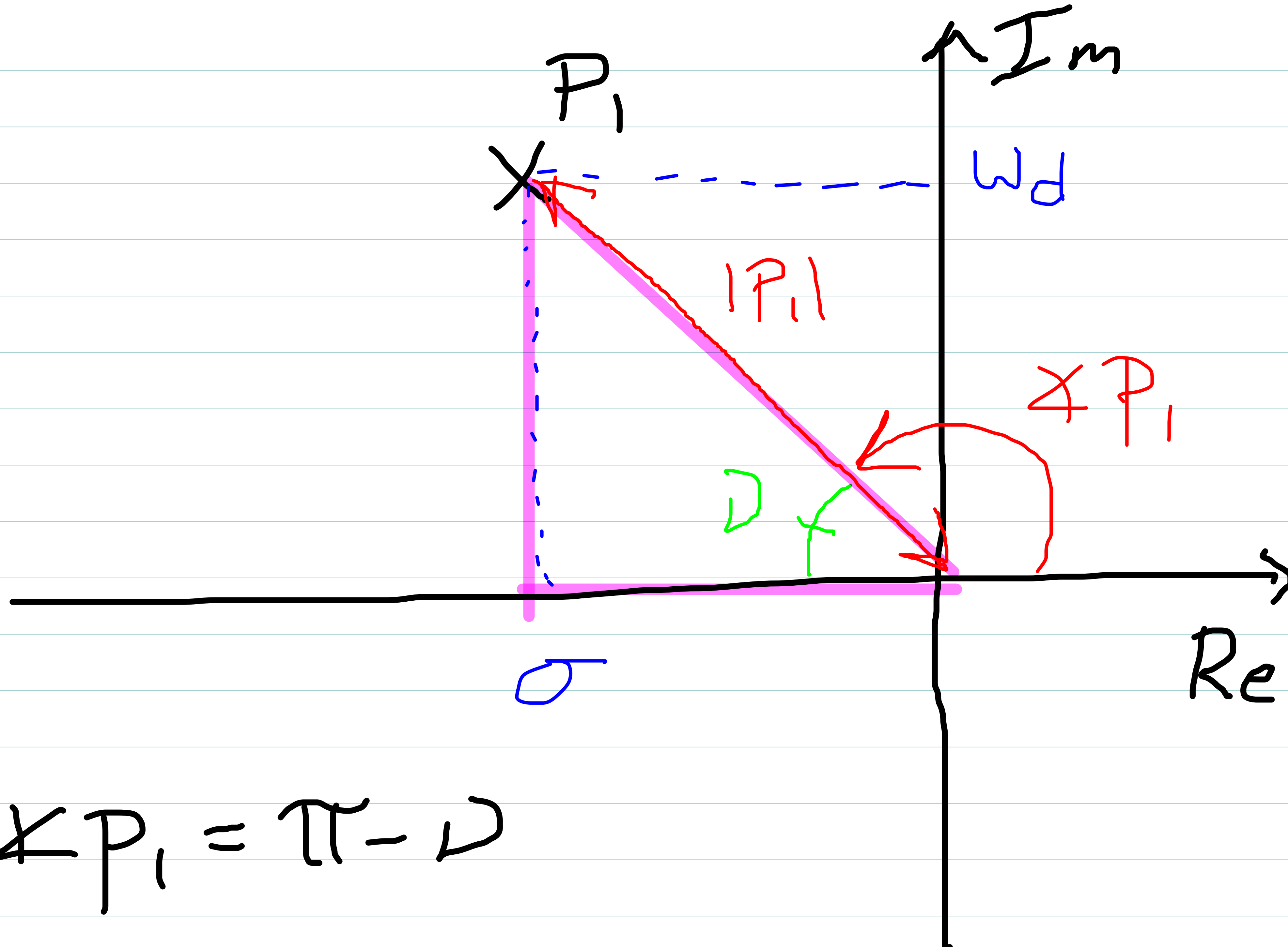
where:
$$B = \frac{\alpha_0}{(j\omega_d)(\sigma + j\omega_d)} = \frac{|p_1|^2}{(j\omega_d)\, p_1}$$

$\Rightarrow$ Transient features completely determined by location of pole $p_1 = \sigma + j\omega_d$ in complex plane

$$|B| = \frac{|p_1|^2}{|j\omega_d| \cdot |p_1|} = \frac{|p_1|}{\omega_d}$$

$$\measuredangle B = \cancel{\measuredangle |p_1|^2} \xrightarrow{} \phi - \left(\measuredangle (j\omega_d) + \measuredangle p_1\right)$$

$$= -\left(\frac{\pi}{2} + \measuredangle p_1\right)$$ must quantify this!



$$\measuredangle P_1 = \pi - \nu$$

<u>Note</u>: $\nu > \measuredangle$ is supplement of $\measuredangle P_1$

So:

$$\measuredangle B = -\left(\frac{\pi}{2} + \measuredangle p_1\right) = -\left(\frac{\pi}{2} + (\pi - \nu)\right)$$

$$= -\frac{3\pi}{2} + \nu$$

and thus:

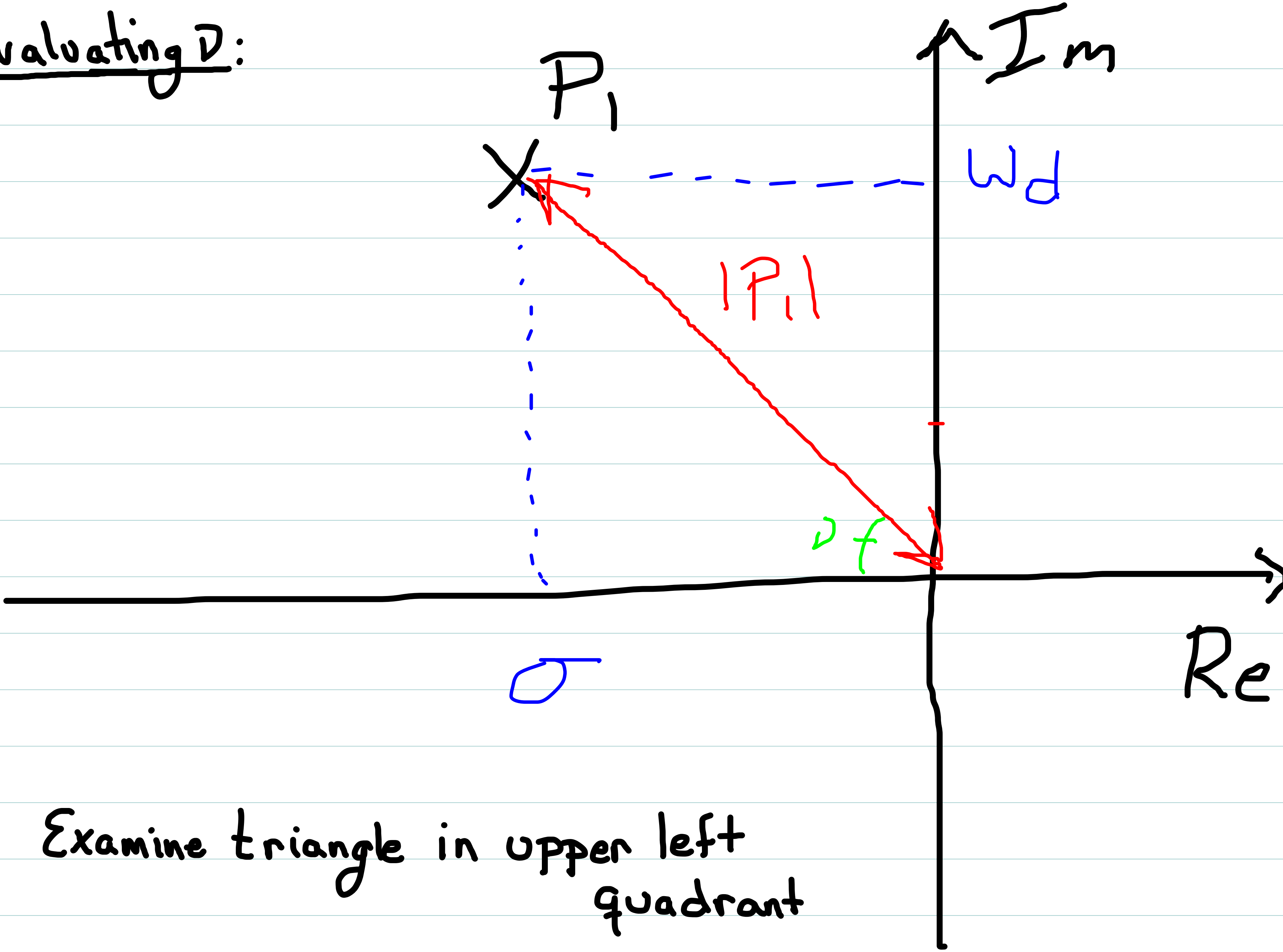
$$y(t) = G(0)\left[1 + \left(\frac{|p_1|}{\omega_d}\right) e^{\sigma t} \cos\left(\omega_d t - \frac{3\pi}{2} + \nu\right)\right]$$

so:

$$\boxed{y(t) = G(0)\left[1 - \left(\frac{|p_1|}{\omega_d}\right) e^{\sigma t} \sin(\omega_d t + \nu)\right]}$$

Need to understand how ν depends on p_1

Evaluating $\mathcal{D}$:



Examine triangle in upper left quadrant

Two Useful Parameters

Define: $\omega_n = |p_1| = \sqrt{\sigma^2 + \omega_d^2}$ "natural" frequency

$\Rightarrow$ purely theoretical! ω_d is physical frequency of transient oscillations

Define: $\xi = \dfrac{|\sigma|}{\omega_n} = \dfrac{|\sigma|}{\sqrt{\sigma^2 + \omega_d^2}}$

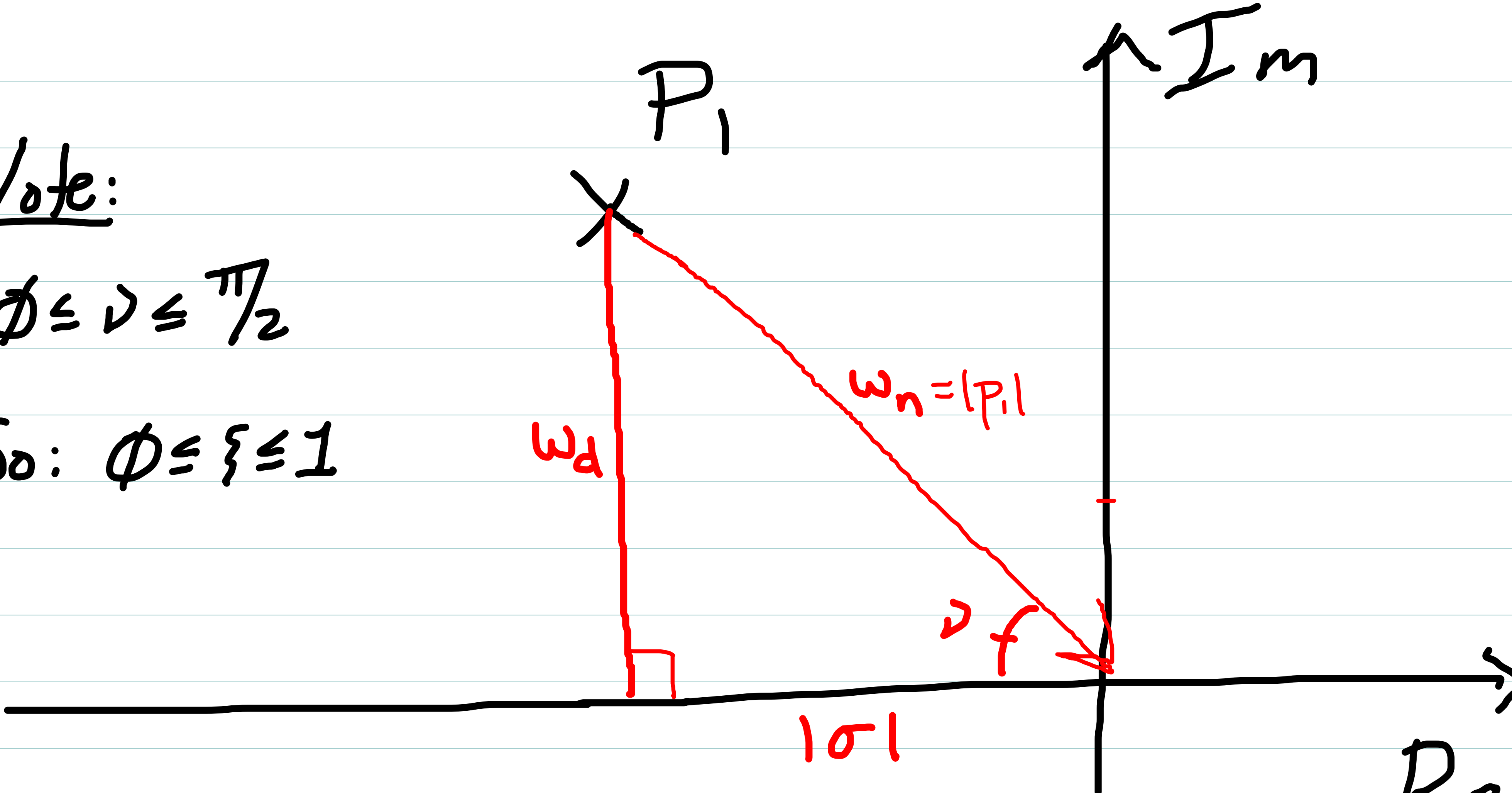
"Damping ratio"

$\Rightarrow$ A normalized measure of the number of transient oscillations observed before amplitude becomes negligible

Note:

$\emptyset \le \nu \le \pi/2$

So: $\emptyset \le \zeta \le 1$



$$\nu = \tan^{-1}\left(\frac{\omega_d}{|\sigma|}\right)$$

$$\nu = \sin^{-1}\left(\frac{\omega_d}{\omega_n}\right)$$

$$\nu = \cos^{-1}\left(\frac{|\sigma|}{\omega_n}\right) = \cos^{-1}\zeta$$ $\Leftarrow$ very useful!

Thus finally, the Case 1 step response is:

$$\Rightarrow \boxed{y(t) = G(0)\left[1 - \left(\frac{\omega_n}{\omega_d}\right) e^{\sigma t} \sin(\omega_d t + \cos^{-1}\zeta)\right]}$$

We can now solve for important transient parameters

$\Rightarrow$ t_c: Solve for first $t > 0$ such that

$$y(t) = y_{ss}(t) = G(0)$$

$$\Rightarrow \sin(\omega_d t + \cos^{-1}\zeta) = 0$$

$$\Rightarrow t_c = \frac{\pi - \cos^{-1}\zeta}{\omega_d}$$

or:

$$\boxed{t_c = \frac{\pi - \rho}{\omega_d}}$$

$\Rightarrow$ For t_p, y_p

Solve for first $t>0$ such that

$$\dot{y}(t) = 0$$

$$\Rightarrow \boxed{t_p = \frac{\pi}{\omega_d}}$$

Substituting:

$$y_p = y(t_p) = G(0)\left[1 + e^{(\sigma\pi/\omega_d)}\right]$$

<u>Define</u>:

$$\boxed{\begin{aligned} M_p &= e^{(\sigma\pi/\omega_d)} \\ \text{then:} \quad y_p &= G(0)\left[1 + M_p\right] \end{aligned}}$$

Peak Overshoot

⇒ M_p is the Normalized peak overshoot

$$y_p = G(0)[1+M_p] \implies M_p = \frac{y_p - G(0)}{G(0)} = \frac{y_p - y_{ss}}{y_{ss}}$$

⇒ M_p is entirely determined by damping ratio ξ

$$M_p = \exp\left[\frac{\sigma\pi}{\omega_d}\right]$$

$$= \exp\left[\frac{(-\xi\omega_n)\pi}{\omega_n\sqrt{1-\xi^2}}\right]$$

OR

$$\boxed{M_p = \exp\left[\frac{-\xi\pi}{\sqrt{1-\xi^2}}\right]}$$

$$\%OS = 100 \times M_p$$

Summary: Case I step response; $p_1 = \sigma + j\omega_d$

"Natural" frequency: $\omega_n = \sqrt{\sigma^2 + \omega_d^2} = |p_1|$

Damping ratio: $\xi = \dfrac{|\sigma|}{\omega_n}$

1st crossing: $t_c = \dfrac{\pi - \cos^{-1}\xi}{\omega_d} = \dfrac{\pi - \nu}{\omega_d}$, $\xi = \cos\nu$

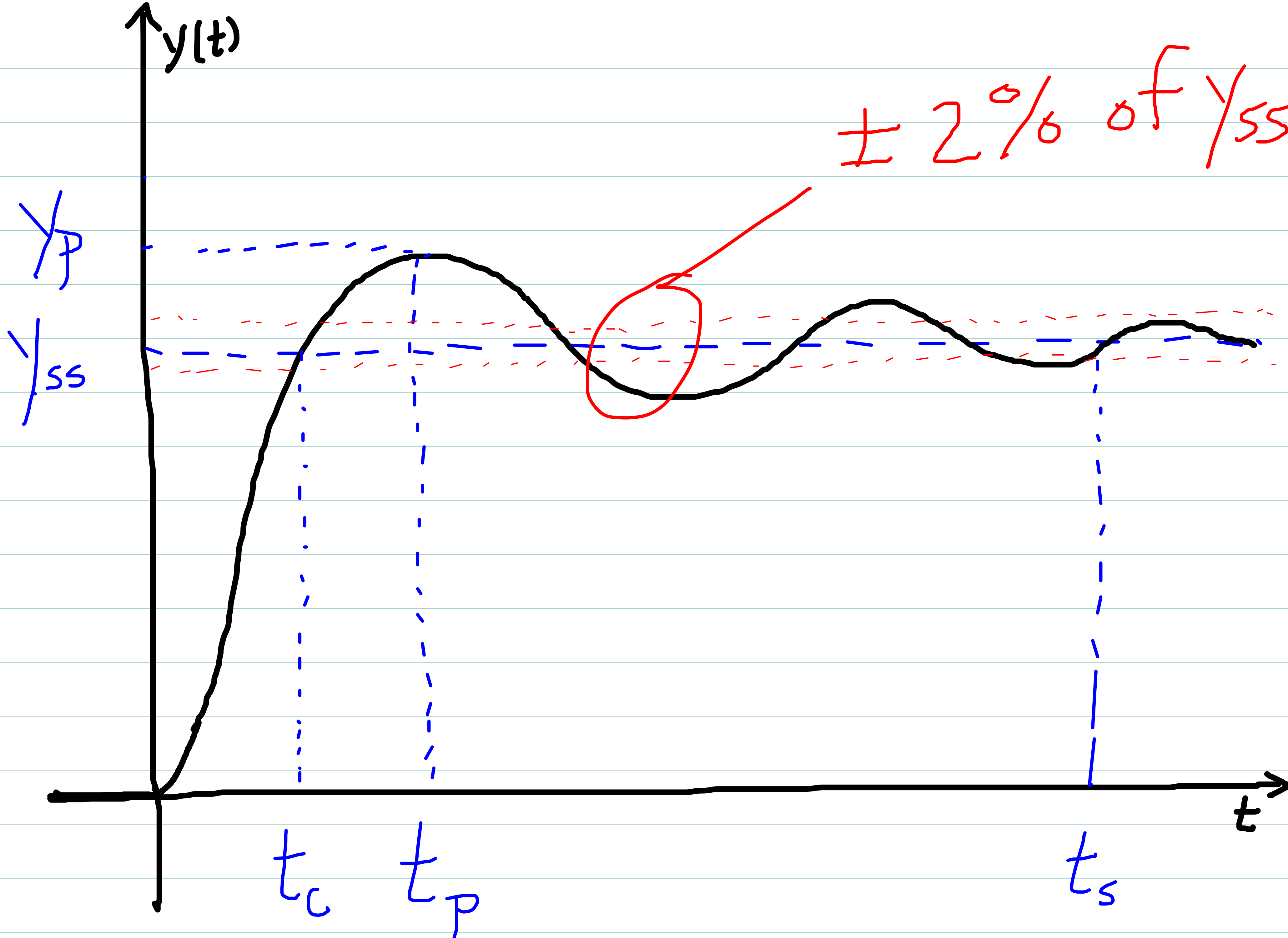
1st peak: $t_p = \dfrac{\pi}{\omega_d}$

Normalized overshoot: $M_p = \exp\left[\dfrac{\sigma\pi}{\omega_d}\right] = \exp\left[\dfrac{-\xi\pi}{\sqrt{1-\xi^2}}\right]$

$$M_p = \left[\frac{y_p - y_{ss}}{y_{ss}}\right]$$

Peak response: $y_p = y_{ss}[1 + M_p]$

$y_{ss} = G(\emptyset)$ for unit step



y(t)
y_p
y_{ss}
$\pm 2\%$ of y_{ss}
t
t_c
t_p
t_s

Settling Time

As usual, we can use the approximation

$$\boxed{t_s \approx \frac{4}{|Re\{p_i\}|} = \frac{4}{|\sigma|}}$$

But t_s is actually a function of ξ also here:

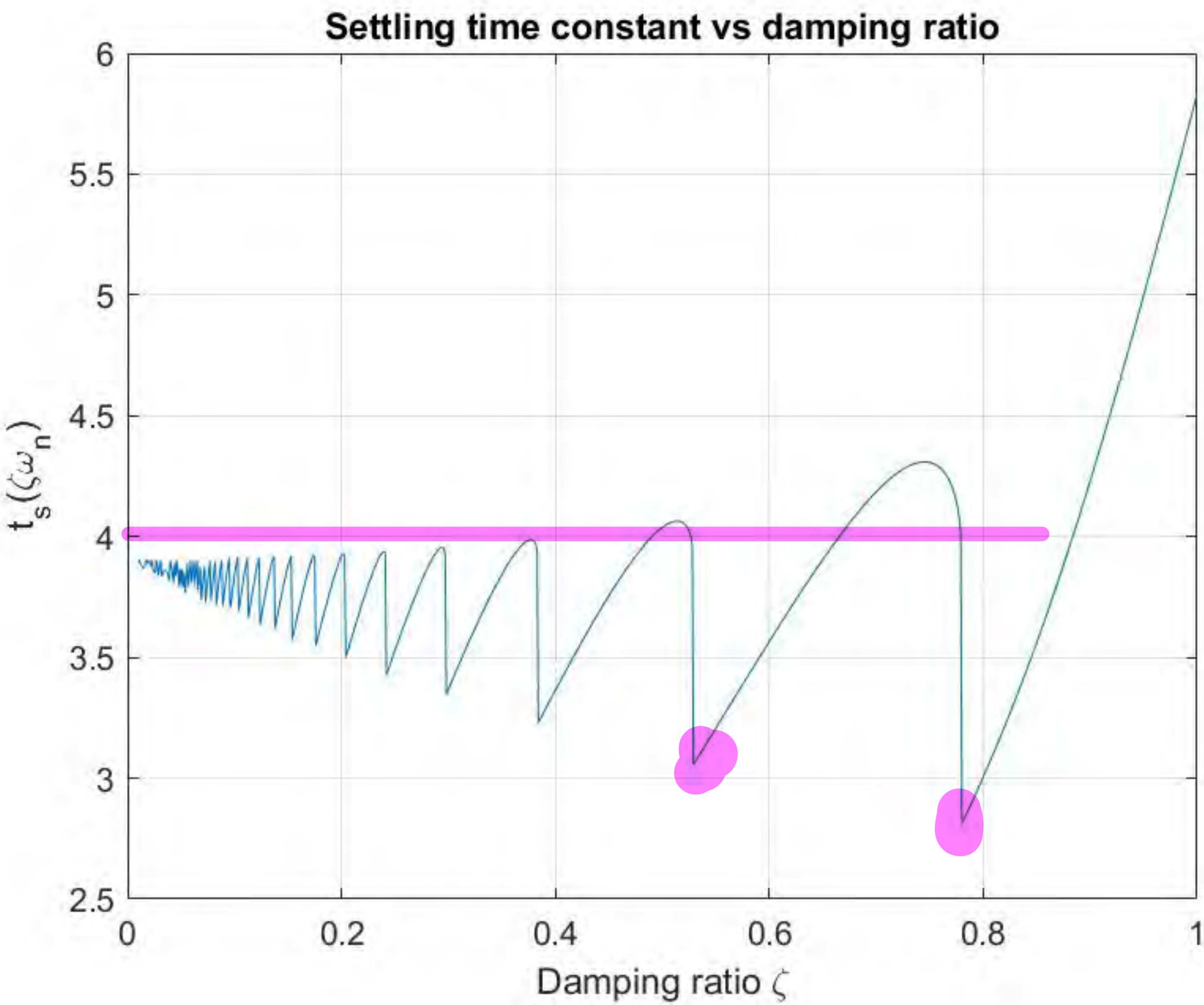
$$t_s = \frac{C(\xi)}{|\sigma|}$$

with $3 \leq C(\xi) \leq 5$ for most $0 \leq \xi < 0.9$

so 4 is an "average" value for $C(\xi)$

However for $0.95 \leq \xi \leq 1$ a better approximation is: $t_s \approx \dfrac{6}{|\sigma|}$

Complex poles $-$ $C(\xi)$ $\quad t_s = \dfrac{C(\xi)}{\xi\omega_n}$



A few more observations

$$\xi = \frac{|\sigma|}{\omega_n} \implies \boxed{\sigma = -\xi\omega_n}$$ (Stable System Assumed)

$$\omega_n = \sqrt{\sigma^2 + \omega_d^2}$$

$$\implies \omega_d^2 = \omega_n^2 - \sigma^2 = \omega_n^2 - (-\xi\omega_n)^2 = \omega_n^2(1-\xi^2)$$

So: $$\boxed{\omega_d = \omega_n\sqrt{1-\xi^2}}$$

Then note:

$$\boxed{\begin{aligned} s^2+\alpha_1 s+\alpha_0 &= (s-p)(s-\bar{p}) \\ &= s^2 - 2\sigma s + (\sigma^2+\omega_d^2) \\ &= s^2 + 2\xi\omega_n s + \omega_n^2 \end{aligned}}$$

all equivalent

Note:

The three possible cases for 2^{nd} order step responses can be categorized by ζ:

Case 1 (complex poles): $0 \le \zeta < 1$

$$\alpha_1^2 < 4\alpha_0 \Rightarrow (2\zeta\omega_n)^2 < 4\omega_n^2 \quad ✓$$

Case 2 (repeated real poles): $\zeta = 1$

$$\alpha_1^2 = 4\alpha_0 \Rightarrow (2\cancelto{1}{\zeta}\omega_n)^2 = 4\omega_n^2 \quad ✓$$

Case 3 (distinct real poles): $\zeta > 1$

$$\alpha_1^2 > 4\alpha_0 \Rightarrow 4\zeta^2\omega_n^2 > 4\omega_n^2 \quad ✓$$

Note:

The three possible cases for 2^{nd} order step responses can be categorized by ξ:

Case 1 (complex poles): $0 \leq \xi < 1$ "underdamped"

$$\alpha_1^2 < 4\alpha_0 \Rightarrow (2\xi\omega_n)^2 < 4\omega_n^2 \quad ✓$$

Case 2 (repeated real poles): $\xi = 1$ "critically damped"

$$\alpha_1^2 = 4\alpha_0 \Rightarrow (2\cancelto{1}{\xi}\omega_n)^2 = 4\omega_n^2 \quad ✓$$

Case 3 (distinct real poles): $\xi > 1$ "overdamped"

$$\alpha_1^2 > 4\alpha_0 \Rightarrow 4\xi^2\omega_n^2 > 4\omega_n^2 \quad ✓$$

Limiting case: $\xi \to \emptyset$

$\xi \to \emptyset \implies \sigma = -\xi\omega_n \to \emptyset \implies P_1 = j\omega_d$ (pure imaginary)

Overshoot $M_p = e^{(\sigma\pi/\omega_d)} \to 1$ (100% OS)

Peak: $y_p = G(\emptyset)[1 + M_p]$

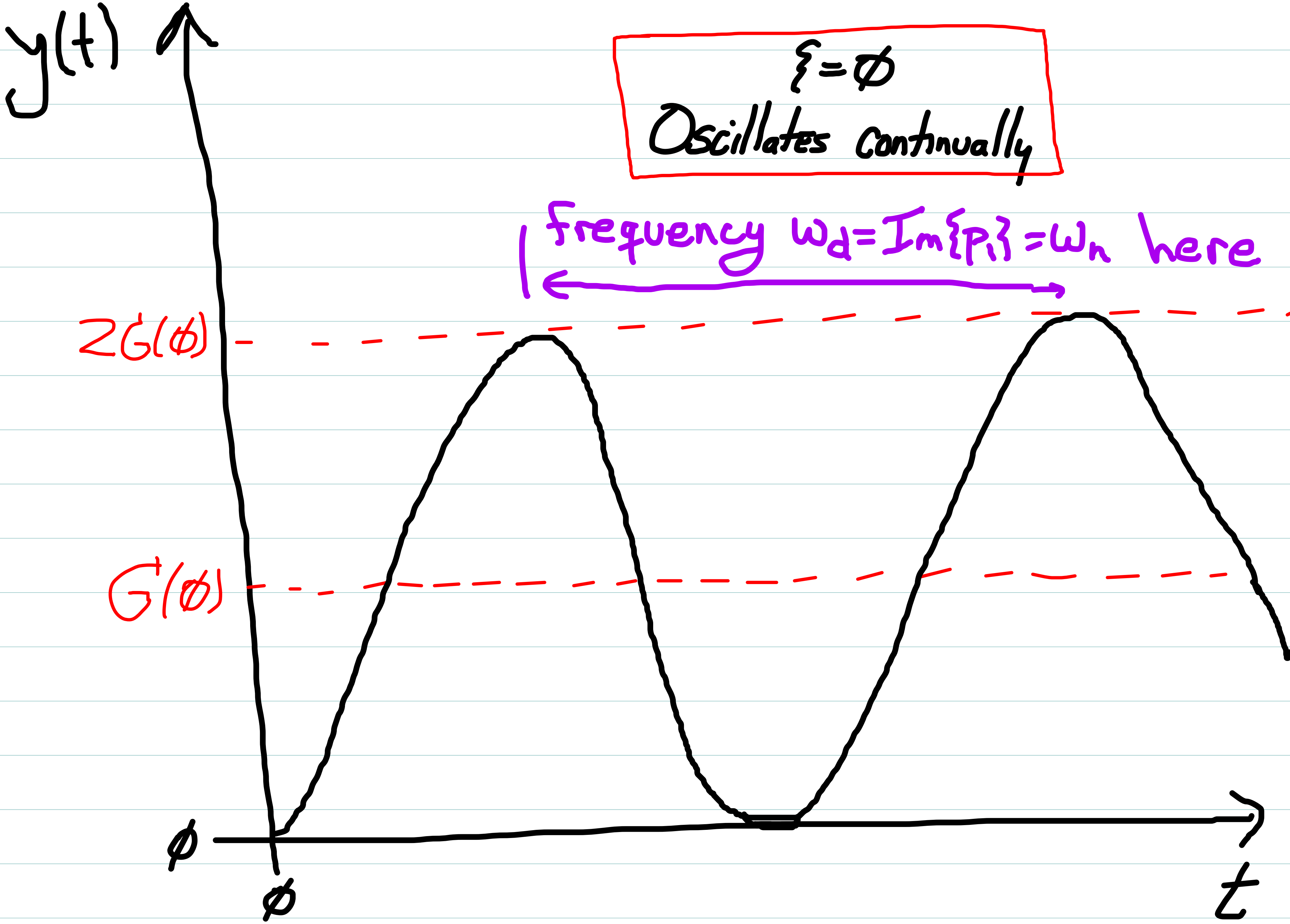
or $y_p = 2y_{ss}$

Settling time: $t_s \approx \dfrac{4}{|\sigma|} = \infty$

Never settles!

Response oscillates infinitely between $\emptyset$ and $2G(\emptyset)$ with frequency $\omega_d = \omega_n\sqrt{1-\xi^2} = \omega_n$

"Undamped"


y(t)
ξ=∅
Oscillates continually
frequency ω_d=Im{p_i}=ω_n here
2G(∅)
G(∅)
∅
∅
t

Limiting Case, $\xi \to 1$

$$\xi \to 1 \Rightarrow \quad \sigma = -\xi\omega_n \to -\omega_n$$

$$\omega_d = \omega_n\sqrt{1-\xi^2} \to \phi$$

Response does not oscillate!

Overshoot:

$$M_p = e^{(\sigma\pi/\omega_d)} = e^{-\omega_n\pi/\phi} = \phi$$

No overshoot

1st crossing:

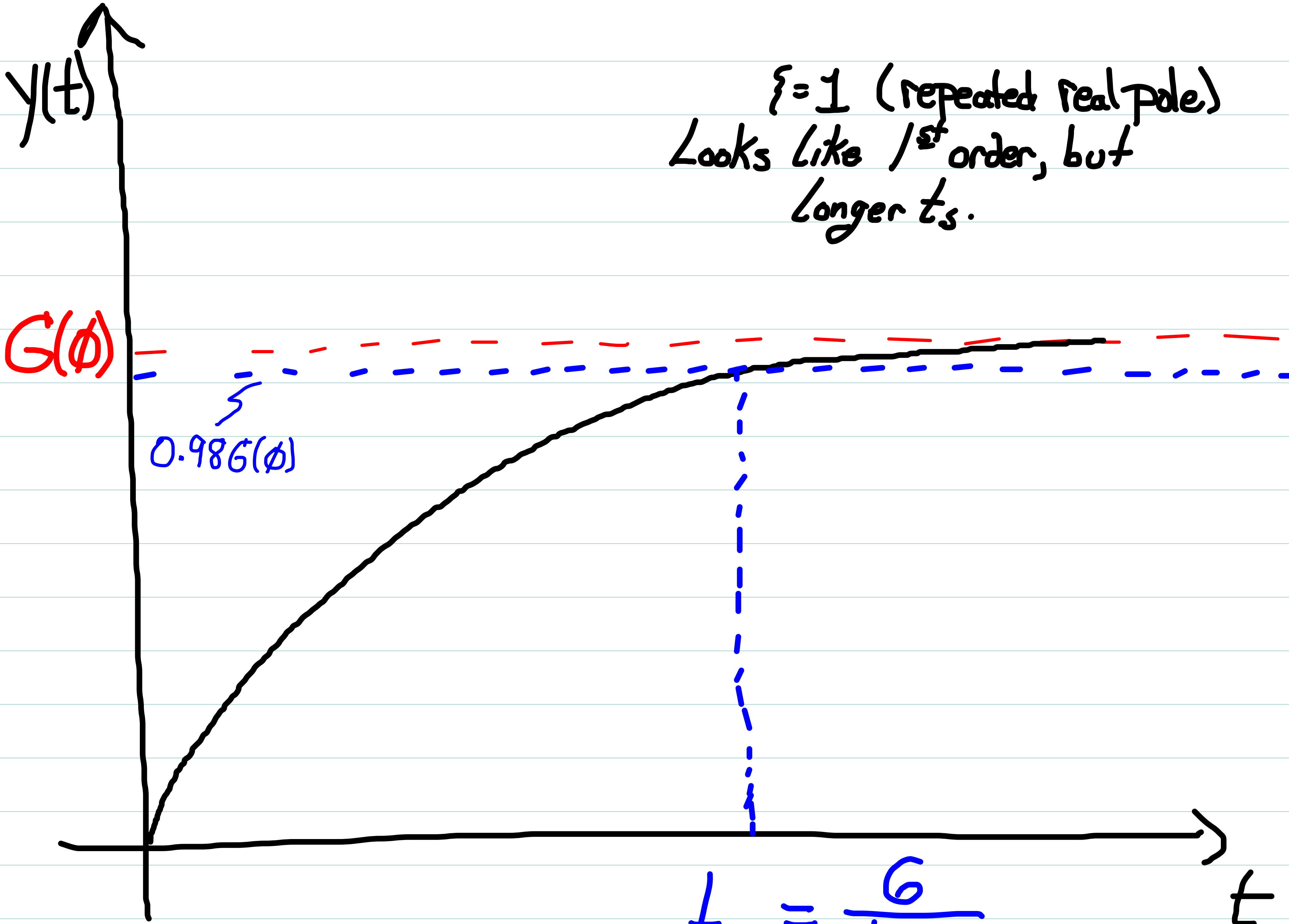
$$t_c = \frac{\pi - \cos^{-1}\xi}{\omega_d} = \pi/2/\phi = \infty$$

$\Rightarrow$ response asymptotes to y_{ss} from below

Settling:

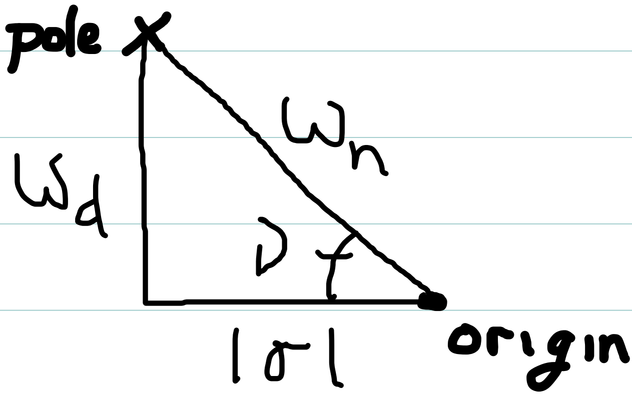
$$t_s \approx \frac{6}{|\sigma|}$$ use 6 here

$\zeta = 1$ (repeated real pole)
Looks like 1^{st} order, but longer t_s.



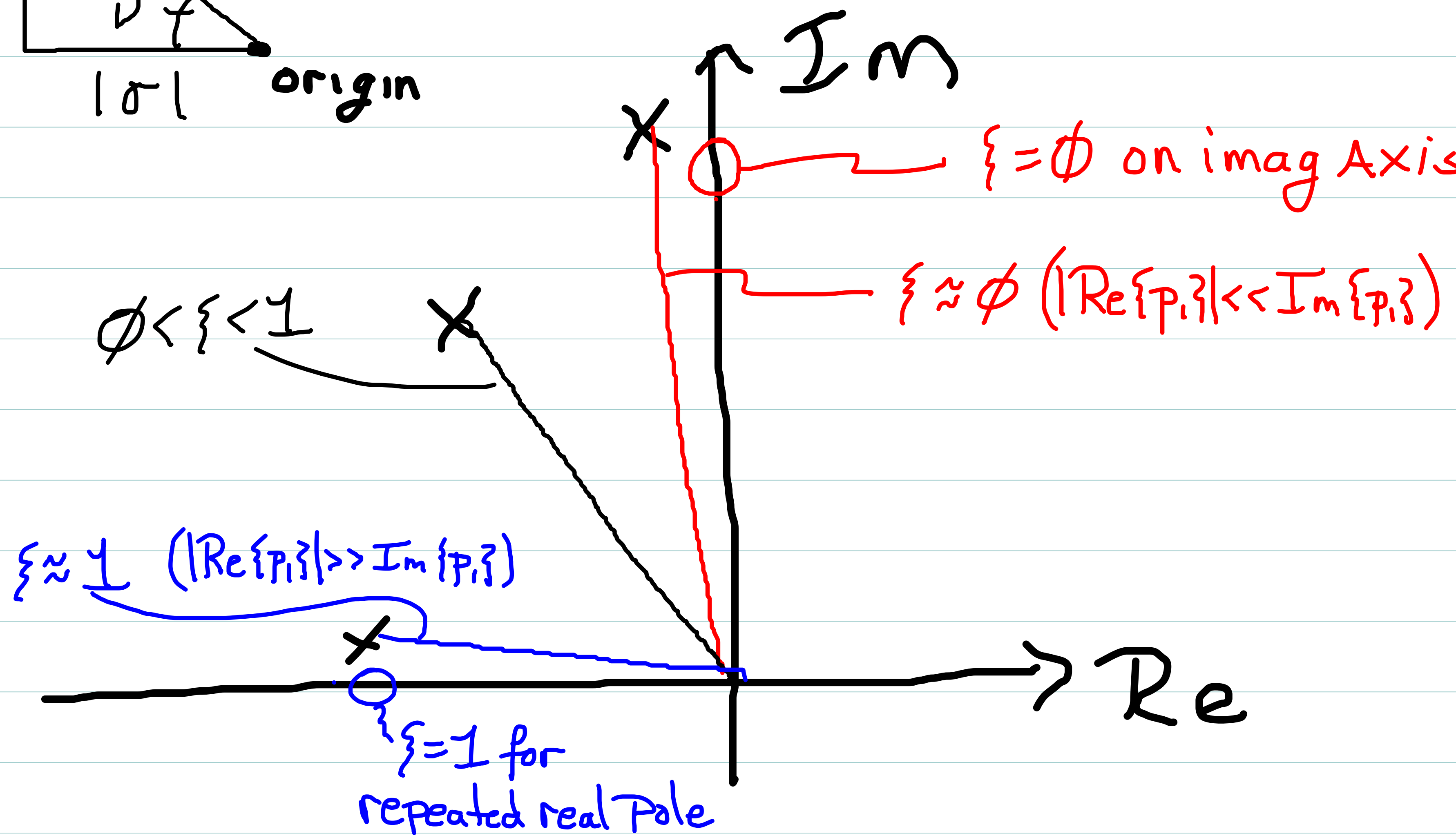
Graphical Interpretation of ξ:

$\boxed{\xi = \cos\nu}$:

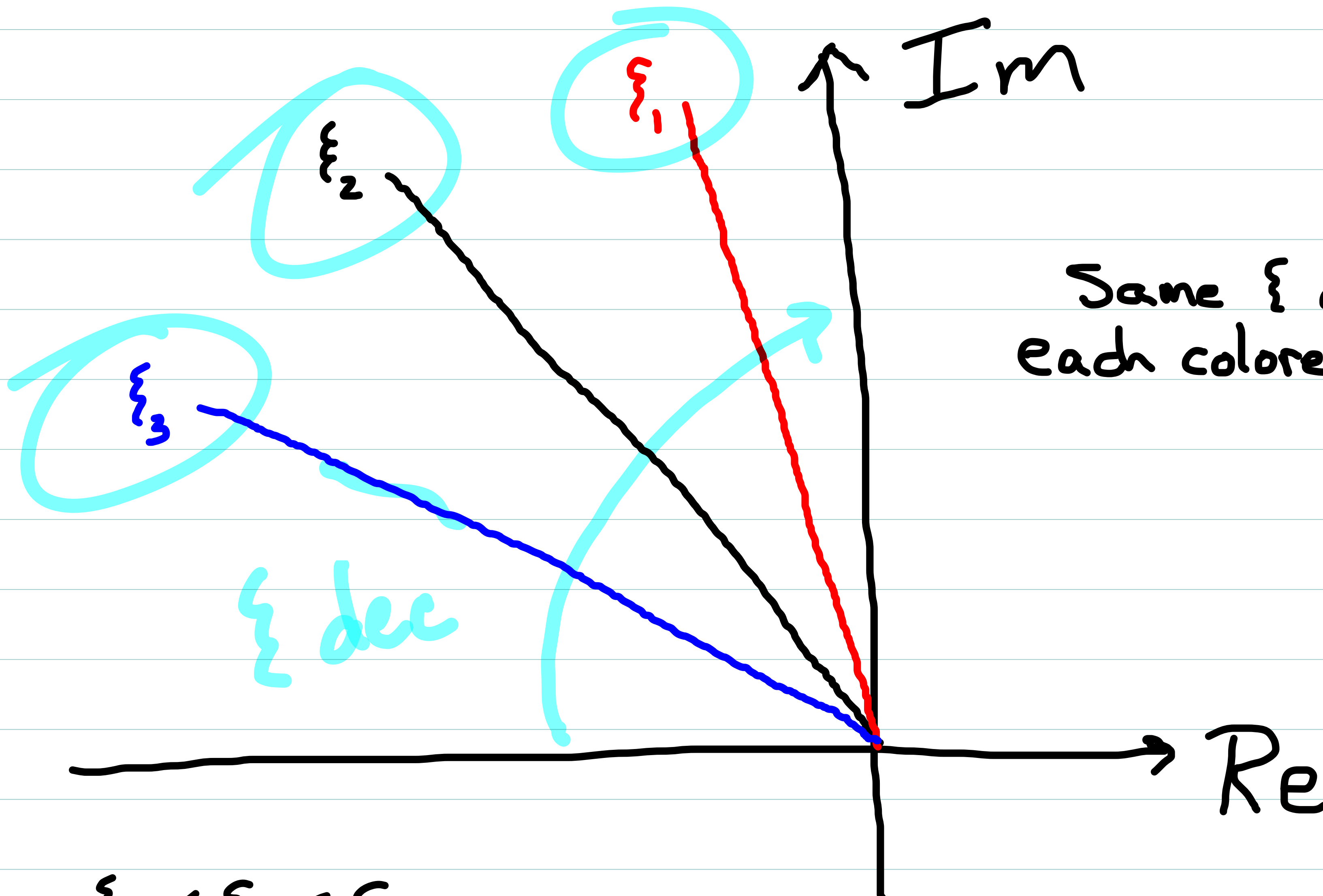


$$\xi \to 0 \Rightarrow \nu \to \pi/2$$

$$\xi \to 1 \Rightarrow \nu \to 0$$



Lines of constant ξ lie on rays in upper left quadrant of complex plane:



Same ξ on each colored line

$\xi_1 < \xi_2 < \xi_3$

=> 1^{st} and 2^{nd} order step responses are "building blocks" by which we can understand response of more complex systems

=> each real pole introduces a new decaying exponential into transient response.

=> each complex pole pair introduces a decaying oscillation into the transient

=> An arbitrary number of poles of different types will typically require numerical simulation to quantify y_p, t_r, t_p, t_s

=> <u>However</u> in some cases we can still accurately predict these features.

Suppose:

$$G(s) = \frac{K}{(s-p_1)(s^2+2\zeta\omega_n s+\omega_n^2)} \quad \text{with } \zeta<1$$

$$= \frac{K}{(s-p_1)(s-p_2)(s-\bar{p}_2)}$$

for a unit step input $u(t) = \mathbb{1}(t)$ we know

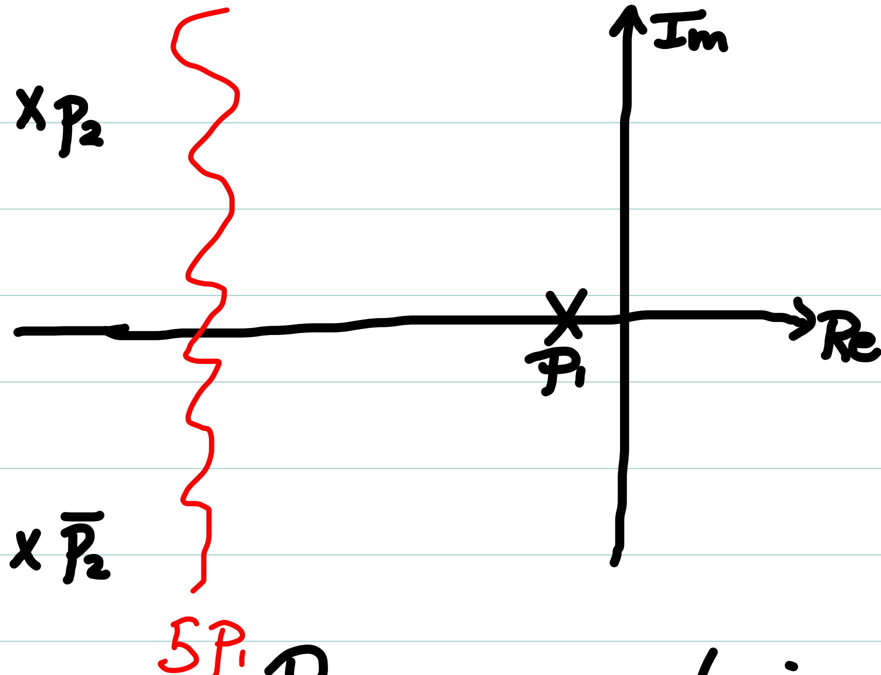
$$y_{ss} = G(0) = \frac{K}{-\omega_n^2 p_1}$$

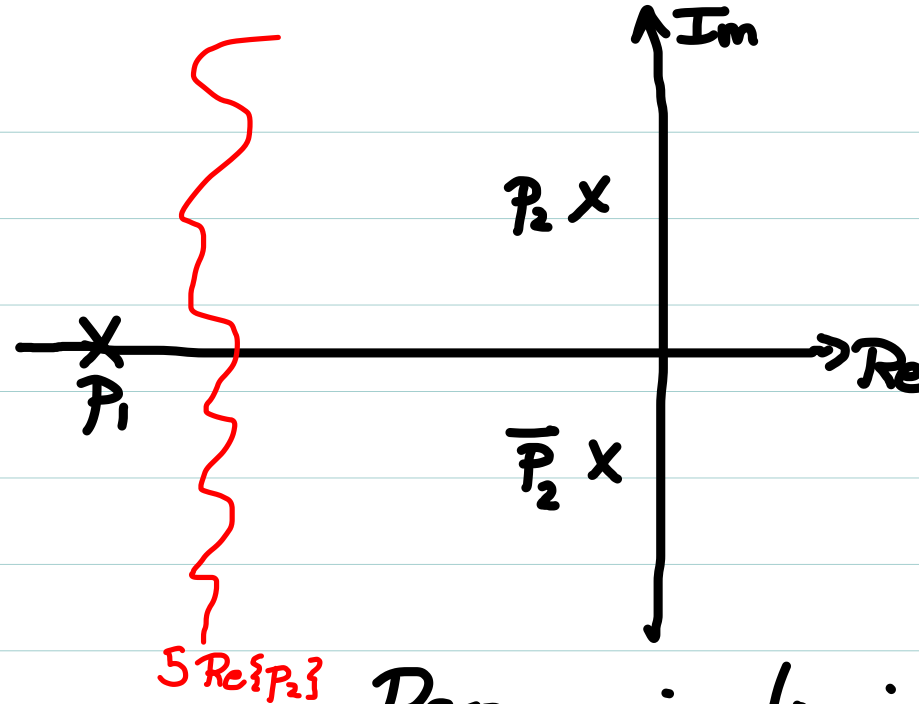
But what can we say about y_p, t_p, t_c, t_s?

In general, not much unless either

$$|p_1| > 5|Re\{p_2\}| \text{ or } |Re\{p_2\}| > 5|p_1|$$

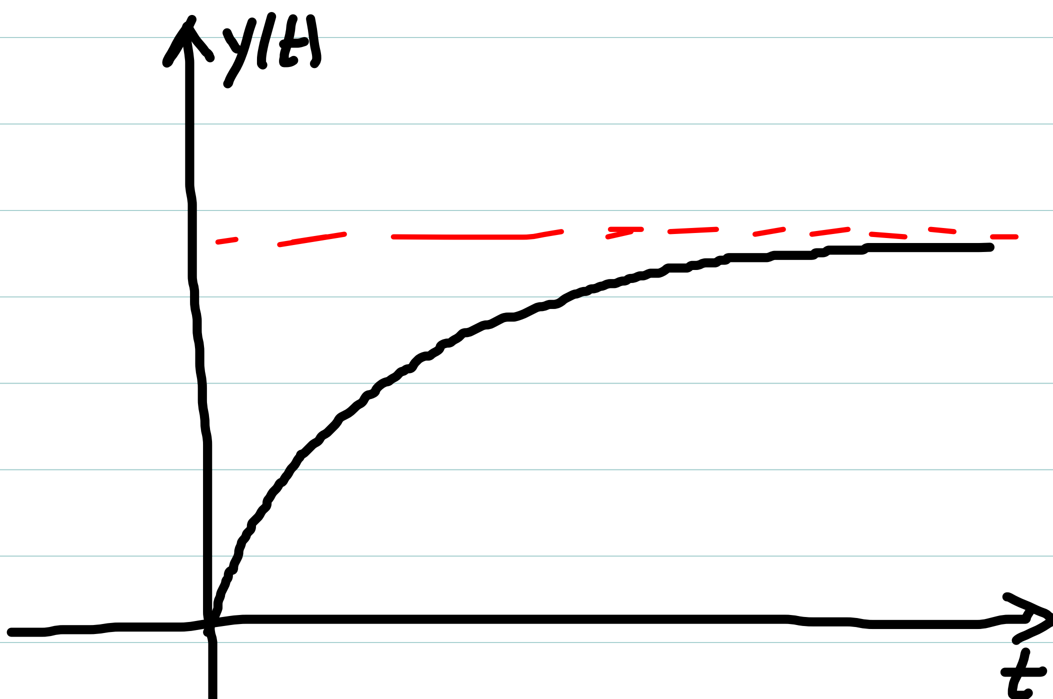
i.e. if one of the modes is <u>dominant</u>.

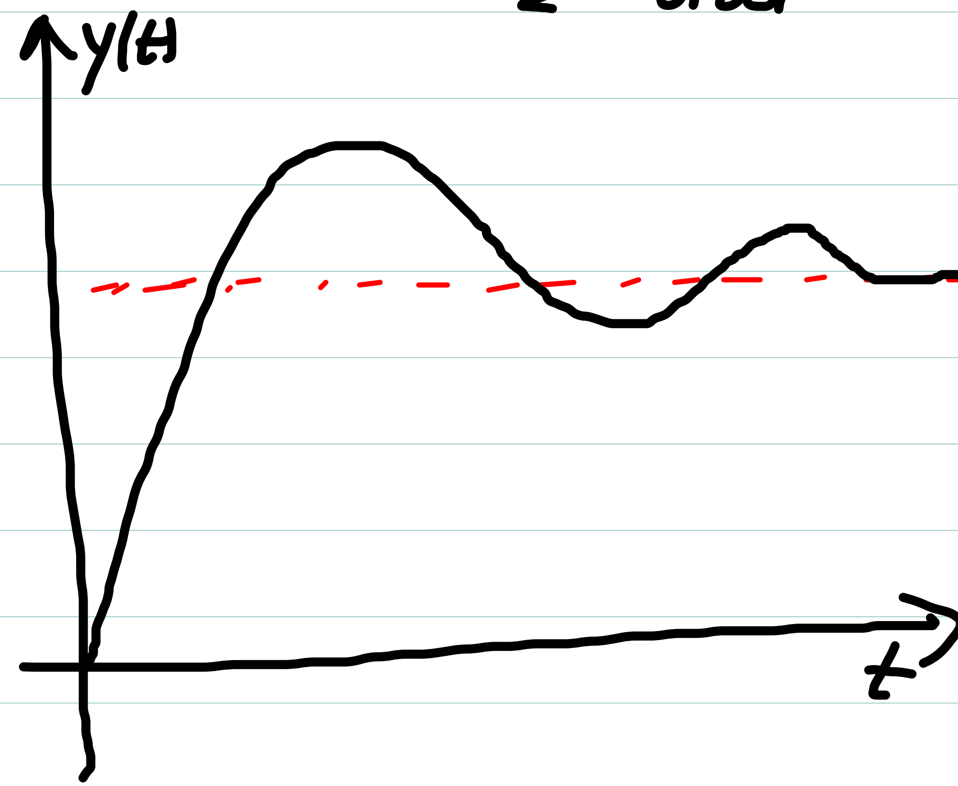




Response is dominantly 1st order

Response is dominant
2nd order

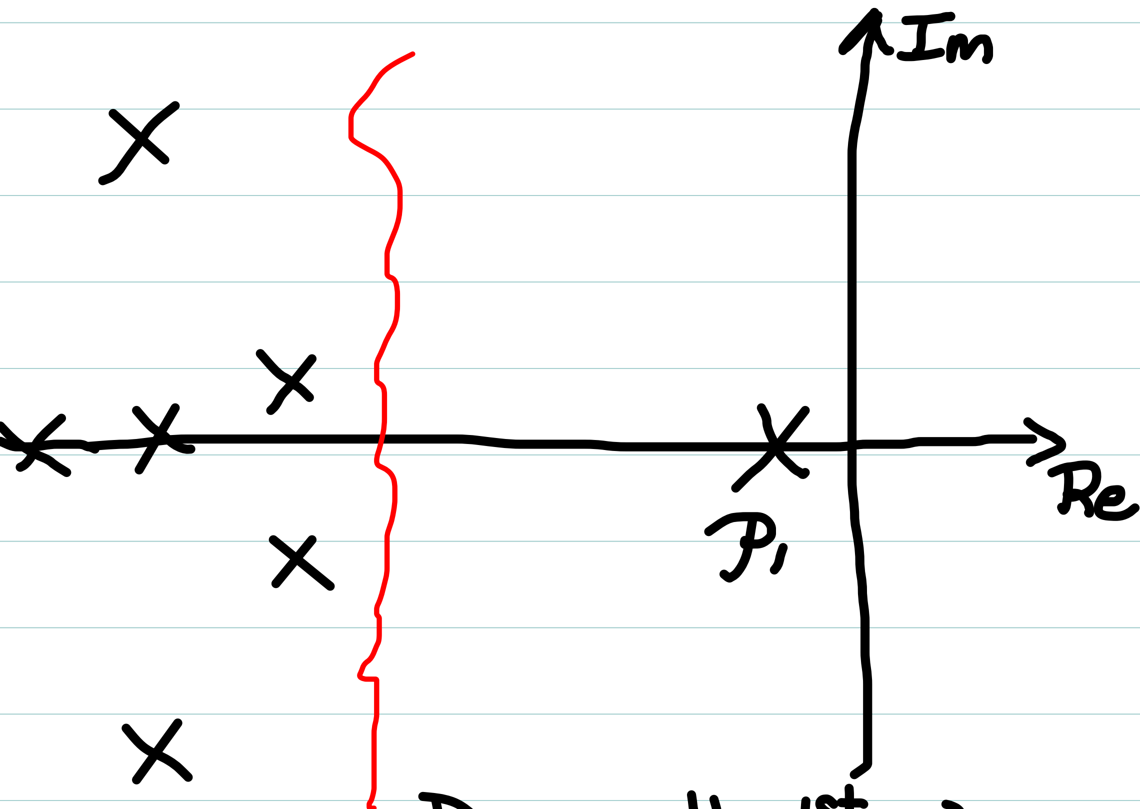




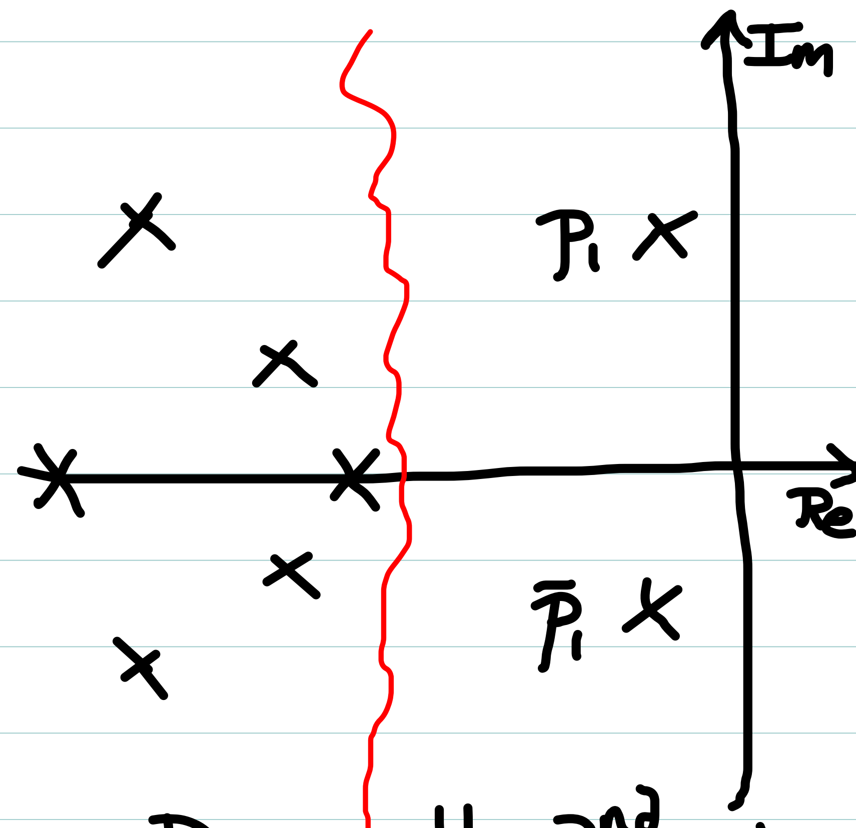
Dominant modes revisited

When a single mode is dominant, we can approximate the features of the response using just that mode

An arbitrarily complex system can be well approximated in this fashion.



Dominantly 1st order



Dominantly 2nd order

Effect of Zeros

Step response of

$$G(s) = \frac{\beta_1 s + \beta_0}{s^2 + \alpha_1 s + \alpha_0}$$

Zero at

$$z_1 = -\beta_0 / \beta_1$$

3 important effects:

1. "Input absorbing" property
2. Transient suppression
3. Transient amplification

Both? Yes!

Depending on system

① Input absorbtion

For unit step response of stable system

$$y_{ss}(t) = G(\emptyset)$$

Suppose $z_1 = -\beta_0/\beta_1 = \emptyset \implies \beta_0 = \emptyset$

zero at origin

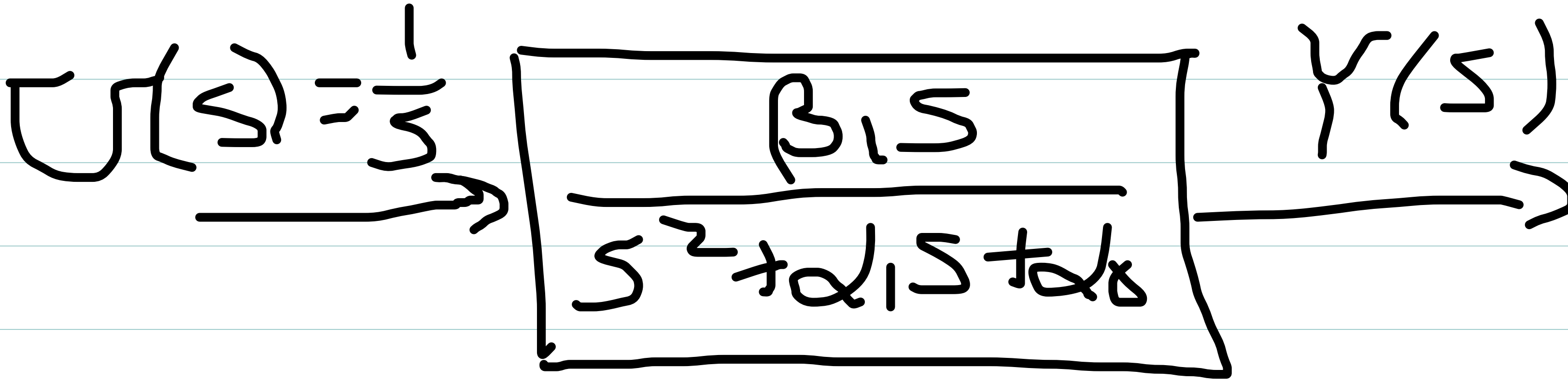
$$G(s) = \frac{\beta_1 s}{s^2 + \alpha_1 s + \alpha_0}$$

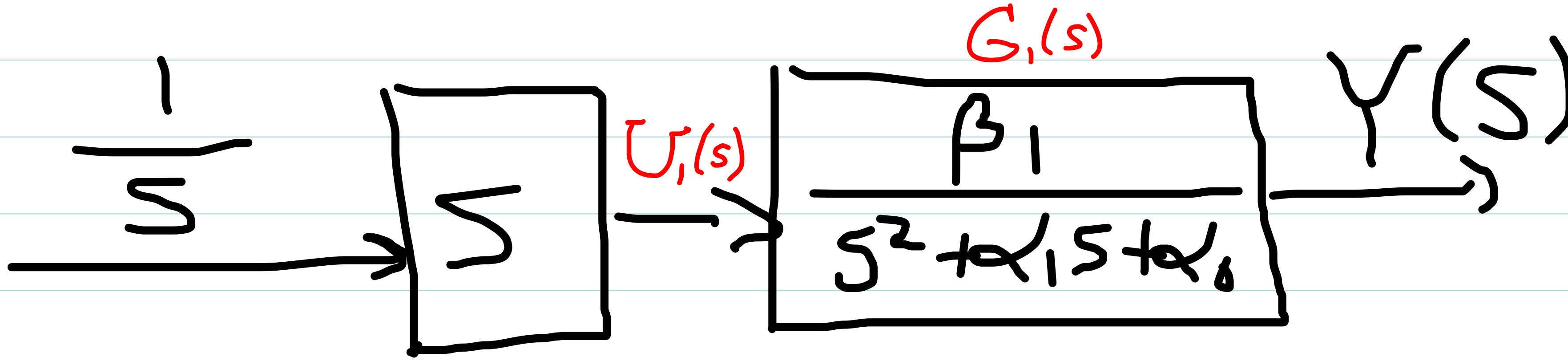
Then $y_{ss}(t) = G(\emptyset) = \emptyset \Leftarrow$ steady-state is zero

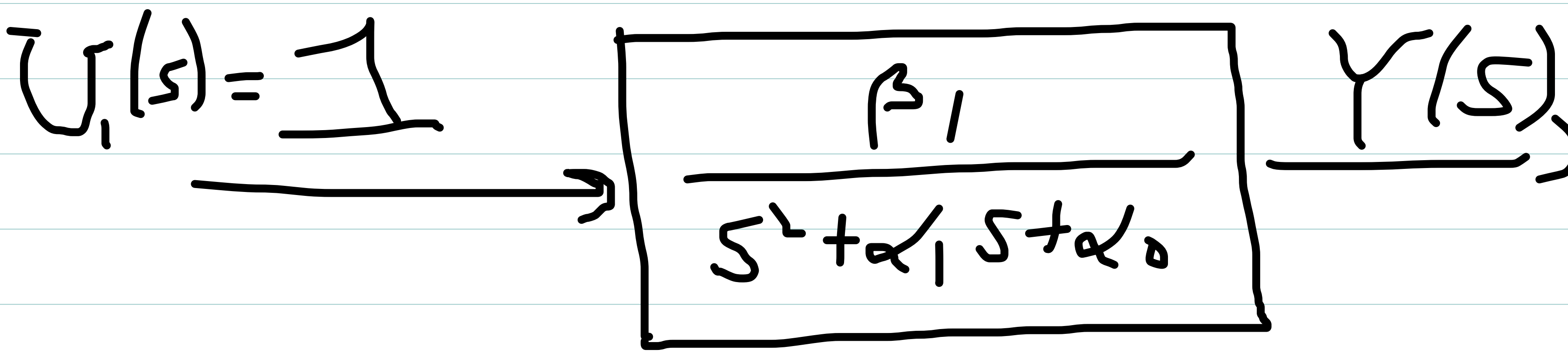
response contains only transient terms

In fact, $y(t)$ is the impulse response of

$$G_1(s) = \frac{\beta_1}{s^2 + \alpha_1 s + \alpha_0}$$


$U(s) = \frac{1}{s}$
$\frac{\beta_1 s}{s^2 + \alpha_1 s + \alpha_0}$
$Y(s)$


$\frac{1}{s}$
s
$U_1(s)$
$G_1(s)$
$\frac{\beta_1}{s^2 + \alpha_1 s + \alpha_0}$
$Y(s)$


$U_1(s) = 1$
$\frac{\beta_1}{s^2 + \alpha_1 s + \alpha_0}$
$Y(s)$

⇒ 1^{st} and 2^{nd} order step responses are "building blocks" by which we can understand response of more complex systems

⇒ each real pole introduces a new decaying exponential into transient response.

⇒ each complex pole pair introduces a decaying oscillation into the transient

⇒ An arbitrary number of poles of different types will typically require numerical simulation to quantify y_p, t_r, t_p, t_s

⇒ <u>However</u> in some cases we can still accurately predict these features.

Suppose:

$$G(s) = \frac{K}{(s-p_1)(s^2+2\zeta\omega_n s+\omega_n^2)} \quad \text{with } \zeta < 1$$

$$= \frac{K}{(s-p_1)(s-p_2)(s-\bar{p}_2)}$$

for a unit step input $u(t) = \mathbb{1}(t)$ we know

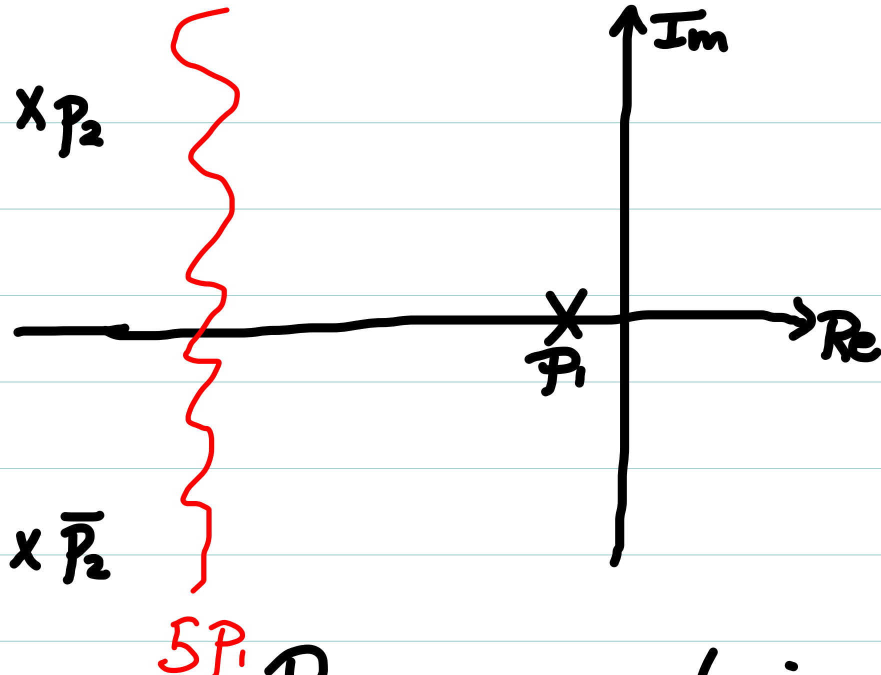
$$y_{ss} = G(0) = \frac{K}{-\omega_n^2 p_1}$$

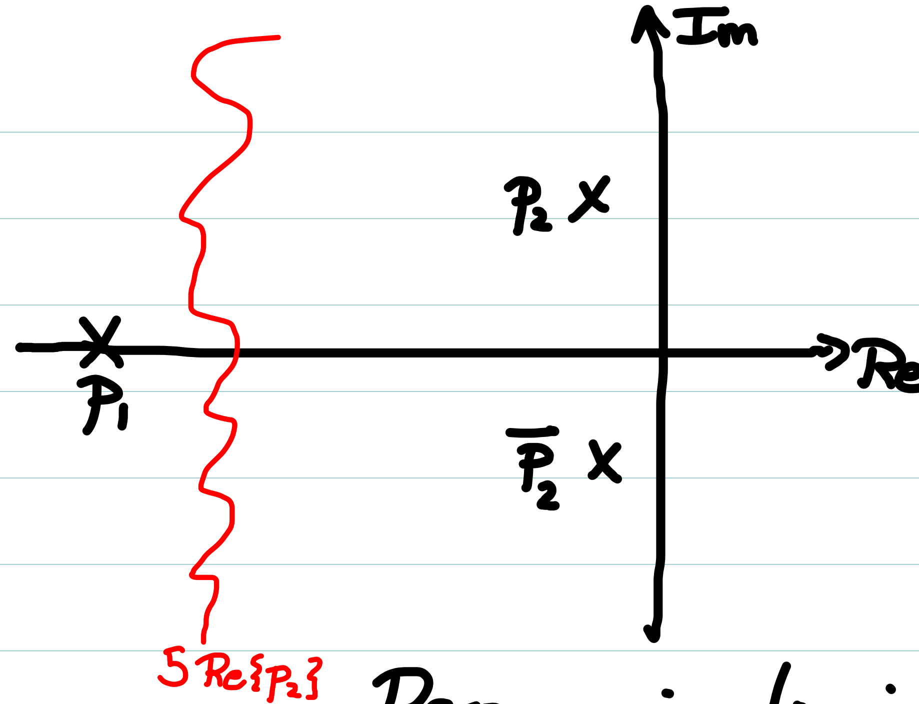
But what can we say about y_p, t_p, t_c, t_s?

In general, not much unless either

$$|p_1| > 5\,|Re\{p_2\}| \text{ or } |Re\{p_2\}| > 5\,|p_1|$$

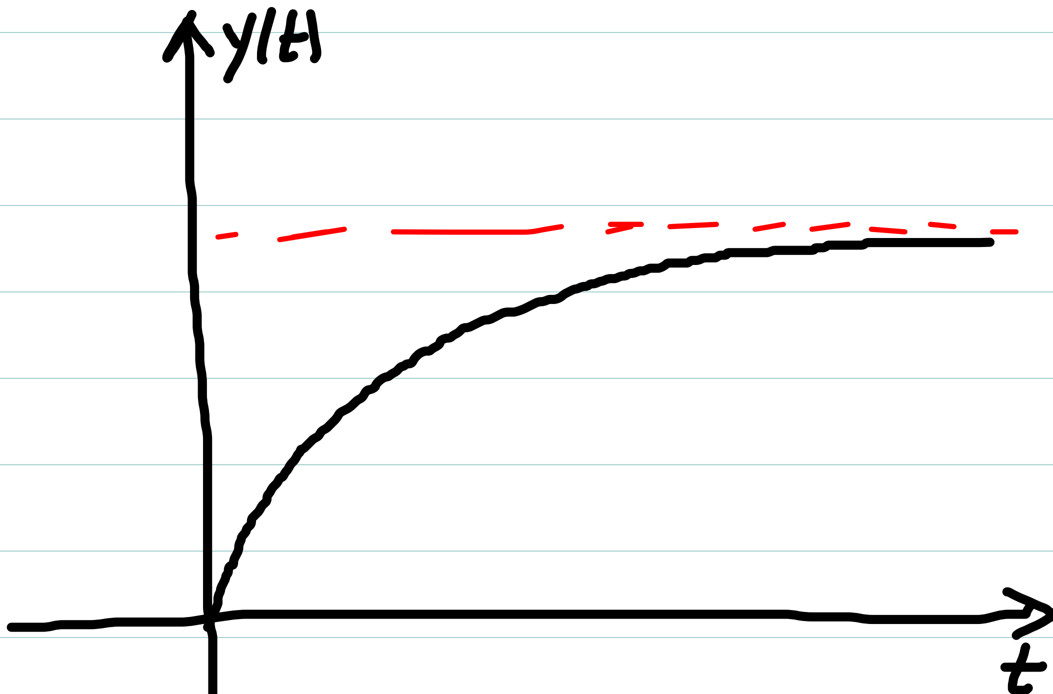
i.e. if one of the modes is <u>dominant</u>.

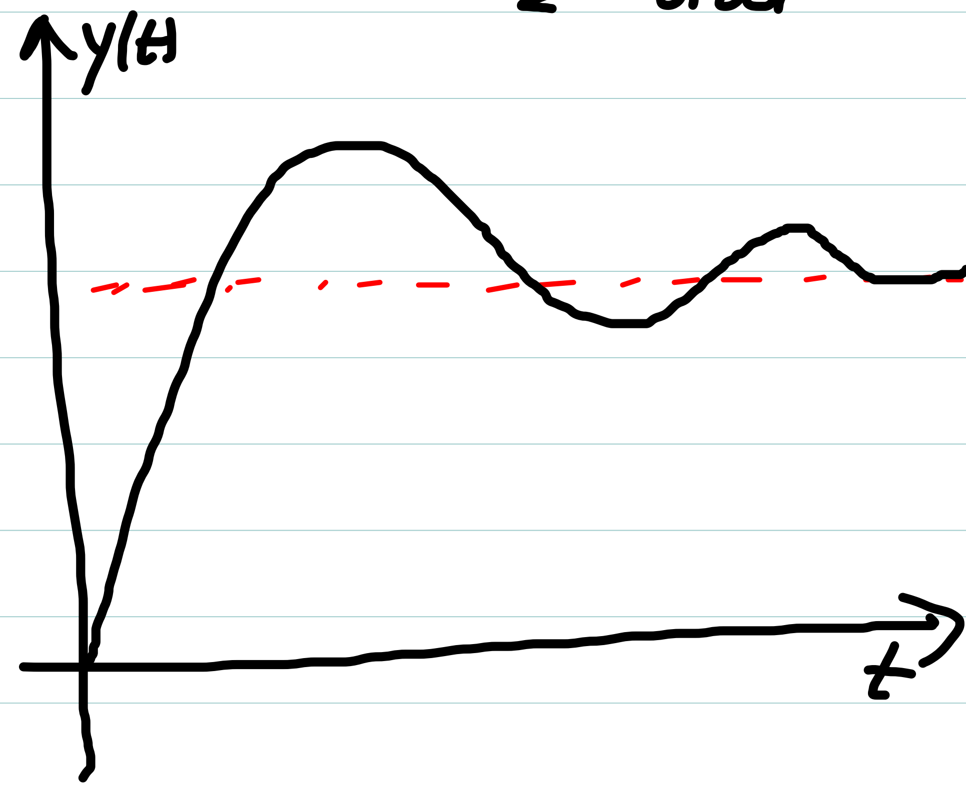




Response is dominantly 1^{st} order

Response is dominant 2^{nd} order

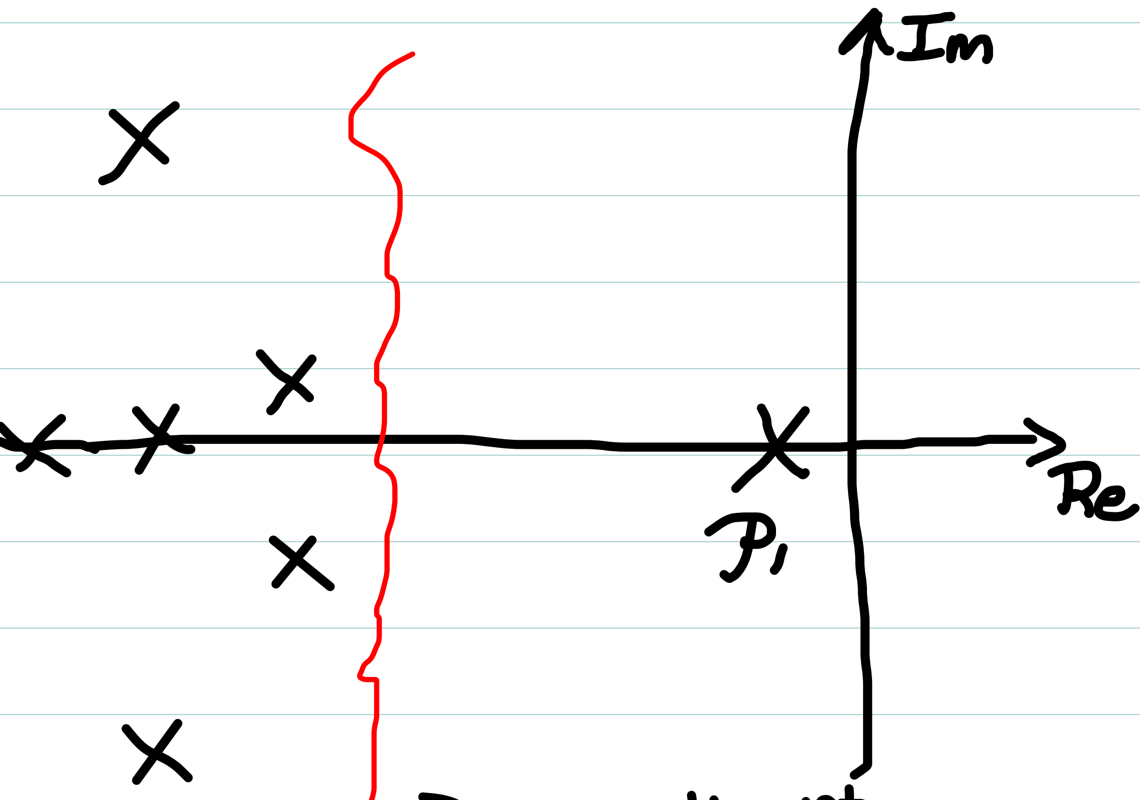




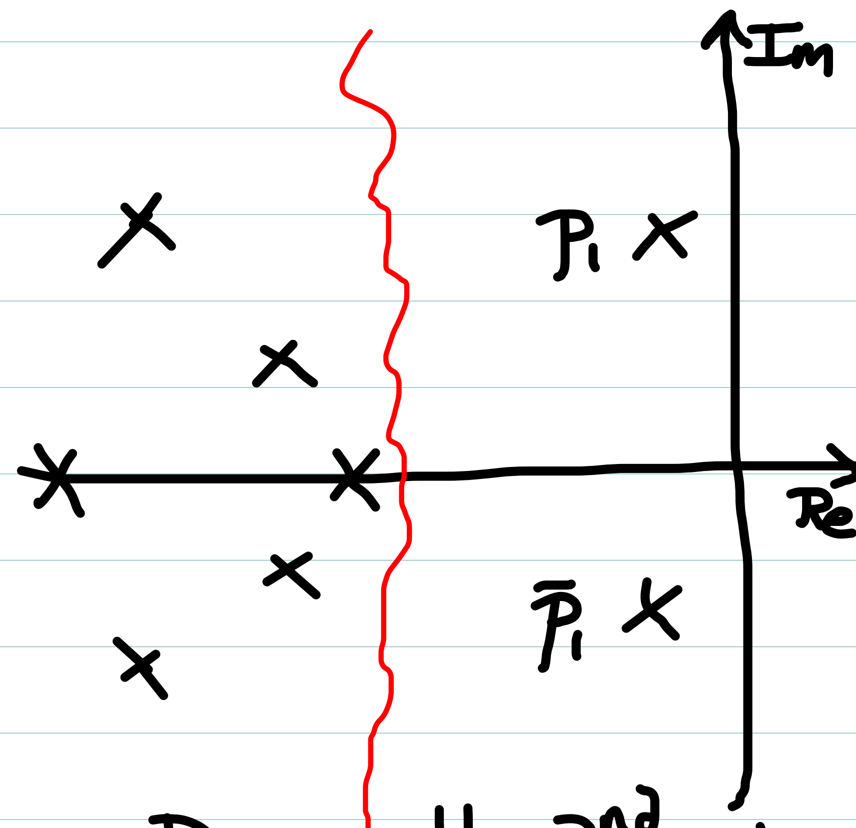
Dominant modes revisited

When a single mode is dominant, we can approximate the features of the response using just that mode.

An arbitrarily complex system can be well approximated in this fashion.



Dominantly 1st order



Dominantly 2nd order

Effect of Zeros

Step response of

$$G(s) = \frac{\beta_1 s + \beta_0}{s^2 + \alpha_1 s + \alpha_0}$$

Zero at $z_1 = -\beta_0/\beta_1$

3 important effects:

1. "Input absorbing" property
2. Transient suppression
3. Transient amplification

Both? Yes!

Depending on system

① Input absorbtion

For <u>unit</u> step response of stable system

$$y_{ss}(t) = G(\emptyset)$$

Suppose $z_1 = {-\beta_0}/{\beta_1} = \emptyset \implies \beta_0 = \emptyset$

zero at origin

$$G(s) = \frac{\beta_1 s}{s^2 + \alpha_1 s + \alpha_0}$$

Then $y_{ss}(t) = G(\emptyset) = \emptyset \Leftarrow$ steady-state is zero response contains only transient terms

In fact, $y(t)$ is the *impulse* response of

$$G_1(s) = \frac{\beta_1}{s^2 + \alpha_1 s + \alpha_0}$$

Effect of Zeros

Step response of

$$G(s) = \frac{\beta_1 s + \beta_0}{s^2 + \alpha_1 s + \alpha_0}$$

Zero at $z_1 = -\beta_0/\beta_1$

3 important effects:

1. "Input absorbing" property
2. Transient suppression
3. Transient amplification

Both? Yes!

Depending on system

② Transient Suppression

Suppose $s^2+\alpha_1 s+\alpha_0=(s-p_1)(s-p_2)$ p_1, p_2 real

So
$$G(s)=\frac{\beta_1(s-z_1)}{(s-p_1)(s-p_2)}$$

Suppose $z_1 \approx p_1$, i.e $|z_1-p_1|=\varepsilon<<1$

We Know $y(t)=G(\phi)+A_1e^{p_1t}+A_2e^{p_2t}$

where $A_1=\left[(s-p_1)Y(s)\right]_{s=p_1}=\dfrac{\beta_1\overbrace{(p_1-z_1)}^{\sim\,<<1}}{p_1(p_1-p_2)}$ is small

so, for sufficiently small ε, the e^{p_1t} term in transient is negligible, and response is equivalent to a <u>1^{st} order</u> system with single pole p_2

Pole-zero Cancellation

Algebraically, if $z_1 \approx p_1$

$$G(s) = \frac{\beta_1 \cancel{(s-z_1)}}{\cancel{(s-p_1)}(s-p_2)} \approx \frac{\beta_1}{(s-p_2)}$$

Usually, if

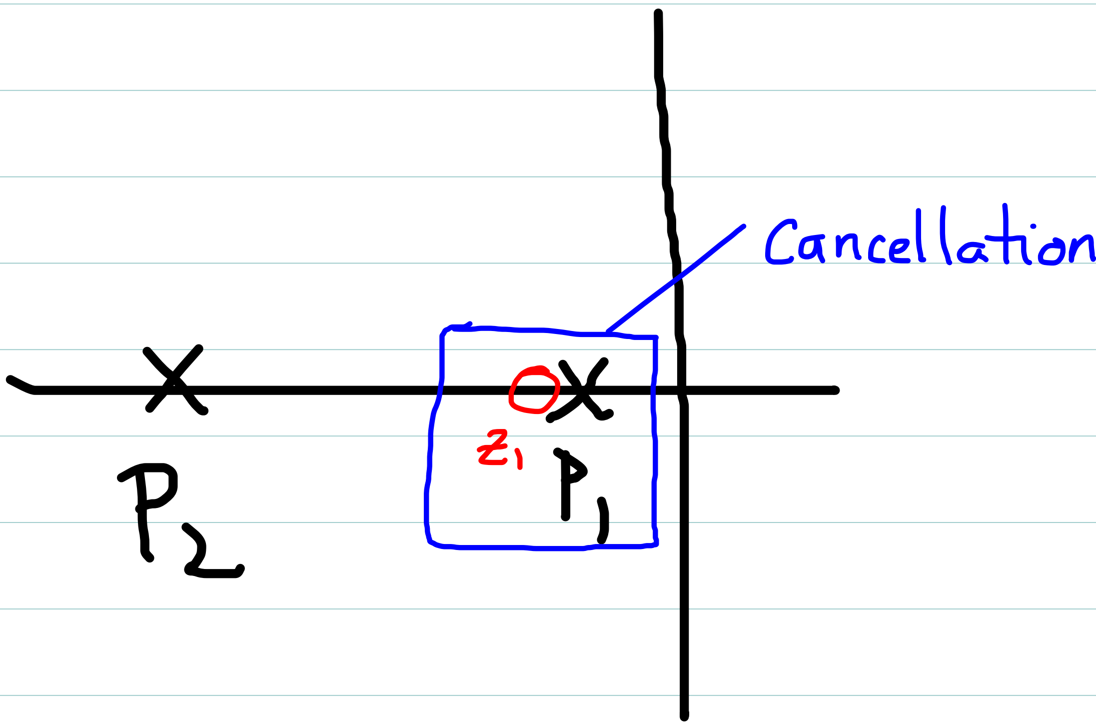
$$0.9 \leq \left|\frac{z_1}{p_1}\right| \leq 1.1$$

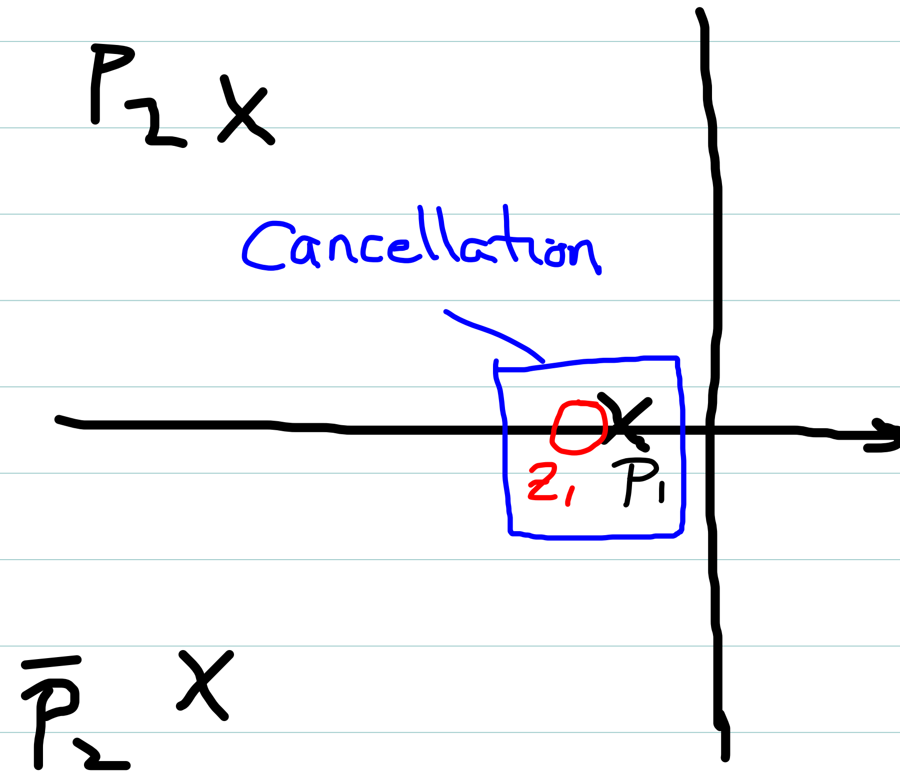
i.e. zero location within 10% of pole location

this is a good approximation

Cancellation and Dominance

Pole-zero cancellations can change dominance calculation





"fast" pole becomes dominant

2^{nd} order poles become dominant

Cancellation is never exact!

⇒ z_1, p_1 come from different coefs. in diff'l eq'n.

⇒ These coefs come from physical properties of system whose values are not known precisely.

⇒ Cancellation should always be considered approx.

⇒ If p_1 is stable, it is a good approximation to cancel it

$$A_1 e^{p_1 t} \propto \varepsilon e^{p_1 t}$$

This term starts small, and gets smaller as t increases

<u>But</u>

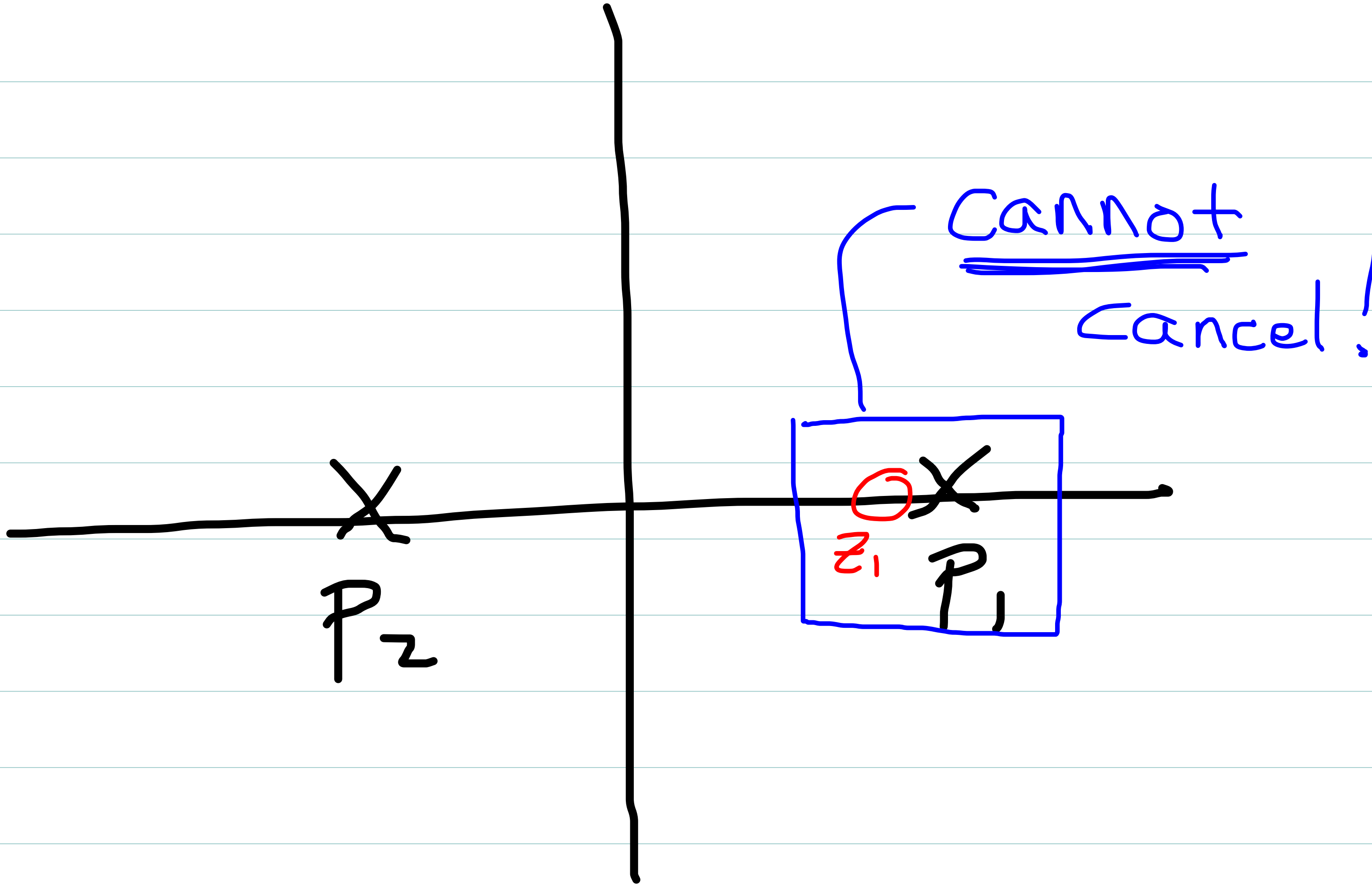
Suppose p_1 Not stable: $p_1 > 0$

then $A_1 e^{p_1 t} \propto \varepsilon e^{p_1 t}$

may start small, but increases w/o bound as t increases

Term will diverge to ∞, regardless how small ε is!

Pole-zero cancellation can <u>Never</u> be performed in RHP


Cannot
cancel!
z_1
P_1
P_2

Moreover...

Generally, if ICs on $y(t)$ are not all zero

$$Y(s) = G(s)U(s) + \frac{C(s)}{r(s)}$$

Nonzero

Will contribute terms to $y(t)$ which contain unstable mode <u>even if</u> this mode "cancels" in $G(s)$

<u>Moral</u>: Can <u>never</u> "cancel" an unstable mode

!!!

Effects of zeros on step response

$$G(s) = \frac{\beta_1 s + \beta_0}{s^2 + \alpha_1 s + \alpha_0}, \quad \text{zero at } z_1 = \frac{-\beta_0}{\beta_1}$$

① Input absorbtion (if $\beta_0 = \phi \Rightarrow z_1 = \phi$)

② Transient suppression via pole-zero cancellation

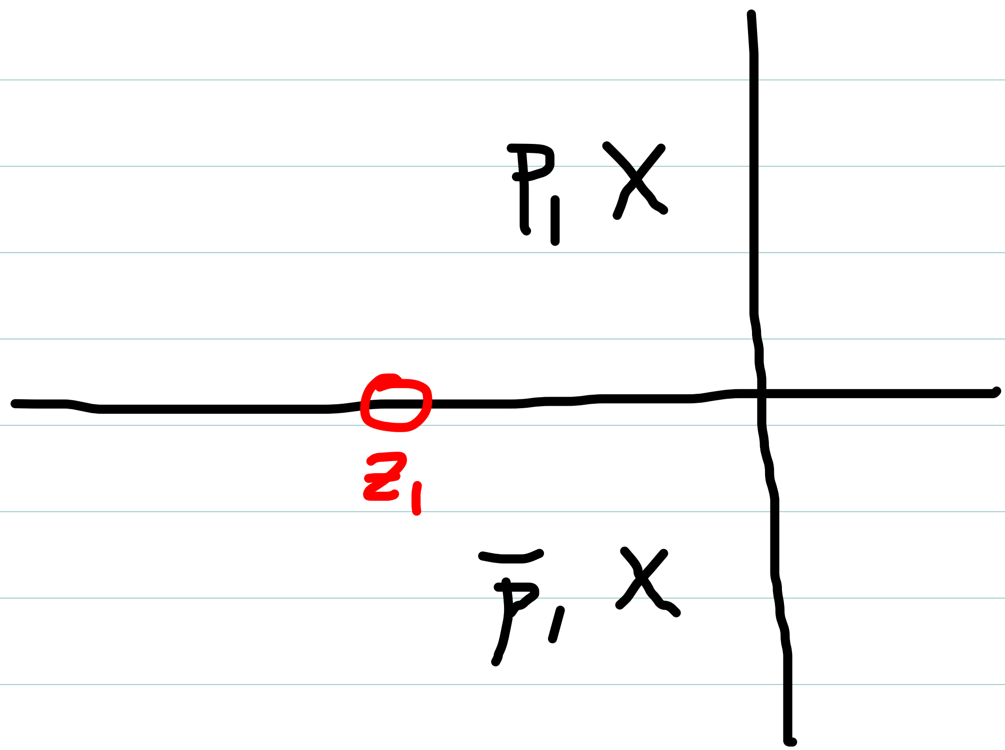
$\Rightarrow$ if $s^2 + \alpha_1 s + \alpha_0 = (s - p_1)(s - p_2)$; p_1, p_2 real
and $z_1 \approx p_1$ (or p_2)

③ Transient amplification $\Rightarrow$ examine this now.

③ Transient Amplification

Now suppose $s^2+\alpha_1 s+\alpha_0 = (s-p_1)(s-\bar{p}_1)$

$p_1 = \sigma + j\omega_d, \quad \omega_d \neq 0$



Pole-zero cancellation cannot occur here

What is the effect of the zero?

$$Y(s) = \frac{\beta_1 s + \beta_0}{s(s-p_1)(s-\bar{p}_1)} = \frac{\beta_1 s}{s(s-p_1)(s-\bar{p}_1)} + \frac{\beta_0}{s(s-p_1)(s-\bar{p}_1)}$$

$$= \left[\left(\frac{\beta_1}{\beta_0}\right)s\right]\left[\frac{\beta_0}{s(s-p_1)(s-\bar{p}_1)}\right] + \left[\frac{\beta_0}{s(s-p_1)(s-\bar{p}_1)}\right]$$

Let

$$Y_1(s) = \left[\frac{\beta_0}{s(s-p_1)(s-\bar{p}_1)}\right]$$

So

$$Y(s) = \left(\frac{\beta_1}{\beta_0}\right)\left[s\,Y_1(s)\right] + Y_1(s)$$

$$\Rightarrow \boxed{y(t) = \left(\frac{\beta_1}{\beta_0}\right)\dot{y}_1(t) + y_1(t)} \quad , \quad y_1(t) = \mathcal{L}^{-1}\{Y_1(s)\}$$

Note: $y_1(t)$ is ideal 2^{nd} order step response

$$y(t) = \left(\frac{\beta_1}{\beta_0}\right) \dot{y}_1(t) + y_1(t)$$

or equivalently:

$$\boxed{y(t) = y_1(t) - \left(\frac{1}{z_1}\right) \dot{y}_1(t)} \quad \left(z_1 = -\beta_0/\beta_1\right)$$

Where $y_1(t)$ is the "ideal" (no zero) step response

The total response $y(t)$ is the sum of the ideal response, and a fraction of the <u>derivative</u> of this response.

Suppose 1st $z_1 < 0$ (LHP zero)

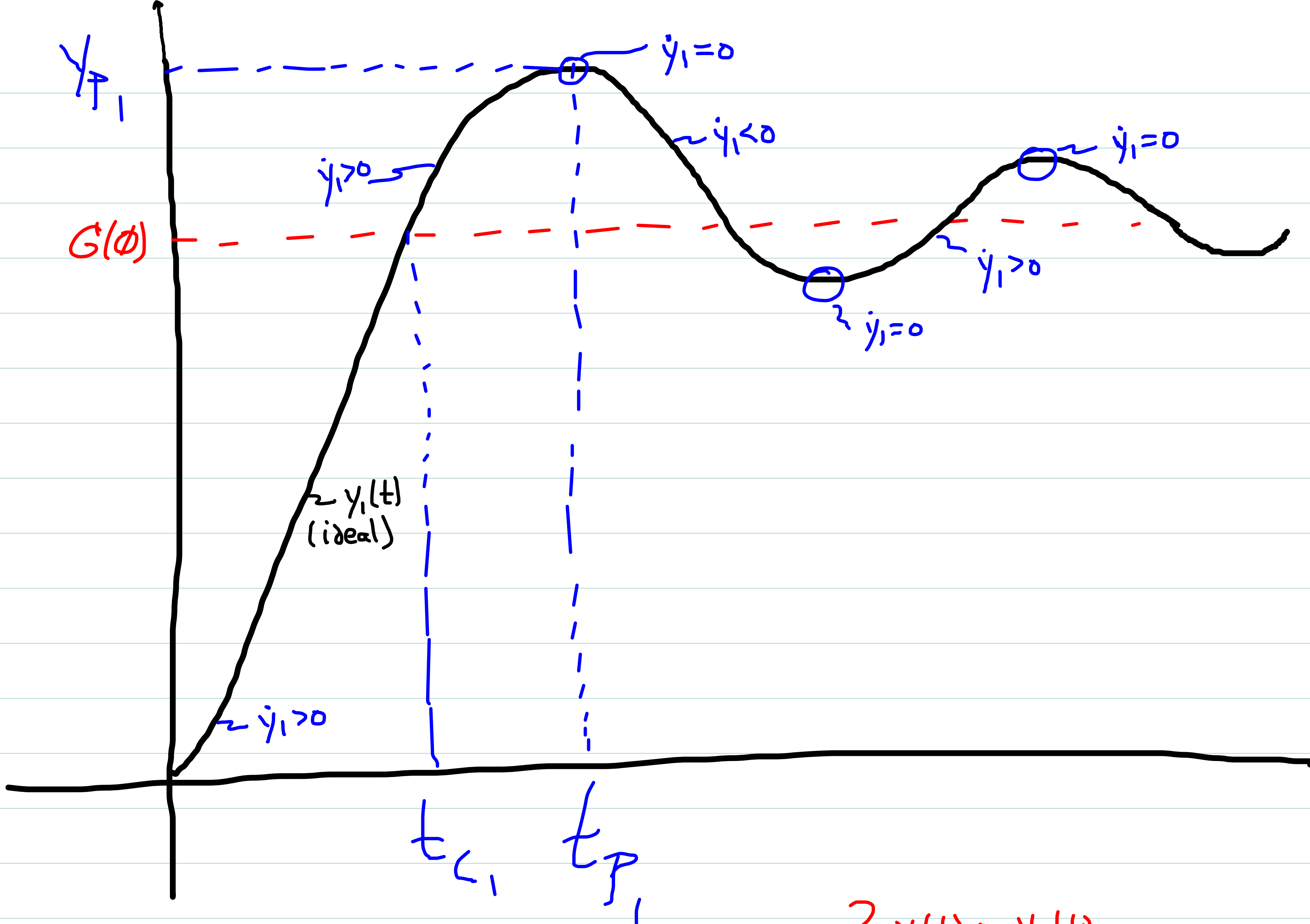
then $z_1 < 0$ and $\left(-\frac{1}{z_1}\right) > 0$ so we can write

$$y(t) = y_1(t) + \left(\frac{1}{|z_1|}\right)\dot{y}_1(t)$$

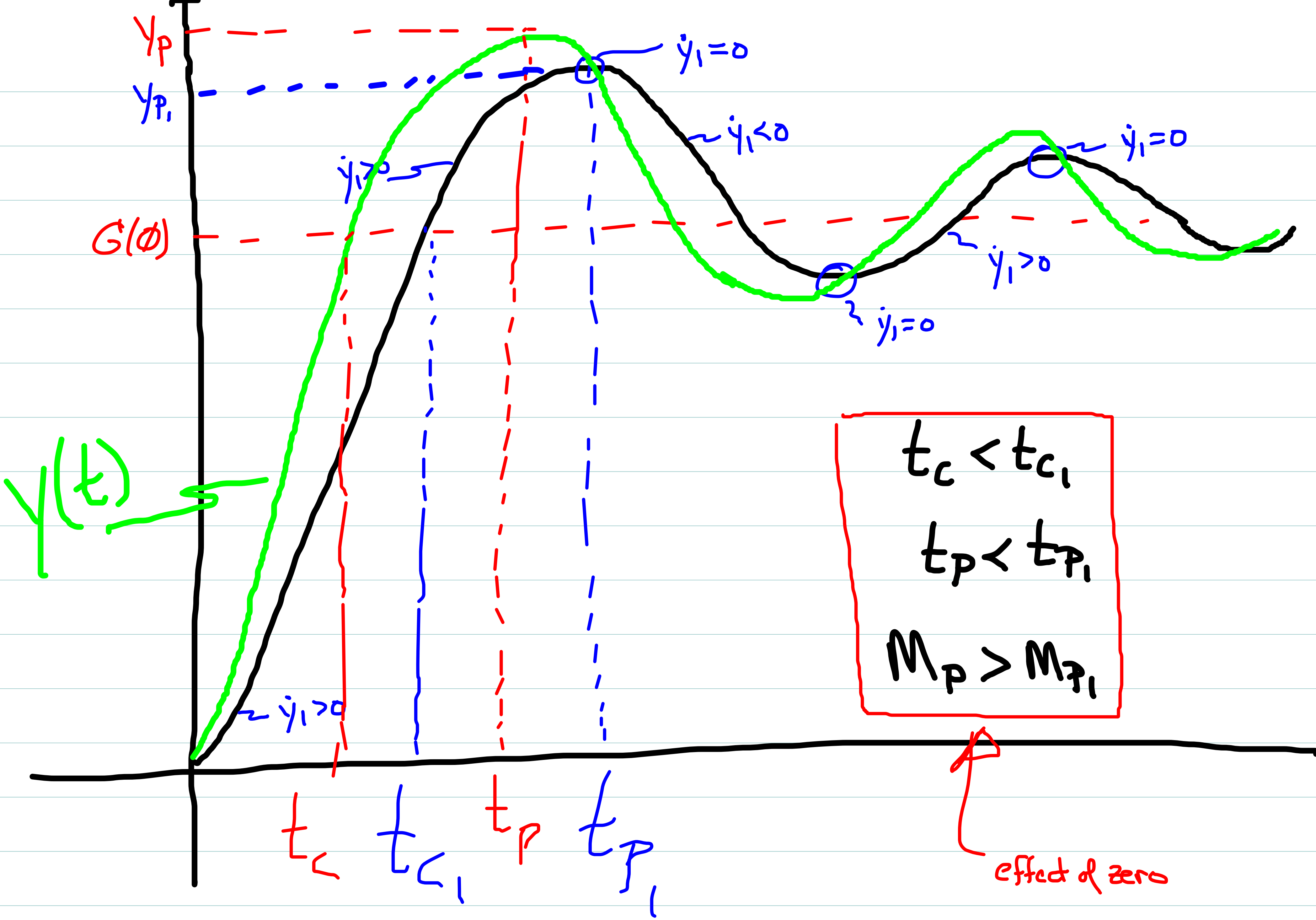
Derivative <u>adds</u> to total response. To understand effect of this, must examine behavior of $\dot{y}_1(t)$

Note that $\dot{y}_1(t) \to \emptyset$ As $t \to \infty$, so the steady-state of the new response will be the <u>Same</u> as the ideal response

$$y_{ss} = G(\emptyset)$$



Note: $\dot{y}_1(t) > \phi$ for all $\phi \le t < t_{p_1}$ } $y(t) > y_1(t)$ in this region



y_p
y_{p_1}
$G(\emptyset)$
$y(t)$
$\dot{y}_1 = 0$
$\dot{y}_1 < 0$
$\dot{y}_1 > 0$
$\dot{y}_1 = 0$
$\dot{y}_1 > 0$
$\dot{y}_1 = 0$
$\dot{y}_1 > 0$
t_c
t_{c_1}
t_p
t_{p_1}
$t_c < t_{c_1}$
$t_p < t_{p_1}$
$M_p > M_{p_1}$
effect of zero

$$y(t) = \left(\frac{\beta_1}{\beta_0}\right)\dot{y}_1(t) + y_1(t)$$

or equivalently:

$$\boxed{y(t) = y_1(t) - \left(\frac{1}{z_1}\right)\dot{y}_1(t)} \quad \left(z_1 = -\beta_0/\beta_1\right)$$

Where $y_1(t)$ is the "ideal" (no zero) step response

The total response $y(t)$ is the sum of the ideal response, and a fraction of the <u>derivative</u> of this response.

Suppose 1^{st} $z_1 < 0$ (LHP zero)

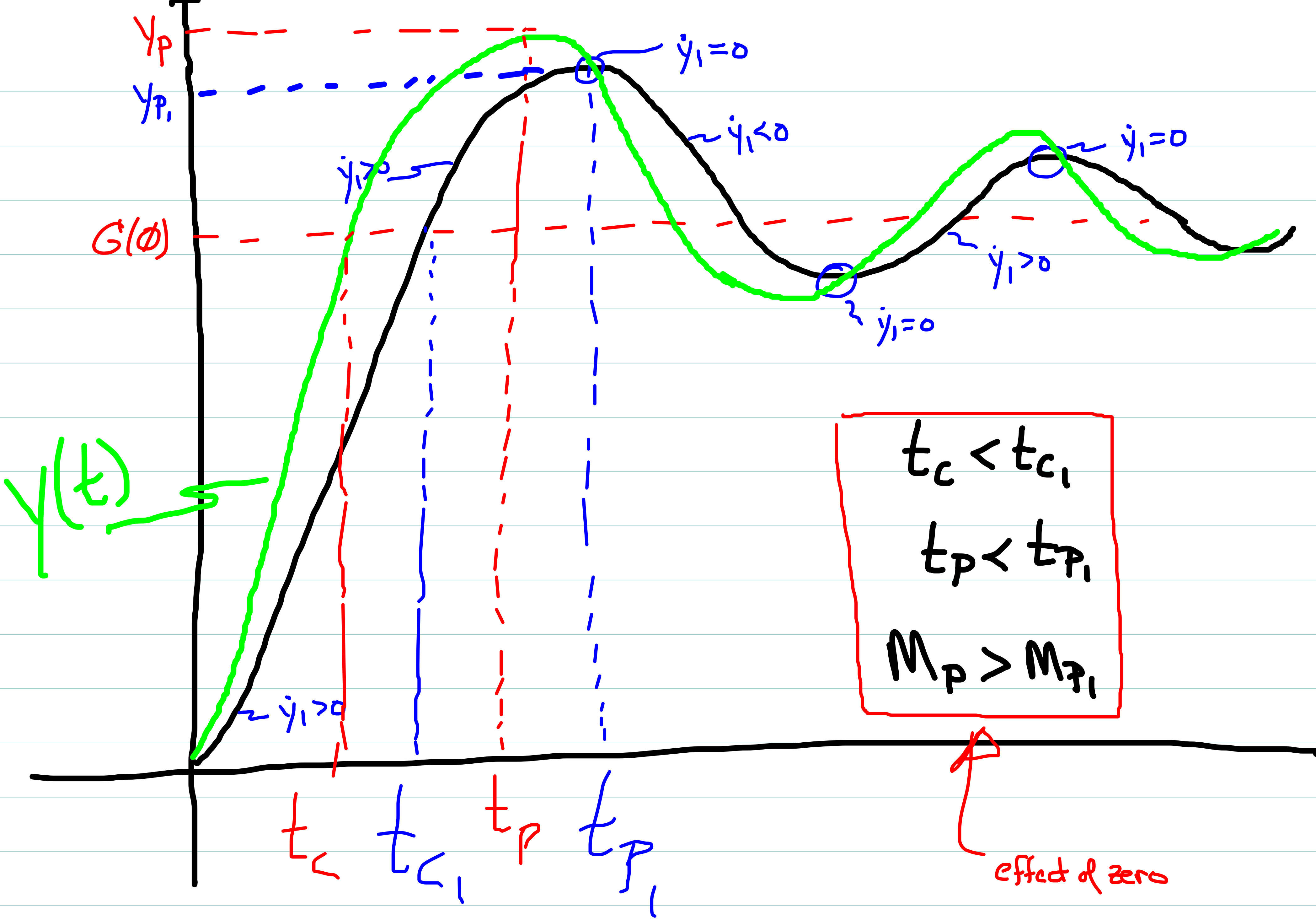
then $z_1 < 0$ and $\left(-\frac{1}{z_1}\right) > 0$ so we can write

$$y(t) = y_1(t) + \left(\frac{1}{|z_1|}\right) \dot{y}_1(t)$$

Derivative <u>adds</u> to total response. To understand effect of this, must examine behavior of $\dot{y}_1(t)$

Note that $\dot{y}_1(t) \to \phi$ As $t \to \infty$, so the steady-state of the new response will be the <u>Same</u> as the ideal response

$$y_{ss} = G(\phi)$$



y_p
y_{p_1}
$\dot{y}_1 = 0$
$\dot{y}_1 < 0$
$\dot{y}_1 > 0$
$\dot{y}_1 = 0$
$G(\emptyset)$
$\dot{y}_1 > 0$
$\dot{y}_1 = 0$
$y(t)$
$\dot{y}_1 > 0$
$t_c < t_{c_1}$
$t_p < t_{p_1}$
$M_p > M_{p_1}$
t_c
t_{c_1}
t_p
t_{p_1}
effect of zero

Summary of observations

A LHP zero changes a 2^{nd} order step response by:

$\Rightarrow$ Increasing overshoot y_p and M_p

$\Rightarrow$ decreasing t_c and t_p

In a sense, system "responds" faster (crosses y_{ss} more quickly), but price is greater overshoot!

$\Rightarrow$ Note: tricky to quantify exact changes to t_c, t_p, y_p based on z_1

$\Rightarrow$ However, note change from "ideal" response is proportional to $\frac{1}{|z_1|}$

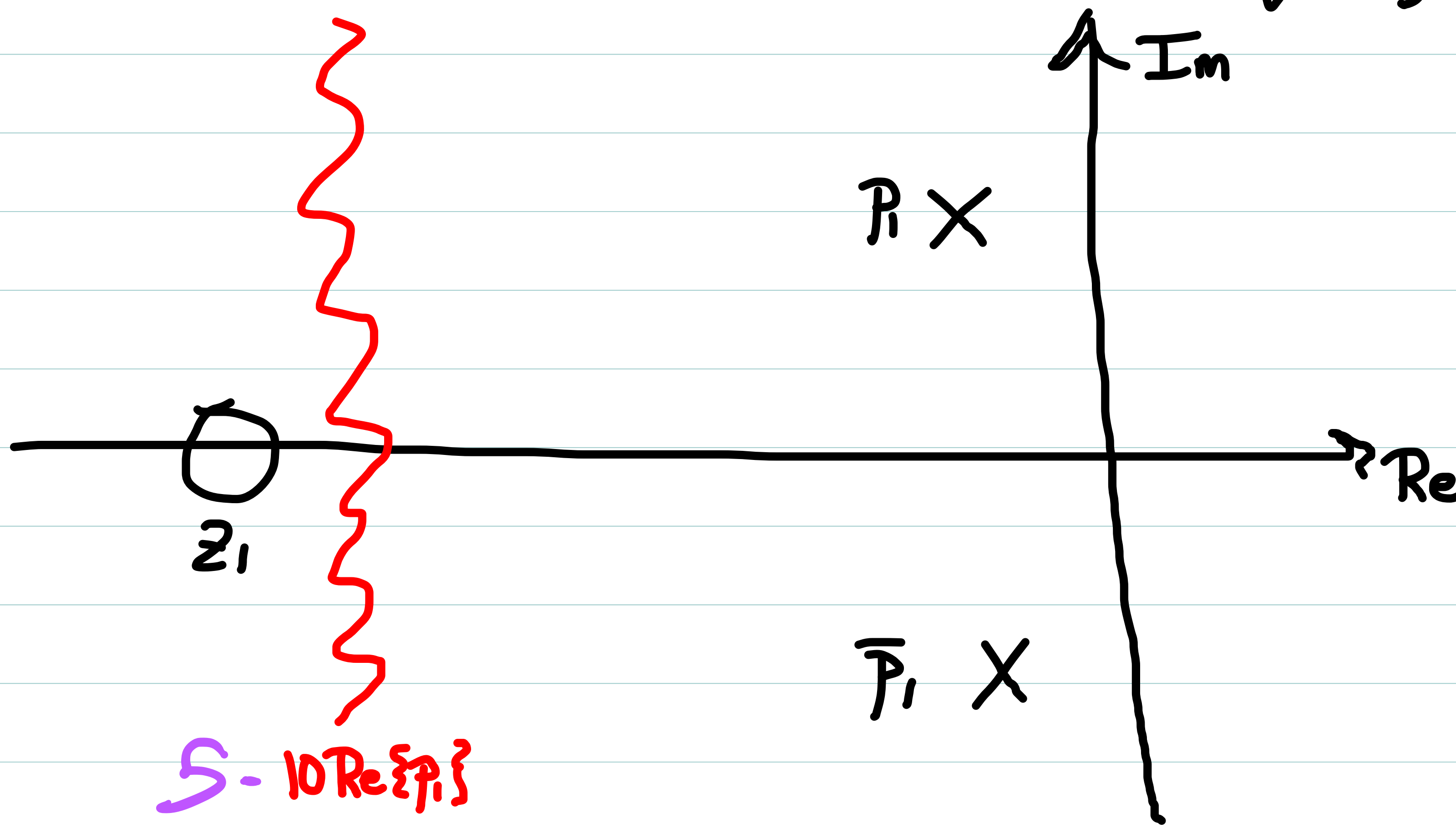
$\Rightarrow$ The further z_1 is from imag axis, the smaller the effect

Rule of Thumb

Effect of zero in this case is negligible if

$$|z_1| > \cancel{10}^{\,5}\,|\mathrm{Re}\{p_1\}|$$

i.e. zero is 10 times further into LHP than complex poles.



Question

$\Rightarrow$ A zero increases (amplifies) the overshoot of a 2^{nd} order system with $\xi < 1$ (complex poles).

$\Rightarrow$ Can it actually create overshoot in a system with 2 real poles ($\xi \geq 1$)?

$\Rightarrow$

Question

$\Rightarrow$ A zero increases (amplifies) the overshoot of a 2^{nd} order system with $\zeta < 1$ (complex poles).

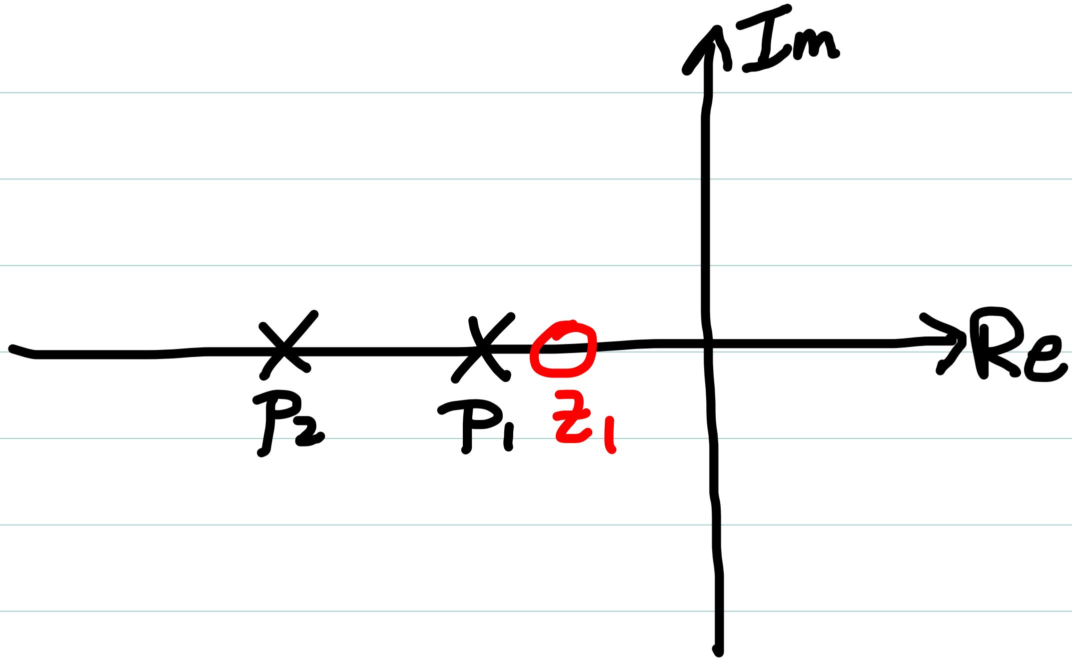
$\Rightarrow$ Can it actually <u>create</u> overshoot in a system with 2 real poles ($\zeta \geq 1$)?

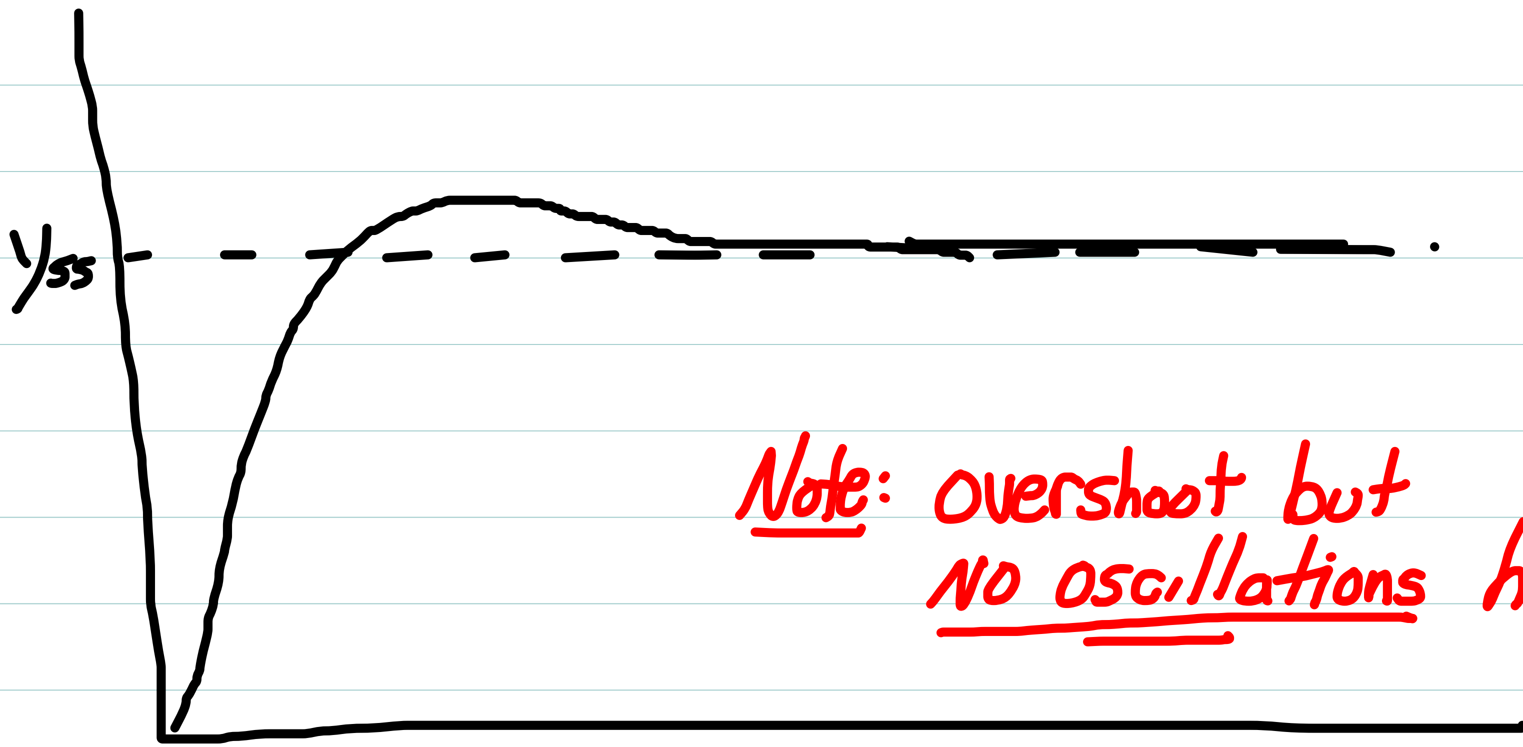
$\Rightarrow$ Yes!

$\Rightarrow$ With 2 real poles p_1 and p_2, $y_p > y_{ss}$ if

$$|z_1| < \min(|p_1|, |p_2|)$$

i.e. if zero is closer to imag axis than ~~either~~ both of the two poles.





Note: overshoot but NO oscillations here

Back to 2^{nd} order ($\zeta < 1$ case)

Suppose $z_1 > \emptyset$, i.e. z_1 in RHP, then

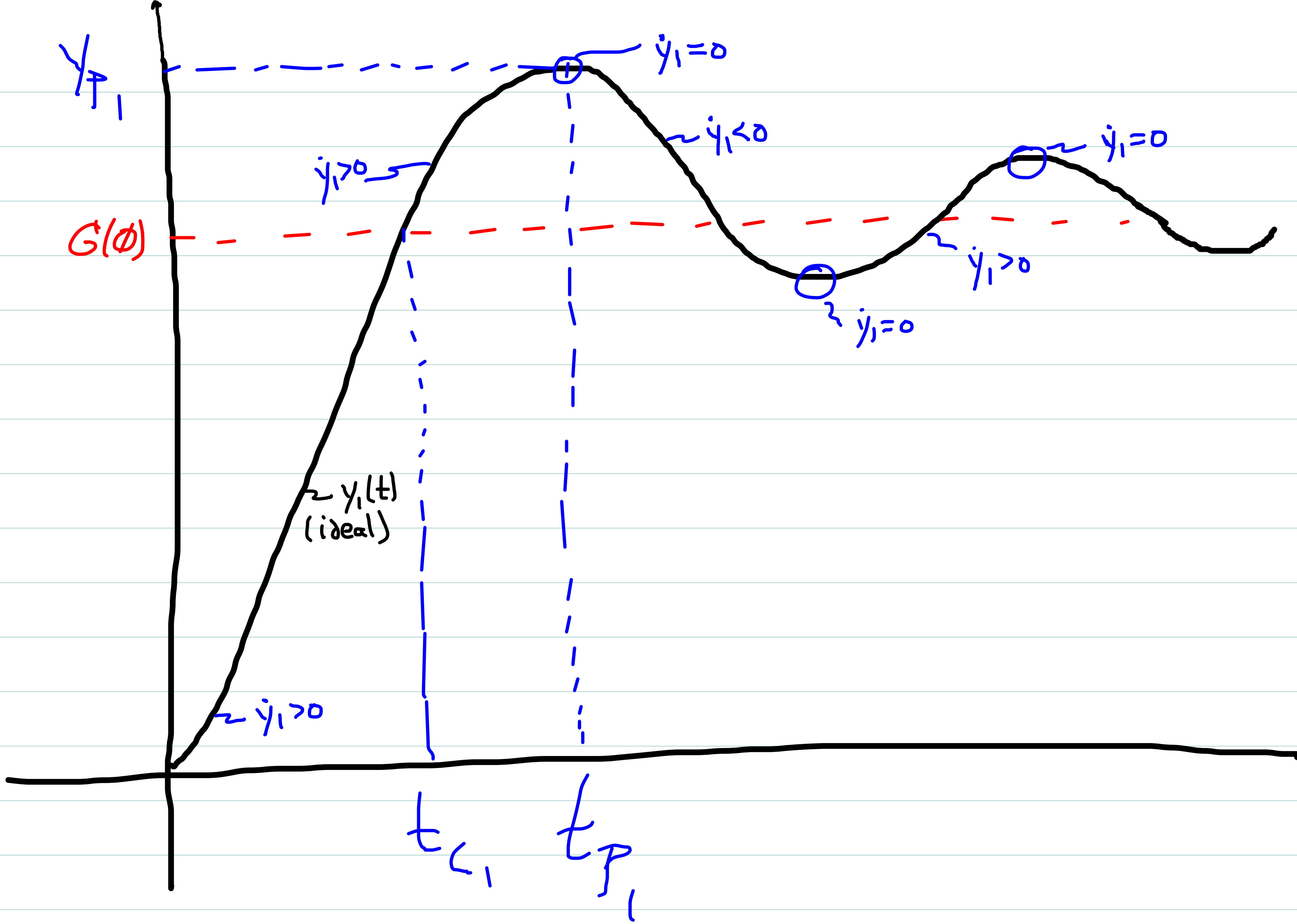
$$y(t) = y_1(t) - \left(\frac{1}{z_1}\right)\dot{y}_1(t)$$

So we are <u>subtracting</u> the derivative from the ideal response.

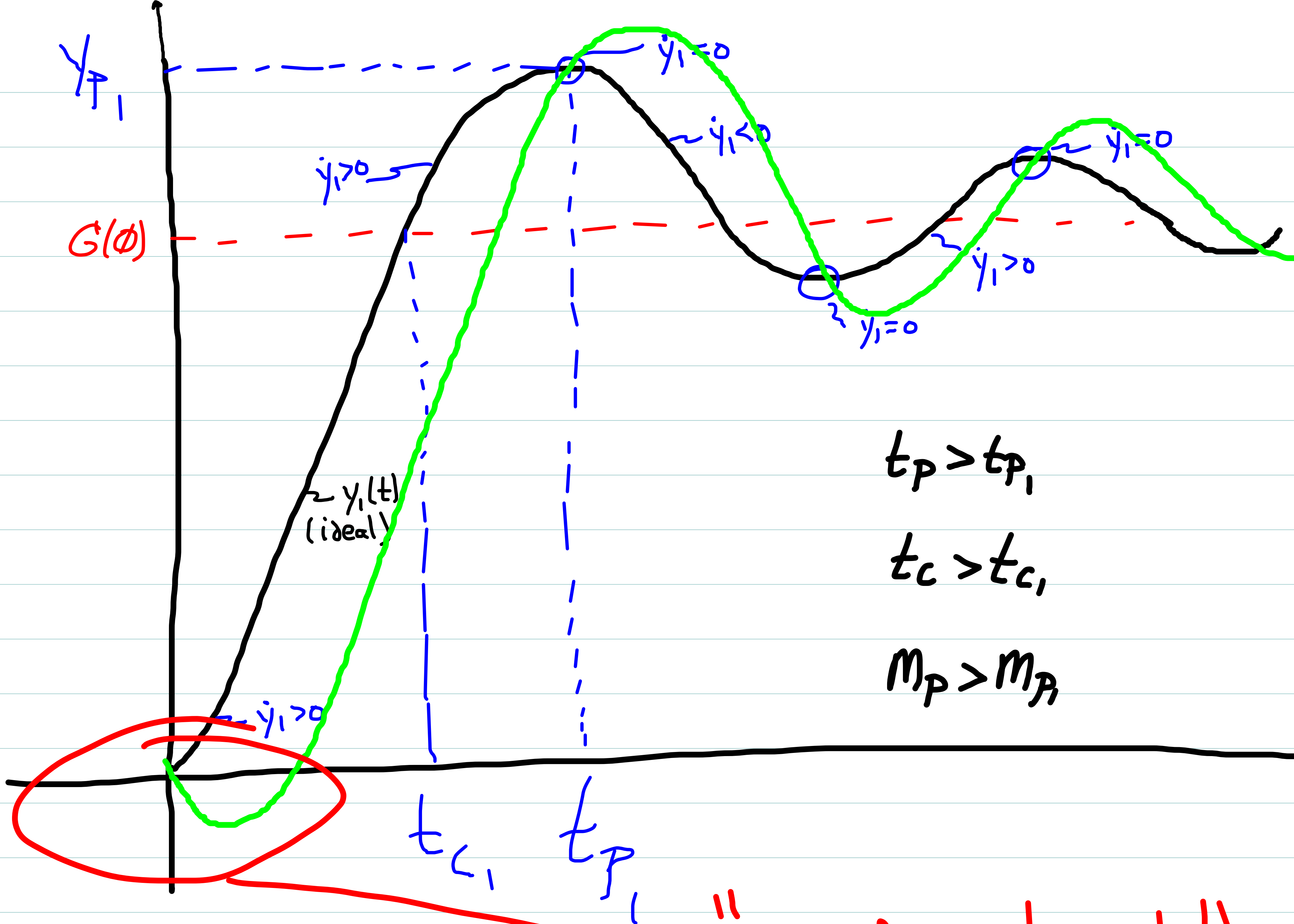
But recall $\dot{y}_1(t) \geq \emptyset$ for $\emptyset < t < t_p$,

And $y_1(t) \approx \emptyset$ for t close to zero

Seems to suggest that $y(t)$ may become <u>negative</u> for times near $t = \emptyset \ldots$?



y_{p_1}
$\dot{y}_1 = 0$
$\dot{y}_1 < 0$
$\dot{y}_1 = 0$
$\dot{y}_1 > 0$
$G(\emptyset)$
$\dot{y}_1 > 0$
$\dot{y}_1 = 0$
$y_1(t)$ (ideal)
$\dot{y}_1 > 0$
t_{c_1}
t_{p_1}



$t_P > t_{P_1}$

$t_c > t_{c_1}$

$M_P > M_{P_1}$

"<u>undershoot</u>"

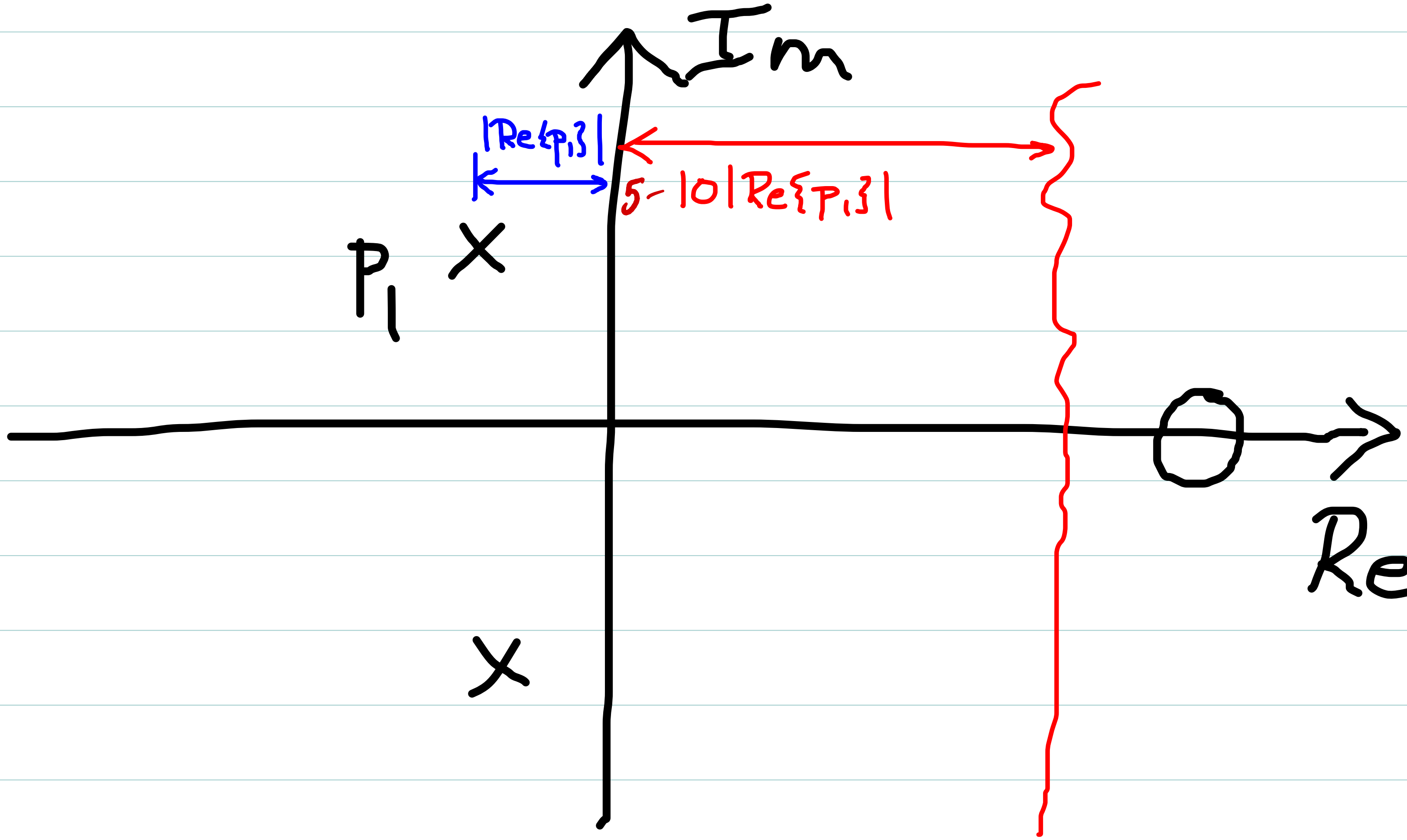
Observations (RHP zero)

- ⇒ Again, the peak response is greater
- ⇒ However, t_c and t_p have increased
- ⇒ Appearence of a new feature: "undershoot"
- ⇒ Response initially heads "in wrong direction" before ultimately returning to the same steady-state
- ⇒ Such behavior is NOT UNstable
- ⇒ It is, however, very tricky to design controllers for such systems.

Effect is still proportional to $\frac{1}{|z_1|}$

hence diminishes as z_1 moves further from Im axis

Again negligible if $|z_1| > 10|Re\{p_1\}|$



Effect on settling time

How a zero, either LHP or RHP, affects t_s is difficult to predict.

$\Rightarrow$ Often, but not always, t_s is longer with zero due to increased amplitude of transient oscillations

$\Rightarrow$ No hard and fast rule here

$\Rightarrow$ Primary effect is increased overshoot and:

- reduction of t_c, t_p (LHP)
- undershoot, with increase of t_c, t_p (RHP)

Performance Specifications

⇒ Step inputs representative for many desired behaviors

- move to new pointing angle (spacraft)
- Move to new altitude or heading (aircraft)

⇒ Required performance often specified as <u>upper limits</u> on acceptable t_s, M_p

- System must settle quickly enough, and not overshoot too much.

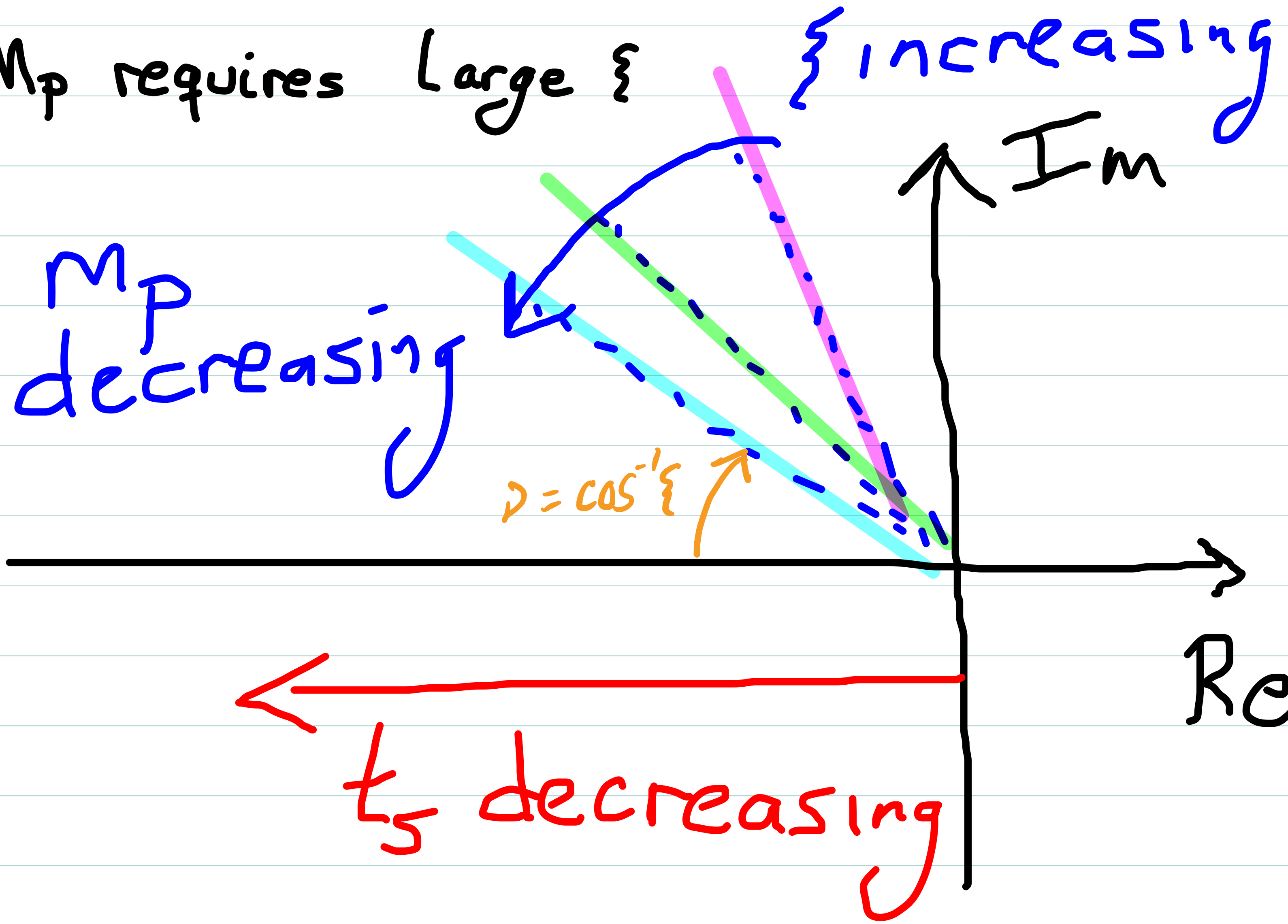
⇒ <u>Recall</u>:

- t_s inversely proportional to $|Re\{p_i\}|$
- M_p a decreasing function of ζ

$$t_s \approx \frac{4}{|Re\{p_1\}|}, \quad \boxed{M_p = \exp\left[\frac{-\zeta\pi}{\sqrt{1-\zeta^2}}\right]}$$

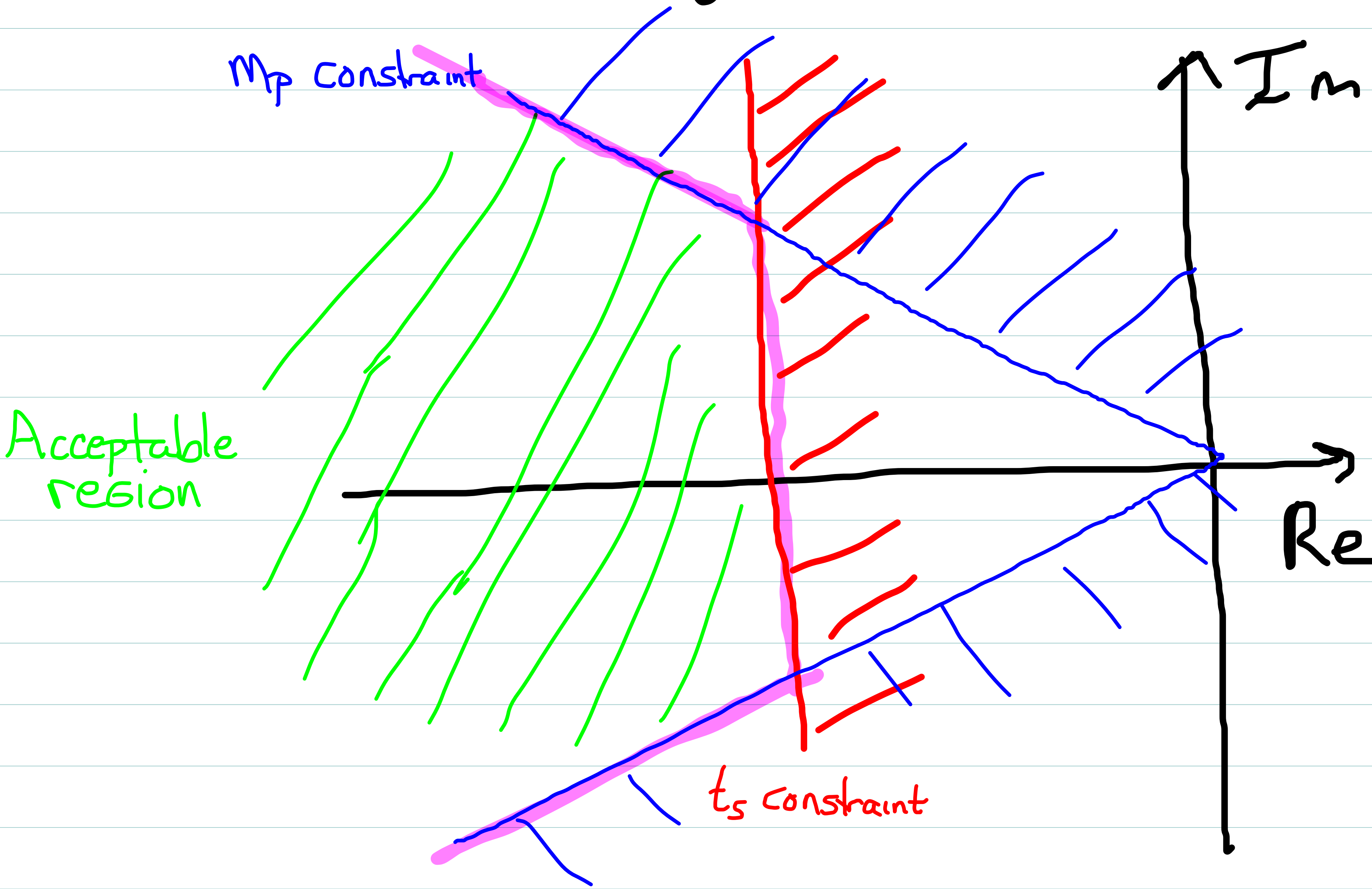
$\Rightarrow$ small t_s requires large $|Re\{p_1\}|$

$\Rightarrow$ small M_p requires large ζ

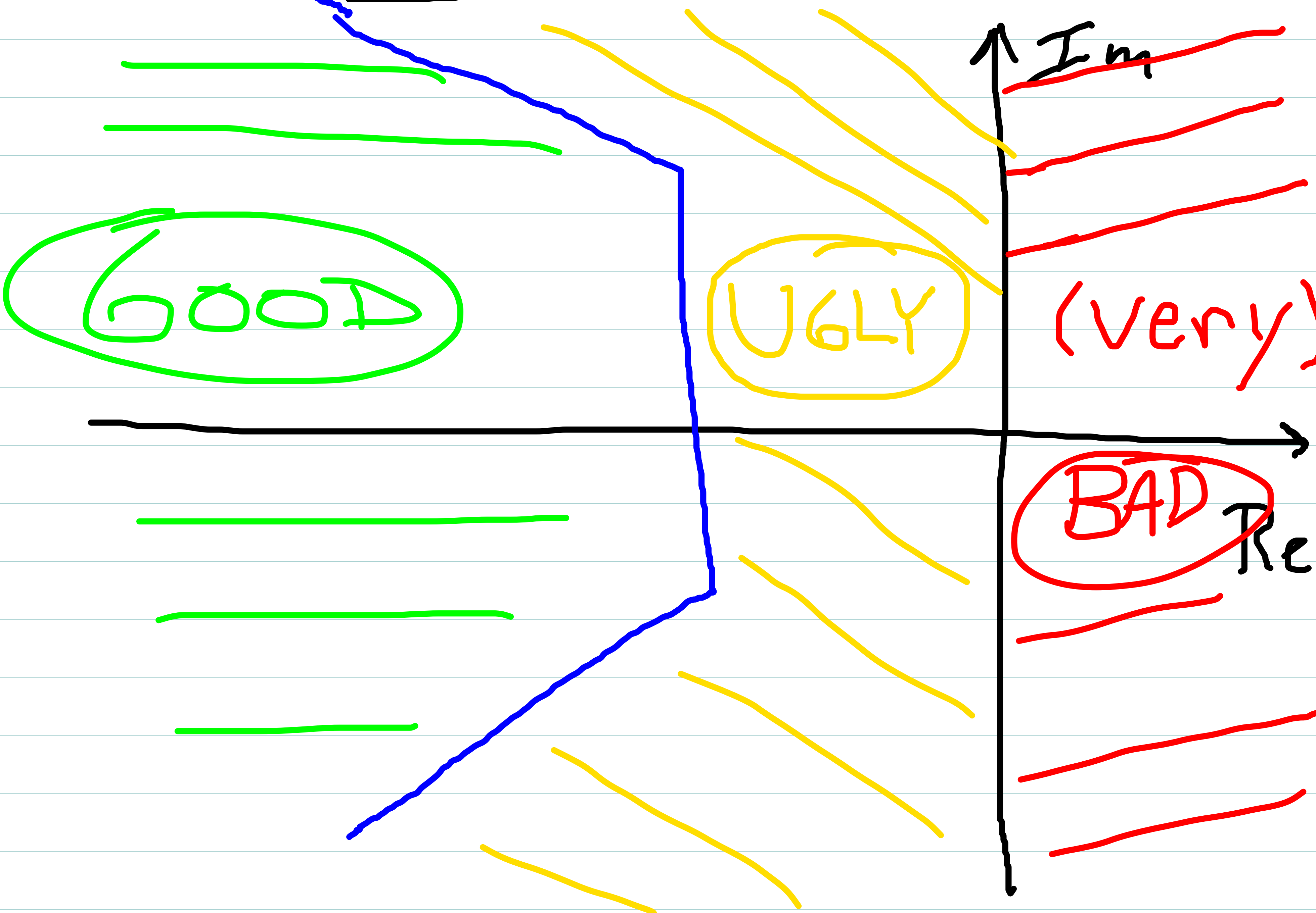


$\Rightarrow$ Upper bound on t_s gives lower bound on $|Re\{p_i\}|$

$\Rightarrow$ Upper bound on M_p gives lower bound on ξ



Desireable Pole Locations



=> "Good" poles satisfy all transient performance constraints (upper bounds on t_s, M_p)

=> "Bad" poles are unstable

=> "Ugly" poles are stable, but have too much overshoot or take too long to settle.

=> Most aerospace system have natural dynamics which are "bad" or "ugly"

=> Goal of control is to make these systems "good"

Feedback "moves" poles

⇒ Already seen this on previous homeworks.

⇒ But it can be tricky!

Suppose $u(t) = K(y_d(t) - y(t))$

If system is modeled with $Y(s) = G(s)U(s)$

where $G(s) = \dfrac{\beta_1 s + \beta_0}{s^2 + \alpha_1 s + \alpha_0}$

Then poles are moved to roots of

$$r_{CL}(s) = s^2 + (\alpha_1 + K\beta_1)s + (\alpha_0 + K\beta_0)$$

⇒ Tricky to predict movement of poles for all possible values of $K, \alpha_0, \alpha_1, \beta_0, \beta_1$

⇒ Even more complicated for $G(s)$ with additional poles and/or zeros

⇒ Need a more systematic tool to predict effectiveness of a control strategy.

⇒ One approach is BASED on a more careful analysis of the behavior of $G(j\omega)$.

Sinusoidal Response

Here we wish to understand the properties of the steady-state response of a $\boxed{\text{stable}}$ system when $\boxed{u(t)=\sin\omega t.}$

Note: our focus is shifting (temporarily) away from the transient response

$$u(t)=\sin\omega t \longrightarrow \boxed{G(s)} \longrightarrow y_{ss}(t)=?$$

$$\Longrightarrow y_{ss}(t) = \operatorname{Im}\{G(j\omega)e^{j\omega t}\}$$

Sinusoidal Response

Here we wish to understand the properties of the steady-state response of a $\boxed{\text{stable}}$ system when $\boxed{u(t)=\sin\omega t.}$

Note: our focus is shifting (temporarily) away from the transient response

$$u(t)=\sin\omega t \longrightarrow \boxed{G(s)} \longrightarrow y_{ss}(t)=\,?$$

Of course, we've already solved this problem:

$$u(t) = \sin\omega t = \mathrm{Im}\{e^{j\omega t}\}$$

$$\Rightarrow y_f(t) = \mathrm{Im}\{G(j\omega)e^{j\omega t}\} = |G(j\omega)|\sin(\omega t + \measuredangle G(j\omega))$$

Then $y(t) = y_f(t) + y_h(t)$

But if system is stable, $y_h(t) \to \phi$ as $t \to \infty$ for any set of initial cond'ns.

Hence $y_{tr}(t) = y_h(t)$ leaving us with

$$y_{ss}(t) = |G(j\omega)|\sin(\omega t + \measuredangle G(j\omega))$$

So:

$$u(t) = \sin \omega t \implies y_{ss}(t) = |G(j\omega)| \sin(\omega t + \measuredangle G(j\omega))$$

Note:

$y_{ss}(t)$ is sinusoidal at same frequency as $u(t)$

But:

Amplitude and phase of $y_{ss}(t)$ different.

Now, more generally suppose:

$$u(t) = B\sin(\omega t + \psi) = \mathrm{Im}\{U e^{j\omega t}\}, \quad U = Be^{j\psi}$$

then $$y_{ss}(t) = \mathrm{Im}\{G(j\omega) U e^{j\omega t}\}$$

$$= |G(j\omega)| \cdot |U| \sin(\omega t + \measuredangle G(j\omega) + \measuredangle U)$$

or $$y_{ss}(t) = |G(j\omega)| B \sin(\omega t + \measuredangle G(j\omega) + \psi)$$

Thus generally:

$$u(t) = B\sin(\omega t + \psi) \implies y_{ss}(t) = A\sin(\omega t + \varphi)$$

where: $A = |G(j\omega)|B$

$\varphi = \measuredangle G(j\omega) + \psi$

<u>Define:</u>

<u>Amplitude ratio:</u> A/B (ratio of output ampl. to input ampl.

<u>Phase shift:</u> $\varphi - \psi$ (Diff. between output and input phase)

<u>Then note:</u>

$$\boxed{\begin{aligned} A/B &= |G(j\omega)| \\ \varphi - \psi &= \measuredangle G(j\omega) \end{aligned}}$$

So generally

$|G(j\omega)|$ quantifies the ratio between output and input amplitude

$\measuredangle G(j\omega)$ quantifies the change in phase of output compared to input

Note: these are frequency dependent

i.e. the amplitude ratio and phase shift depend on frequency of input.

Very useful to quantify this dependence!

Example

$$G(s) = \frac{3}{s+2}$$

$$|G(j\omega)| = \frac{3}{\sqrt{\omega^2+4}} \qquad \measuredangle G(j\omega) = -\tan^{-1}\left(\frac{\omega}{2}\right)$$

$\omega = 1/2 \Rightarrow |G(j/2)| = 3/\sqrt{4.25} \approx 1.46$

$\measuredangle G(j/2) = -\tan^{-1}(1/4) = -.245$ rad or $-14.04°$

$\omega = 2 \Rightarrow |G(2j)| = 3/\sqrt{8} \approx 1.06$

$\measuredangle G(2j) = -\tan^{-1}(1) = -\pi/4 = -45°$

$\omega = 20 \Rightarrow |G(20j)| = 3/\sqrt{404} = 0.15$

$\measuredangle G(20j) = -\tan^{-1}(10) = -1.47 \approx -84.3°$

$\Rightarrow$ Want to learn to predict these changes based on ZPK structure of $G(s)$

$\Rightarrow$ Useful also to visualize graphically

$\Rightarrow$ Three methods

1. Plot $|G(j\omega)|$ and $\measuredangle G(j\omega)$ vs. $\omega \geq 0$ (2 plots)

2. Plot $G(j\omega)$ as ω varies from 0 to ∞ as points in complex plane.

3. Plot $|G(j\omega)|$ vs. $\measuredangle G(j\omega)$ for $0 \leq \omega < \infty$

$\Rightarrow$ Want to learn to predict these changes based on ZPK structure of $G(s)$

$\Rightarrow$ Useful also to visualize graphically

$\Rightarrow$ Three methods

1. Plot $|G(j\omega)|$ and $\measuredangle G(j\omega)$ vs. $\omega \geq 0$ (2 plots) "Bode diagrams"

2. Plot $G(j\omega)$ as ω varies from 0 to ∞ as points in complex plane. "Polar diagram"

3. Plot $|G(j\omega)|$ vs. $\measuredangle G(j\omega)$ for $0 \leq \omega < \infty$ "Nichols chart"

Bode is most fundamental, start there

$\Rightarrow$ want to see behavior for large range of $\omega \geq \emptyset$

$\Rightarrow |G(j\omega)|$ will vary enormously in size

$\Rightarrow$ Use logarithmic scales for plots.

$\Rightarrow$ Horizontal Axis on Bode diagram is freq on a log scale

$\Rightarrow$ equally spaced divisions on this scale are factors of 10 apart.

$\Rightarrow$ We call one of these divisions a "decade"

$$\left.\begin{array}{l} 1/10 \to 1 \\ 2 \to 20 \end{array}\right\} \text{one decade} \qquad \left.\begin{array}{l} 1/10 \to 10 \\ 2 \to 200 \end{array}\right\} \text{two decades}$$

Decibels

$|G(j\omega)|$ is shown on Bode diagrams in special units called <u>decibels</u>.

<u>Def'n</u>: For any real number $X \geq 0$

$$X_{db} = 20 \log X$$

Conversely

$$X = 10^{(X_{db}/20)}$$

Example (from above): $X = 1.46 \Rightarrow X_{db} = 3.25$

$$X = 1.06 \Rightarrow X_{dB} = 0.51$$

$$X = 0.15 \Rightarrow X_{dB} = -16.5$$

Common Shorthand

$$X = 0.15 = -16.5\,dB$$

Note common conversions

X	X (dB)
.01	-40
.1	-20
1	0
10	20
100	40

Important ⟹

Zero on dB axis means magnitude of 1 !!

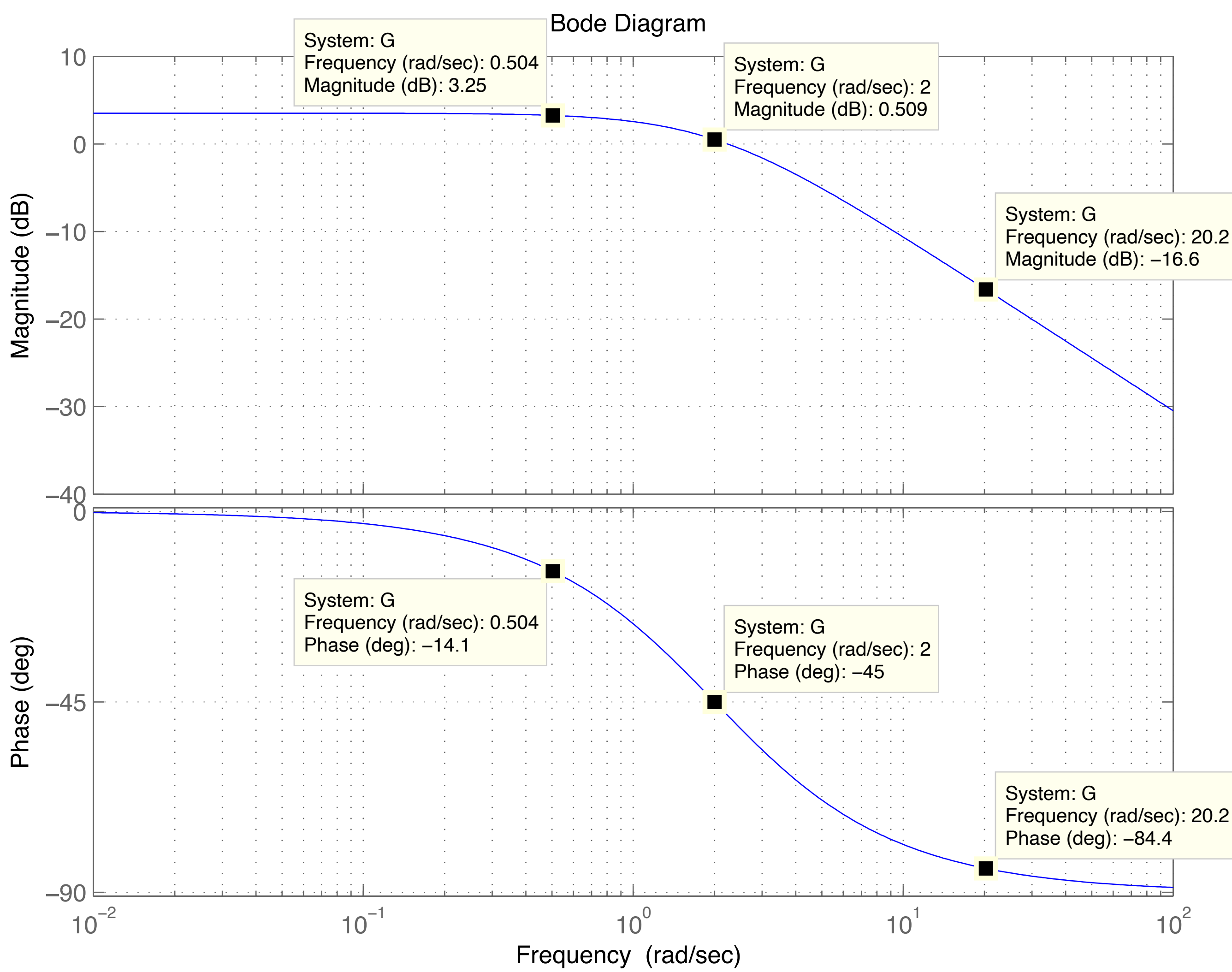
Bode diagrams show

1. $|G(j\omega)|$ in dB vs ω on a log scale
2. $\measuredangle G(j\omega)$ in deg " " "

See example

Note: there are no negative frequencies on a Bode diagram!

The left limit of the horizontal scale corresponds to $\omega \to \emptyset$!


Bode Diagram
System: G
Frequency (rad/sec): 0.504
Magnitude (dB): 3.25
System: G
Frequency (rad/sec): 2
Magnitude (dB): 0.509
System: G
Frequency (rad/sec): 20.2
Magnitude (dB): −16.6
Magnitude (dB)
10
0
−10
−20
−30
−40
System: G
Frequency (rad/sec): 0.504
Phase (deg): −14.1
System: G
Frequency (rad/sec): 2
Phase (deg): −45
System: G
Frequency (rad/sec): 20.2
Phase (deg): −84.4
Phase (deg)
0
−45
−90
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
Frequency (rad/sec)

Recap: Frequency Response Analysis

$$u(t) = B\sin(\omega t + \psi) \Rightarrow y_{ss}(t) = A\sin(\omega t + \varphi)$$

$$A = B|G(j\omega)|, \quad \varphi = \measuredangle G(j\omega) + \psi$$

Bode diagrams: show

$|G(j\omega)|$ (dB) vs. ω (log scale) "Magnitude diagram"

$\measuredangle G(j\omega)$ (deg) vs. ω (log scale) "Phase diagram"

Want to learn to rapidly predict the shapes of these diagrams from the ZPK structure of transfer function $G(s)$

How?

Will Show:

1. Effect of each pole p_K and zero z_i is concentrated in a narrow band of frequencies near $\omega = |p_K|$ (or $|z_i|$, as appropriate)

   $\Rightarrow$ remember: $\omega \geq 0$ on Bode diagrams. There are no negative frequencies shown!

2. Effect of individual poles/zeros on *total* Bode diagrams are <u>additive</u>

"Bode form" of transfer function

ZPK form:

$$G(s) = K\left[\frac{\prod_{i=1}^{m}(s-z_i)}{\prod_{k=1}^{n}(s-p_k)}\right]$$

Bode form:

$$G(s) = K_B \frac{\prod_{i=1}^{m}(1-s/z_i)}{s^N \prod_{k=N+1}^{n}(1-s/p_k)}$$

N = # of poles at origin "Type" of system

K_B = "Bode gain"; note $N=\emptyset \Rightarrow K_B = G(\emptyset)$

Bode and ZPK forms are two different ways of writing the <u>same</u> transfer function

<u>Example</u>:

$$G(s) = \frac{5(s+2)}{s(s+3)(s+4)} \qquad \text{(ZPK)}$$

$$\text{(Bode)} \qquad = \left(\frac{5}{6}\right)\left[\frac{(1+s/2)}{s(1+s/3)(1+s/4)}\right]$$

Here $N=1$ and $K_B = 5/6$

<u>Algebraically equivalent</u> to ZPK form.

i.e. both are the <u>same</u> TF

So:

$$G(j\omega) = K_B \left[\frac{\prod_{i=1}^{m} \left(1 - j\omega/z_i\right)}{(j\omega)^N \prod_{k=N+1} \left(1 - j\omega/p_k\right)} \right]$$

for any real $\omega \geq 0$, $G(j\omega)$ is complex and so are each individual factor (except K_B, which is real)

recall for any $s_1, s_2 \in \mathbb{C}$

$$\boxed{\begin{aligned} &\measuredangle(s_1 s_2) = \measuredangle s_1 + \measuredangle s_2 \\ &\measuredangle\left(\frac{s_1}{s_2}\right) = \measuredangle s_1 - \measuredangle s_2 \\ &\measuredangle s_1^N = N \measuredangle s_1 \end{aligned}}$$

Thus:

$$\measuredangle G(j\omega) = \measuredangle K_B + \sum_{i=1}^{m} \measuredangle\left(1 - {}^{j\omega}/_{z_i}\right) - N\measuredangle(j\omega) - \sum_{K=N+1}^{n} \measuredangle\left(1 - {}^{j\omega}/_{p_K}\right)$$

Note: ① Each factor contributes additively

② Zeros add to angle, poles subtract

③ $\measuredangle K_B$ same for any ω:

$$\measuredangle K_B = \emptyset \;\; (K_B > \emptyset), \quad \measuredangle K_B = \pm 180° \;\; (K_B < \emptyset)$$

③ $\measuredangle(j\omega)$ is same for any $\omega \geq \emptyset$

$$\measuredangle(j\omega) = 90°$$

④ Changes to $\measuredangle G(j\omega)$ as ω varies depends on specific z_i and nonzero p_K.

What about Magnitudes?

Recall: for $s_1, s_2 \in \mathbb{C}$

$$|s_1 s_2| = |s_1|\,|s_2|$$

$$\left|\frac{s_1}{s_2}\right| = \frac{|s_1|}{|s_2|}$$

$$|s_1^N| = |s_1|^N$$

So:

$$|G(j\omega)| = |K_B| \frac{\prod_{i=1}^{m} \left|1 - j\omega/z_i\right|}{|j\omega|^N \prod_{k=N+1}^{n} \left|1 - j\omega/p_k\right|}$$

UGLY . . .

But Bode shows $|G(j\omega)|$ <u>in dB</u>

i.e. $20\log|G(j\omega)|$

<u>Now recall:</u>

$$\log(xy) = \log x + \log y$$

$$\log\left(\frac{x}{y}\right) = \log x - \log y$$

$$\log(x^N) = N\log x$$

Hence in dB:

$$|G(j\omega)|_{dB} = |K_B|_{dB} + \boxed{\sum_{i=1}^{m} \left|1 - {j\omega}/{z_i}\right|_{dB}} - N|j\omega|_{dB} \boxed{- \sum_{K=N+1}^{n} \left|1 - \frac{j\omega}{p_K}\right|_{dB}}$$

Notes:

1. Magnitudes in dB are additive for each factor
2. zeros add to magnitude, poles subtract
3. $|K_B|$ is constant for all ω, like with phase
4. $|j\omega|$ is Not constant, unlike phase.

So, we see effect of individual parts of $G(s)$ contribute additively to

$$\angle G(j\omega) \quad \text{and} \quad |G(j\omega)|_{dB}$$

Look at effect of individual factors

Look at how each $(1 - j\omega/z_i)$ or $(1 - j\omega/p_k)$

Changes with ω.

To simplify notation, we'll look at $(1 + j\omega\tau)$, where

$\tau = -1/z_i$ or $\tau = -1/p_k$ as appropriate

Then:

$$|1 + j\omega\tau| = \sqrt{1 + \omega^2\tau^2}$$

and

$$\measuredangle(1 + j\omega\tau) = \tan^{-1}\omega\tau$$

Study how these vary with ω

Consider first magnitude

$$|1+j\omega\tau| = \sqrt{1+(\omega\tau)^2} = \begin{cases} 1 & \text{if } \omega \ll 1/|\tau| \\ \sqrt{2} & \text{if } \omega = 1/|\tau| \\ \omega|\tau| & \text{if } \omega \gg 1/|\tau| \end{cases}$$

and thus:

$$|1+j\omega\tau|_{dB} = \begin{cases} \emptyset & \omega \ll 1/|\tau| & \text{"Low freq. Limit"} \\ 3 & \omega = 1/|\tau| & \\ 20\log\omega|\tau| & \omega \gg 1/|\tau| & \text{"high freq Limit"} \end{cases}$$

Look at 3rd case:

$$20\log\omega|\tau| = 20\left[\log\omega + \log|\tau|\right]$$

Note when $\omega = 1/|\tau|$, $\log\omega = -\log|\tau|$ & 3rd case evaluates to $\emptyset$.

Also:

in high freq limit $\omega \gg 1/|\tau|$

$$|1+j\omega\tau|_{dB} = 20\left[\log\omega + \log|\tau|\right]$$

Suppose we have two freqs, ω_1, ω_2 both $\gg 1/|\tau|$

with $\omega_2 = 10\omega_1$, then:

$$|1+j\omega_2\tau|_{dB} = |1+j(10\omega_1)\tau|_{dB}$$

$$= 20\left[\log(10\omega_1) + \log|\tau|\right]$$

$$= 20\left[\log\omega_1 + \log 10 + \log|\tau|\right]$$

$$= 20\left[\log\omega_1 + \log|\tau|\right] + 20$$

So

$$|1+j\omega_2\tau|_{dB} = |1+j\omega_1\tau|_{dB} + 20 \Leftarrow \text{+20 dB increase}$$

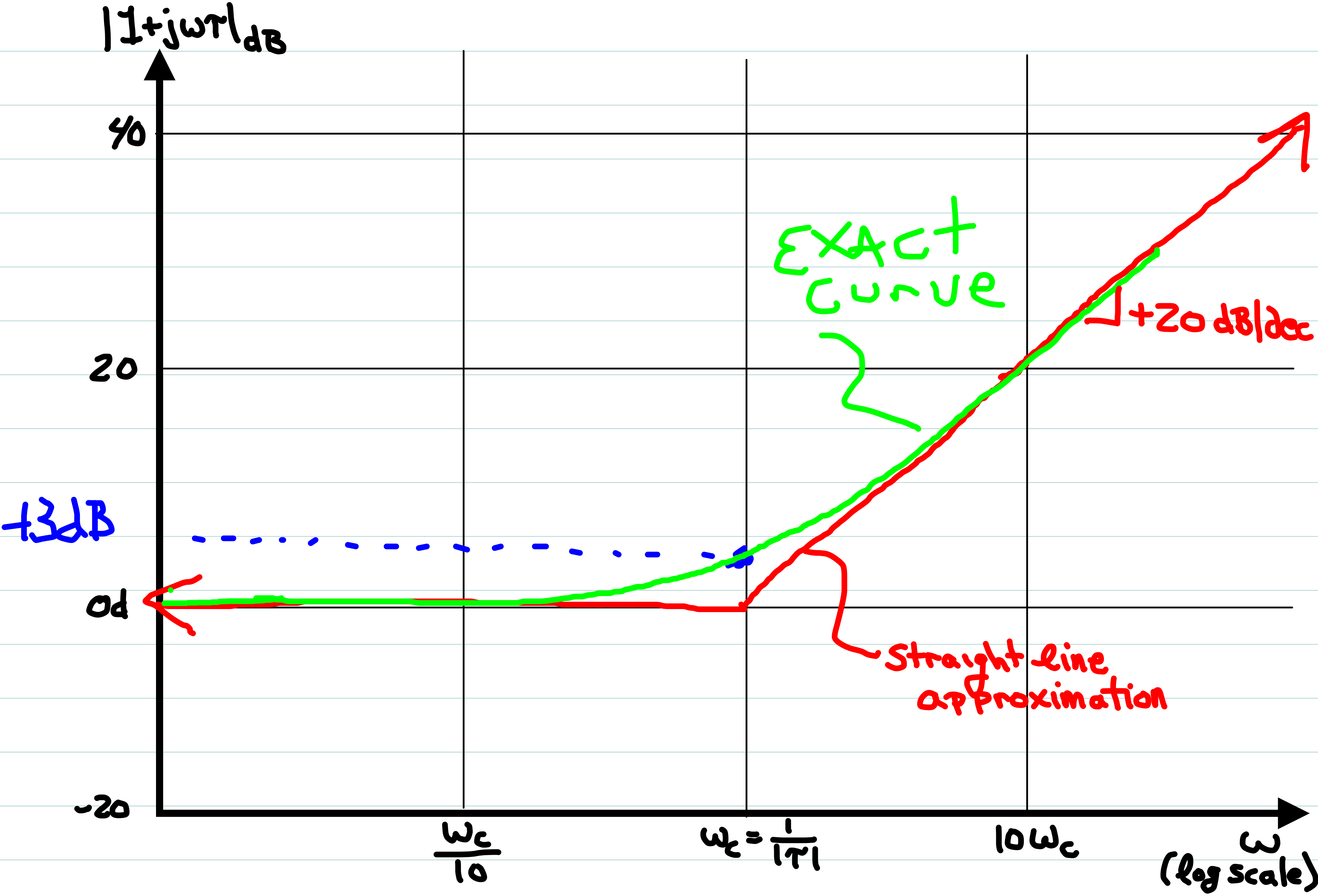
Hence:

in high frequency region $|1+j\omega T|_{dB}$ increases by 20dB for every factor of 10 increase in frequency (decade)

$\Rightarrow$ graph has a slope of 20dB/decade in high freq. region

$\Rightarrow$ Recall graph is constant at $\emptyset$ dB in low freq. region

$\Rightarrow$ The two limiting cases come together at the "corner frequency", $\omega_c = \frac{1}{|T|}$.



|1+jωτ|dB
40
20
+3dB
0dB
-20
EXACT curve
+20 dB/dec
straight line approximation
ωc/10
ωc = 1/|τ|
10ωc
ω (log scale)

Things to note:

- Graph changes slope by +20 dB/dec
- Think in terms of this slope change, not the total shape
- Recall $(1+j\omega\tau)$ is a generic representation of a factor of $G(s)$, either

$$\left(1-\frac{j\omega}{z_i}\right) \quad \text{or} \quad \left(1-\frac{j\omega}{p_K}\right)$$

i.e. $\tau = -1/z_i$ or $\tau = -1/p_K$

Thus the corner freq. $\omega_c = 1/|\tau| = |z_i|$ or $|p_K|$

$\boxed{\text{Corner freq is the absolute VALUE of a pole or zero of } G(s)}$

$\Rightarrow$ Because $|G(j\omega)|_{dB}$ is the sum of the effects of the individual terms $|1 - j\omega/z_i|_{dB}$ $|1 - j\omega/p_k|_{dB}$

Each pole or zero will create a "corner" on the complete graph

$\Rightarrow$ The total graph will have corners at every freq. corresponding to $|z_i|$ and $|p_k|$.

$\Rightarrow$ zeros add to overall $|G(j\omega)|_{dB}$ $\Rightarrow$ slope changes of $+20$ dB/dec at $\omega = |z_i|$, $i = 1 \ldots m$

$\Rightarrow$ Poles subtract from overall $|G(j\omega)|_{dB}$ $\Rightarrow$ Slope changes of -20 dB/dec at $\omega = |p_k|$.

Example #1

$$G(s) = (10s+1)(s/10+1)$$

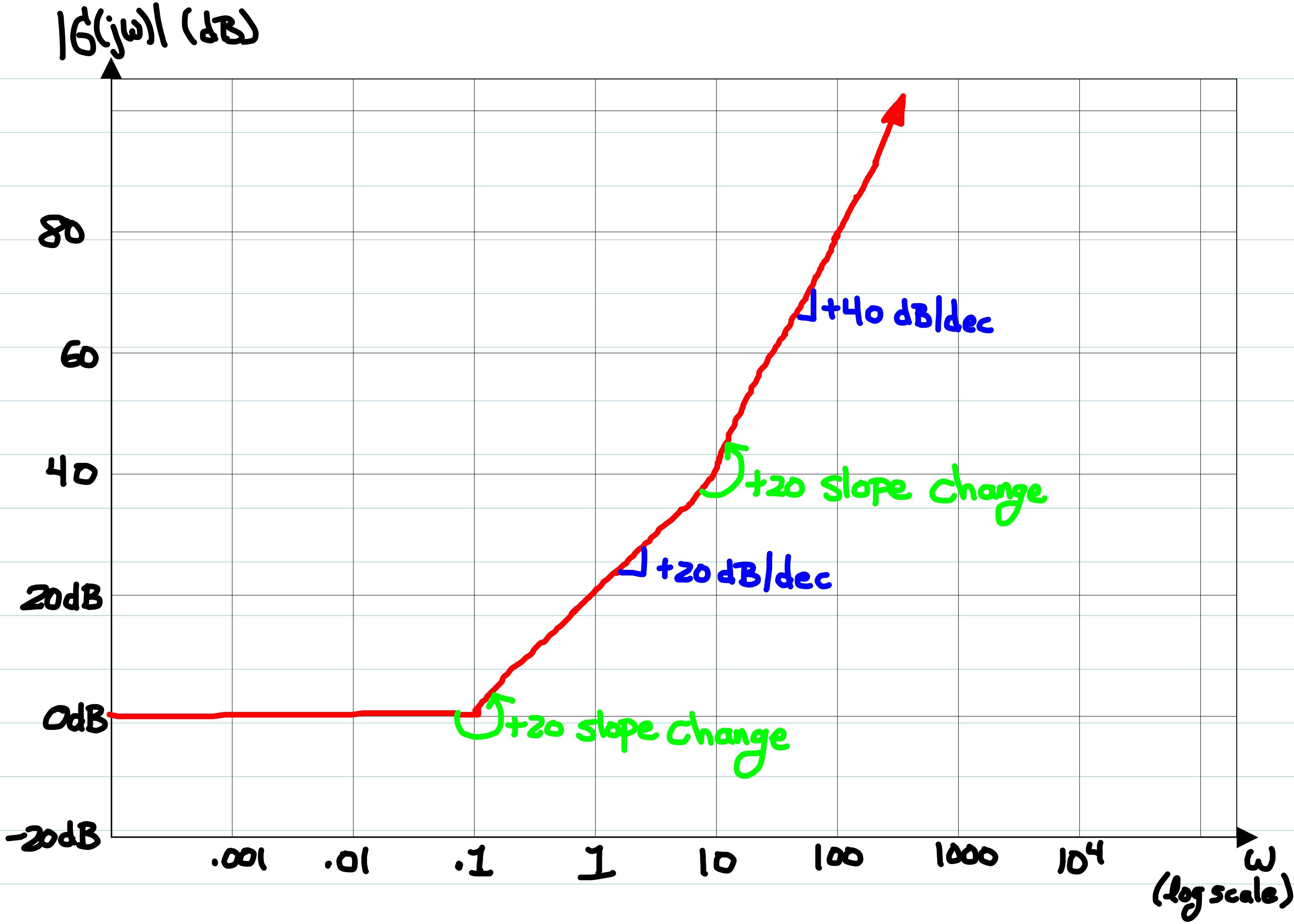
No poles; zeros at $z_1 = -10$, $z_2 = 1/10$

$|G(j\omega)|_{dB}$ will show +20 dB/dec changes at

$\omega = 1/10$ and $\omega = 10$

Below $\omega = 1/10$ the graph will be constant at 0 dB.

Graph bends up by +20 dB/dec at $\omega = 1/10$, and again at $\omega = 10$.



|G(jω)| (dB)
80
60
40
20dB
0dB
-20dB
.001
.01
.1
1
10
100
1000
10^4
ω
(log scale)
+20 slope change
+20 dB/dec
+20 slope change
+40 dB/dec

Example #2

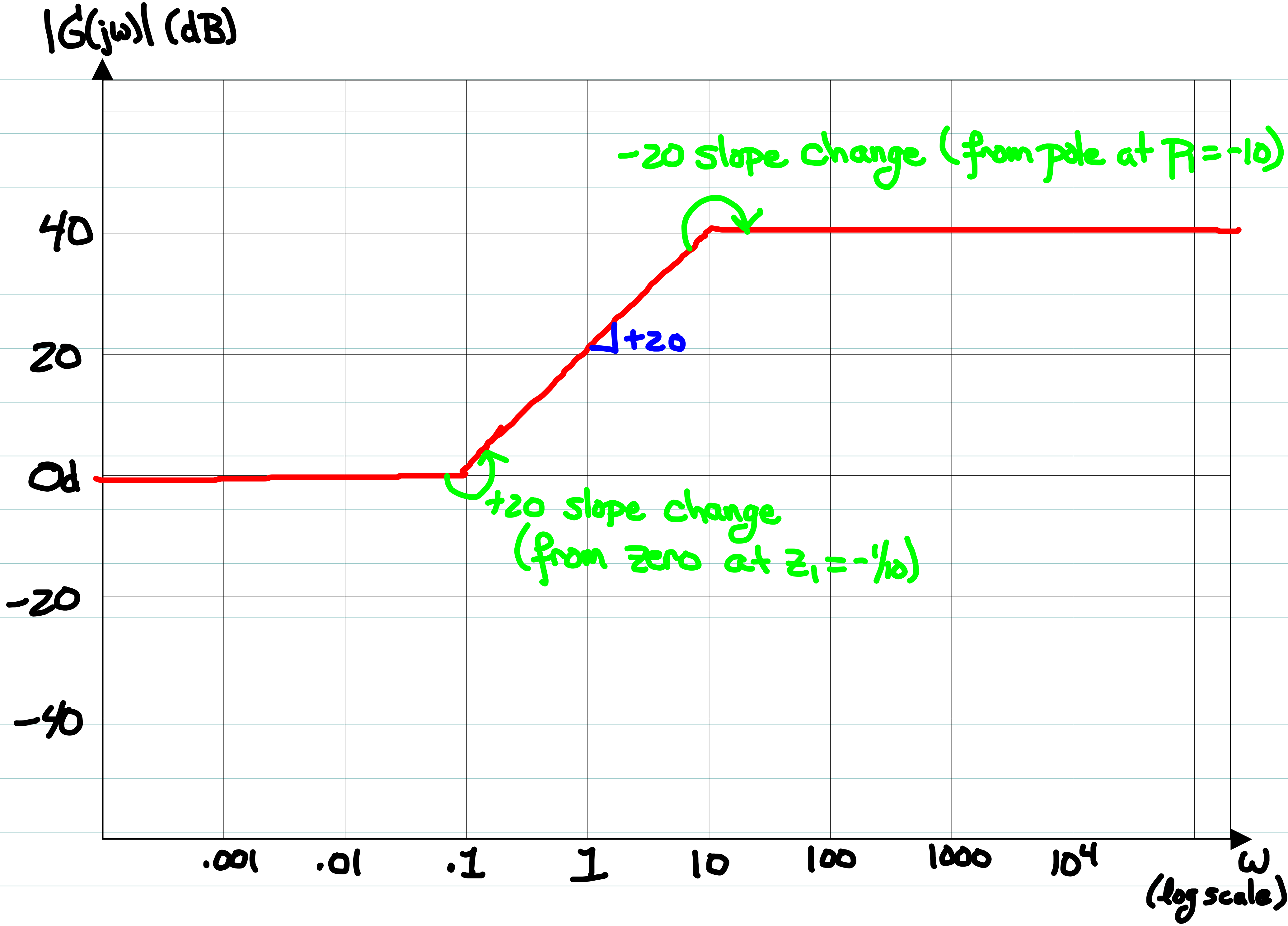
$$G(s) = \frac{(10s+1)}{(s/10+1)}$$

Zero at $z_1 = -1/10$, pole at $p_1 = -10$

Corners at $\omega = 1/10$ and $\omega = 10$ again

But now: at $\omega = 1/10$ slope increases by +20dB/dec

at $\omega = 10$ slope decreases by -20 dB/dec


$|G(j\omega)|$ (dB)
40
20
0d
-20
-40
-20 slope change (from pole at $P_1 = -10$)
+20
+20 slope change (from zero at $z_1 = -1/10$)
.001
.01
.1
1
10
100
1000
10^4
ω (log scale)

Gain effect is additive also, and <u>constant</u> for all ω:

$$|K_B(1+j\omega\tau)|_{dB} = |K_B|_{dB} + |1+j\omega\tau|_{dB}$$

$\Rightarrow$ entire graph shifts <u>up or down</u> by $|K_B|_{dB} = 20\log|K_B|$

shifts <u>up</u> if $|K_B|_{dB} > \emptyset$

Shifts <u>down</u> if $|K_B|_{dB} < \emptyset$

Gain effect is additive also, and <u>constant</u> for all ω:

$$|K_B(1+j\omega\tau)|_{dB} = |K_B|_{dB} + |1+j\omega\tau|_{dB}$$

$\Rightarrow$ entire graph shifts <u>up or down</u> by $|K_B|_{dB} = 20\log|K_B|$

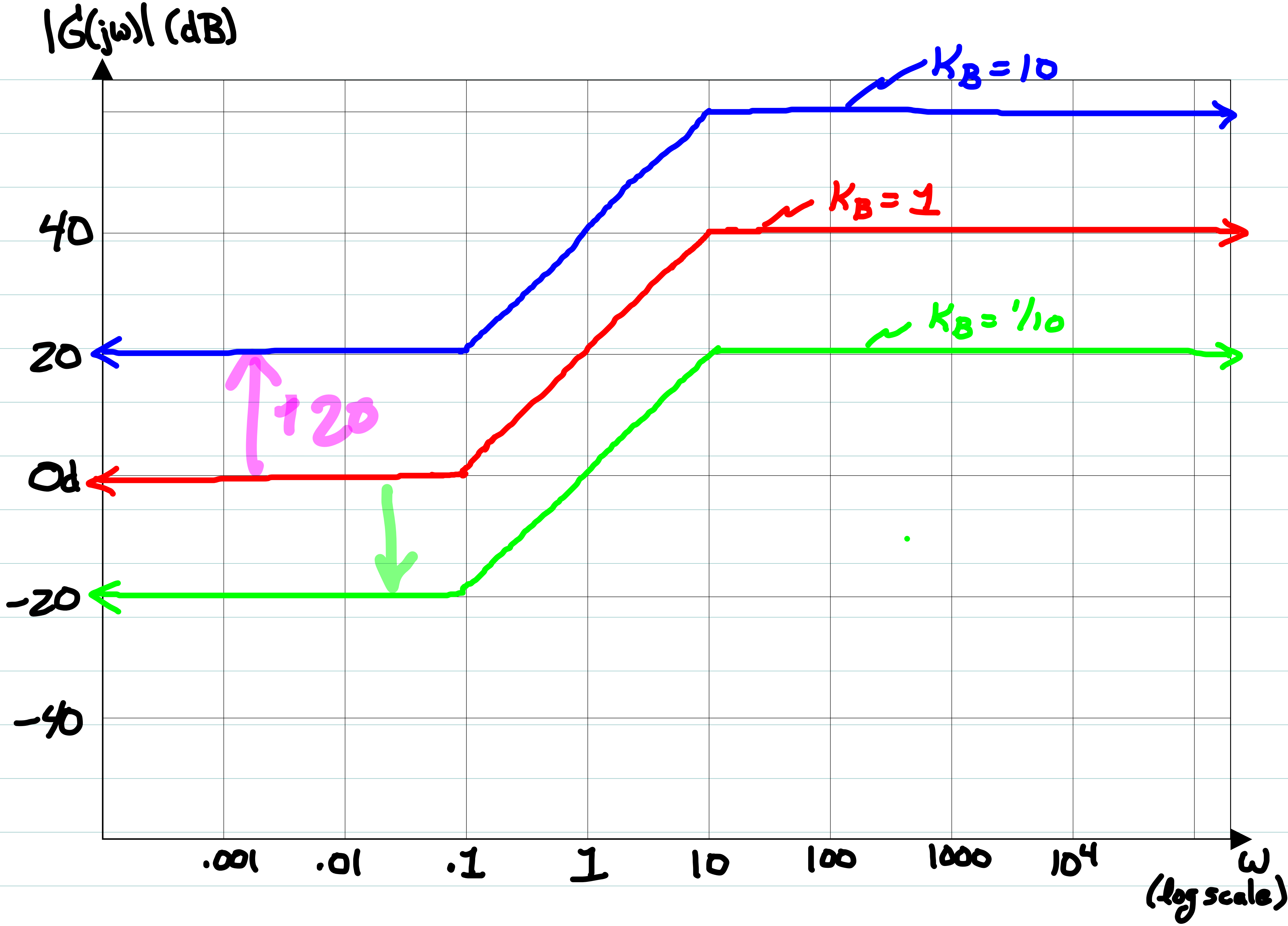
shifts <u>up</u> if $|K_B|_{dB} > \emptyset \Rightarrow |K_B| > 1$

shifts <u>down</u> if $|K_B|_{dB} < \emptyset \Rightarrow |K_B| < 1$

Remember the <u>sign</u> of K_B has <u>No effect</u> on the magnitude diagram!

<u>Example #3:</u>

$$G(s) = K_B\left[\frac{(10s+1)}{(s/10+1)}\right]$$



|G(jω)| (dB)
40
20
0d
-20
-40
$K_B = 10$
$K_B = 1$
$K_B = 1/10$
+20
.001
.01
.1
1
10
100
1000
10^4
ω
(log scale)

Repeated factors

$(1+j\omega\tau)^{\ell}$, ℓ integer ≥ 1

$$|(1+j\omega\tau)^{\ell}|_{dB} = 20 \log |1+j\omega\tau|^{\ell}$$

$$= (20\ell) \log |1+j\omega\tau|$$

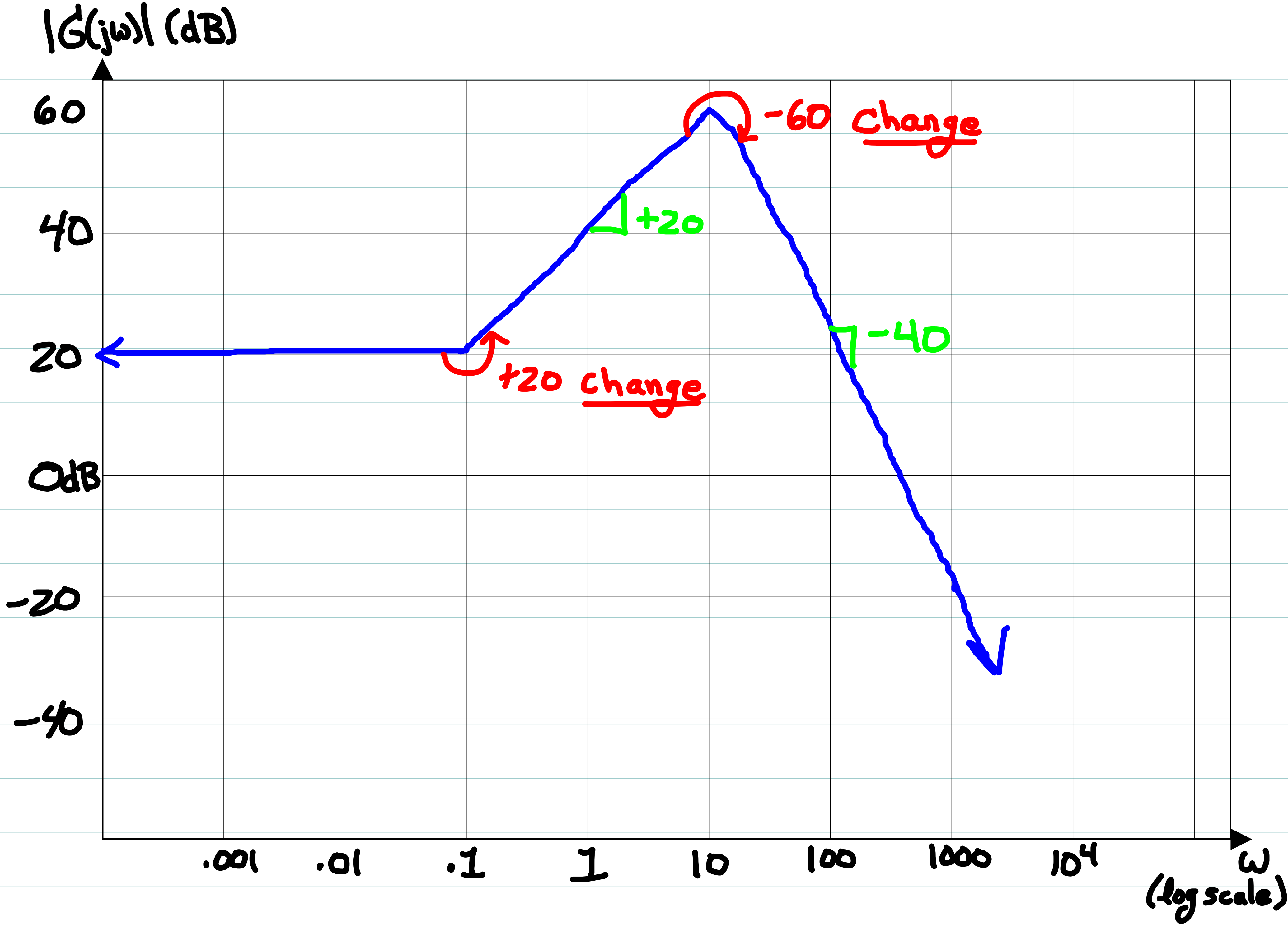
$\Rightarrow$ slope change is $\pm 20\ell$ at $\omega = |1/\tau|$

(positive for zero, negative for pole)

Example #4:

$$G(s) = 10\left[\frac{(10s+1)}{(s/10+1)^3}\right]$$

$+20$ slope change at $\omega = 1/10$, $\underline{-60}$ change at $\omega = 10$.



|G(jω)| (dB)
60
40
20
0dB
-20
-40
-60 change
+20
+20 change
-40
.001
.01
.1
1
10
100
1000
10^4
ω
(log scale)

Summary (so far)

$\Rightarrow$ Poles p_k and zeros z_i cause changes in $|G(j\omega)|_{dB}$ graph at corner frequencies $|p_k|$ and $|z_i|$

$\Rightarrow$ Slope of graph changes at these corners

$\Rightarrow$ Zero corners "bend up", i.e. change slope by +20 dB/dec

$\Rightarrow$ Pole corners "bend down", i.e. change slope by -20dB/dec

$\Rightarrow$ If $|K_B| \neq 1$, entire graph is raised or lowered by $|K_B|_{dB}$

Poles/zeros at origin

Poles at origin (type $N > \emptyset$) or zeros at origin ($N < \emptyset$)

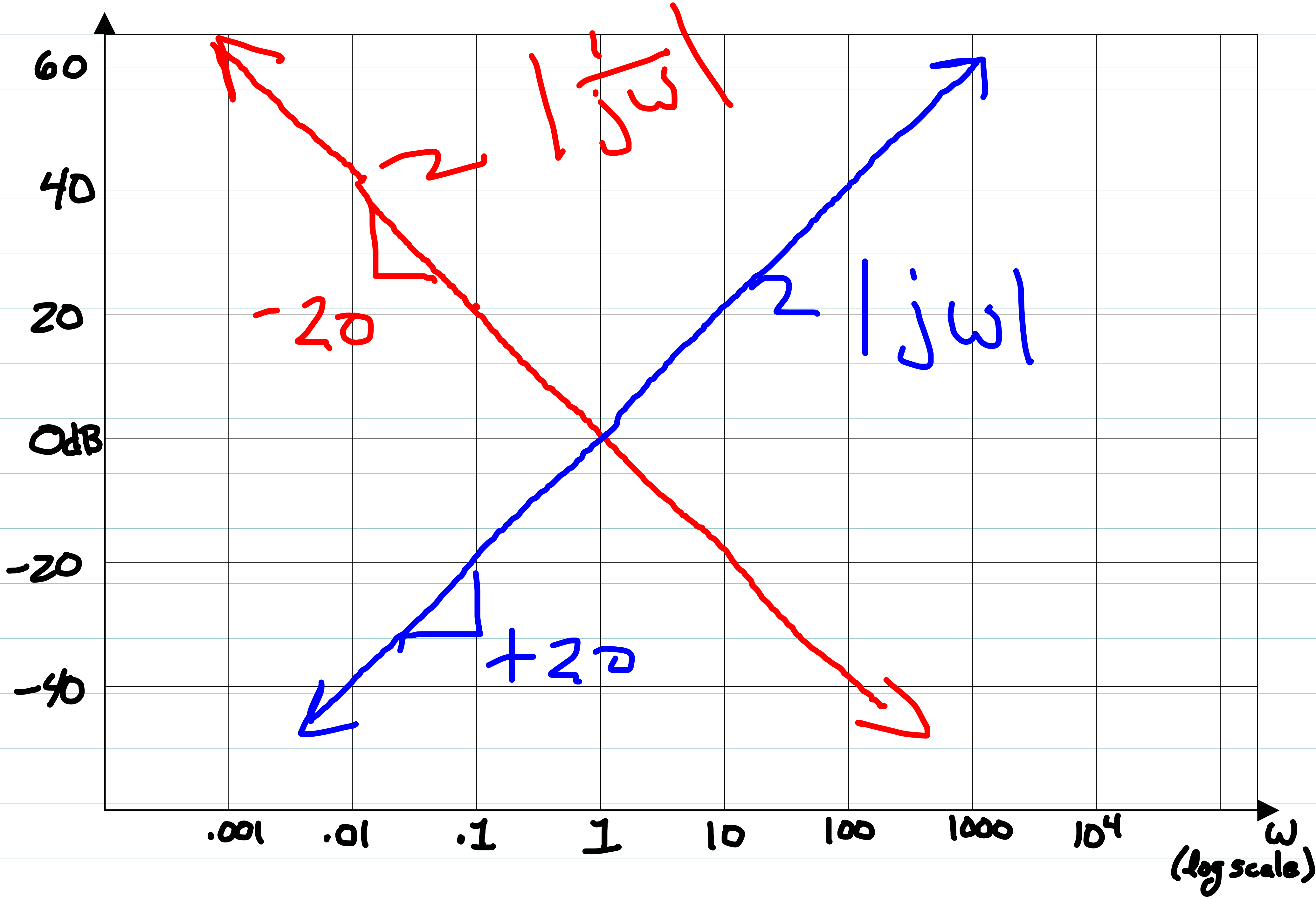
have corner frequencies at $\omega = 0$

$\Rightarrow$ infinitely far to left on horizontal frequency Axis.

These factors do not produce "visible" corners, instead contribute a constant slope of $-20N$ dB/dec for <u>all</u> freqs.

Note also: $|(j\omega)^N| = 1$ at $\omega = 1$ for any N

so graph of $|(j\omega)^N|_{dB}$ will pass through 0 dB at $\omega = \underline{1}$



60
40
20
0dB
-20
-40
.001
.01
.1
1
10
100
1000
10^4
ω (log scale)
$20 \log|j\omega|$
$-20 \log|j\omega|$
-20
+20

For $G(s)$ with poles/zeros at origin:

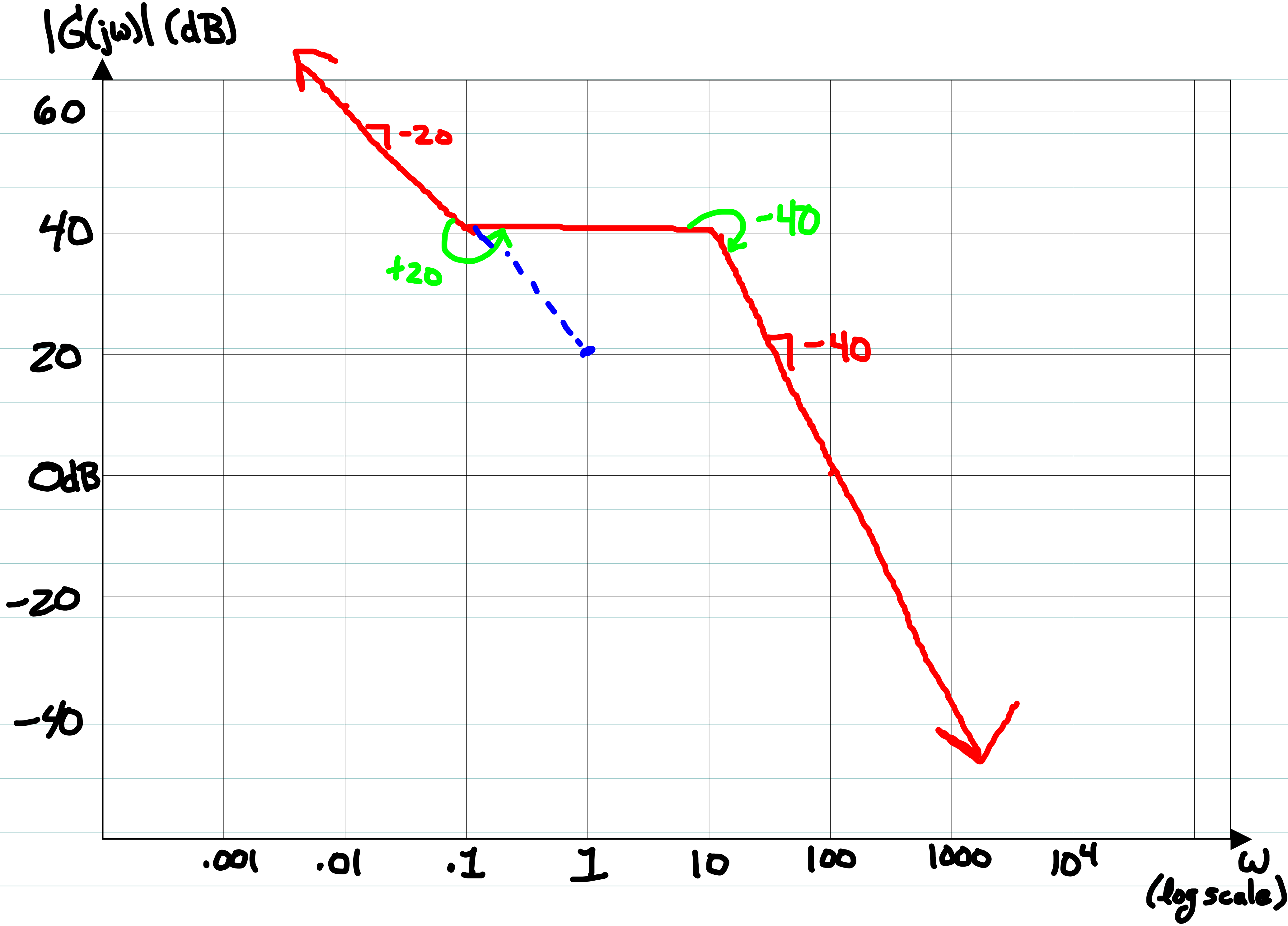
$\Rightarrow$ Start diagram by sketching effect of these poles at low frequencies

$\Rightarrow$ Note if $|K_B| \neq 1$, then this low frequency asymptote will pass through $|K_B|_{dB}$ at $\omega = 1$

$\Rightarrow$ Then add bends due to nonzero z_i and p_K as usual.

<u>Example:</u>

$$G(s) = 10\left[\frac{(10s+1)}{s(s/10+1)^2}\right]$$



|G(jw)| (dB)
60
40
20
0dB
-20
-40
-20
+20
-40
-40
.001
.01
.1
1
10
100
1000
10^4
ω
(log scale)

What about phase?

Recall:

$$\angle G(j\omega) = \angle K_B - N\angle(j\omega) + \sum_{i=1}^{m} \angle\left(1-\frac{j\omega}{z_i}\right) - \sum_{k=N+1}^{n} \left(1-\frac{j\omega}{p_k}\right)$$

$$\angle K_B = \begin{cases} \emptyset & K_B > \emptyset \\ -180 & K_B < \emptyset \end{cases} \quad \text{for all } \omega \geq \emptyset$$

$$\angle(j\omega) = 90^\circ \quad \text{for all } \omega \geq \emptyset$$

So, low frequency phase is constant at

$$-90N \quad \text{if } K_B > \emptyset$$

$$-180-90N \quad \text{if } K_B < \emptyset$$

Other poles/zeros will cause "bends" at higher freqs.

Phase response from other poles/zeros

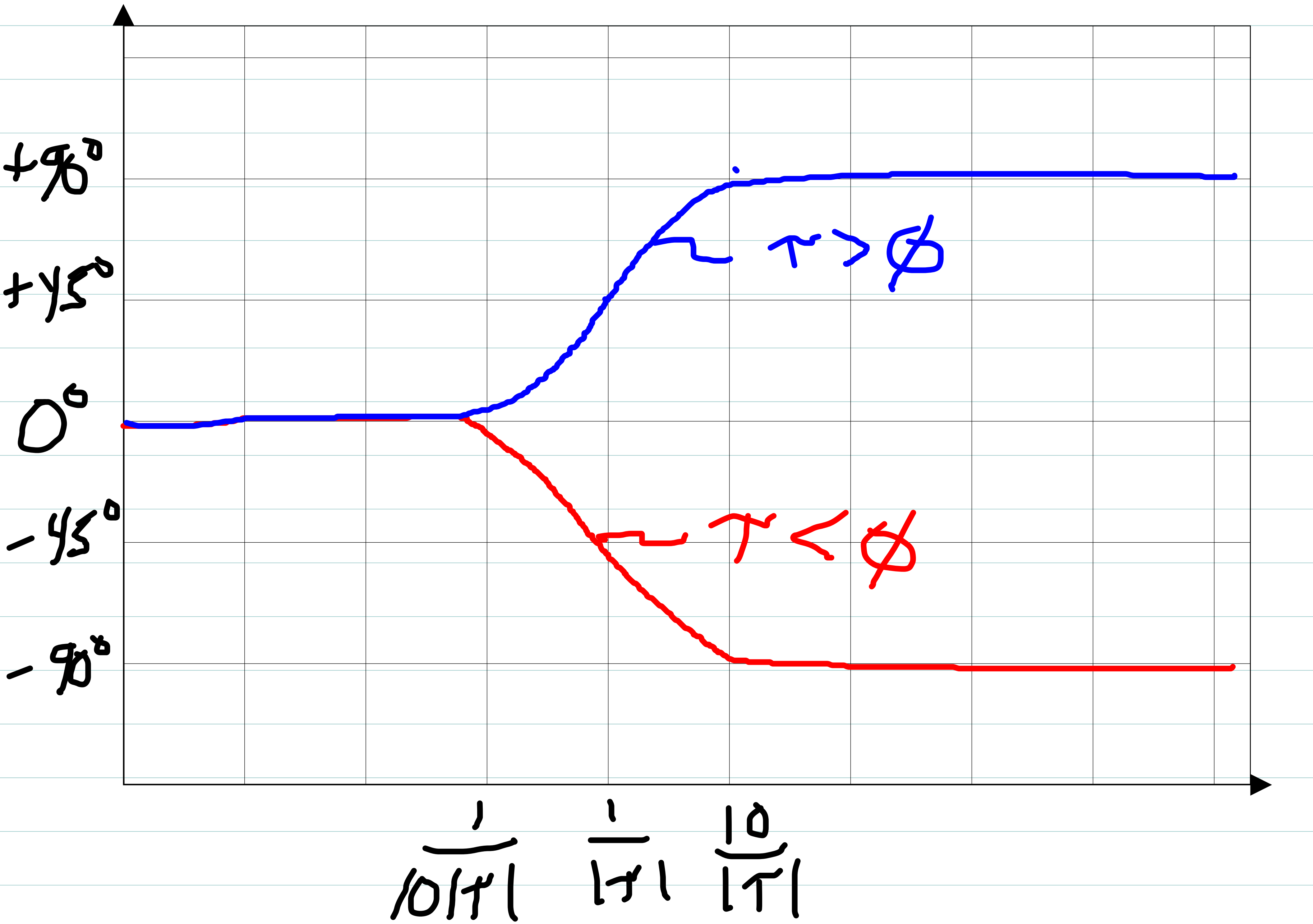
Consider again in generic form $(1+j\omega\tau)$ with $\tau = -1/z_i$ or $\tau = -1/p_k$

$$\measuredangle(1+j\omega\tau) = \tan^{-1}\omega\tau$$

$$= \begin{cases} \emptyset & \text{if } \omega \ll 1/|\tau| \\ +45^\circ & \text{if } \omega = 1/|\tau| \\ +9\emptyset^\circ & \text{if } \omega \gg 1/|\tau| \end{cases}$$

above is for $\tau > 0$. If instead $\tau < 0$

$$\measuredangle(1+j\omega\tau) = -\tan^{-1}\omega|\tau| = \begin{cases} \emptyset & \text{if } \omega \ll 1/|\tau| \\ -45^\circ & \text{if } \omega = 1/|\tau| \\ -90^\circ & \text{if } \omega \gg 1/|\tau| \end{cases}$$


+90°
+45°
0°
-45°
-90°
τ > 0
τ < 0
1/(10|τ|)
1/|τ|
10/|τ|

Observations

$\Rightarrow$ Phase change due to a single factor occurs in a 2 decade band of frequencies centered at the *magnitude corner frequency* $1/|\tau|$

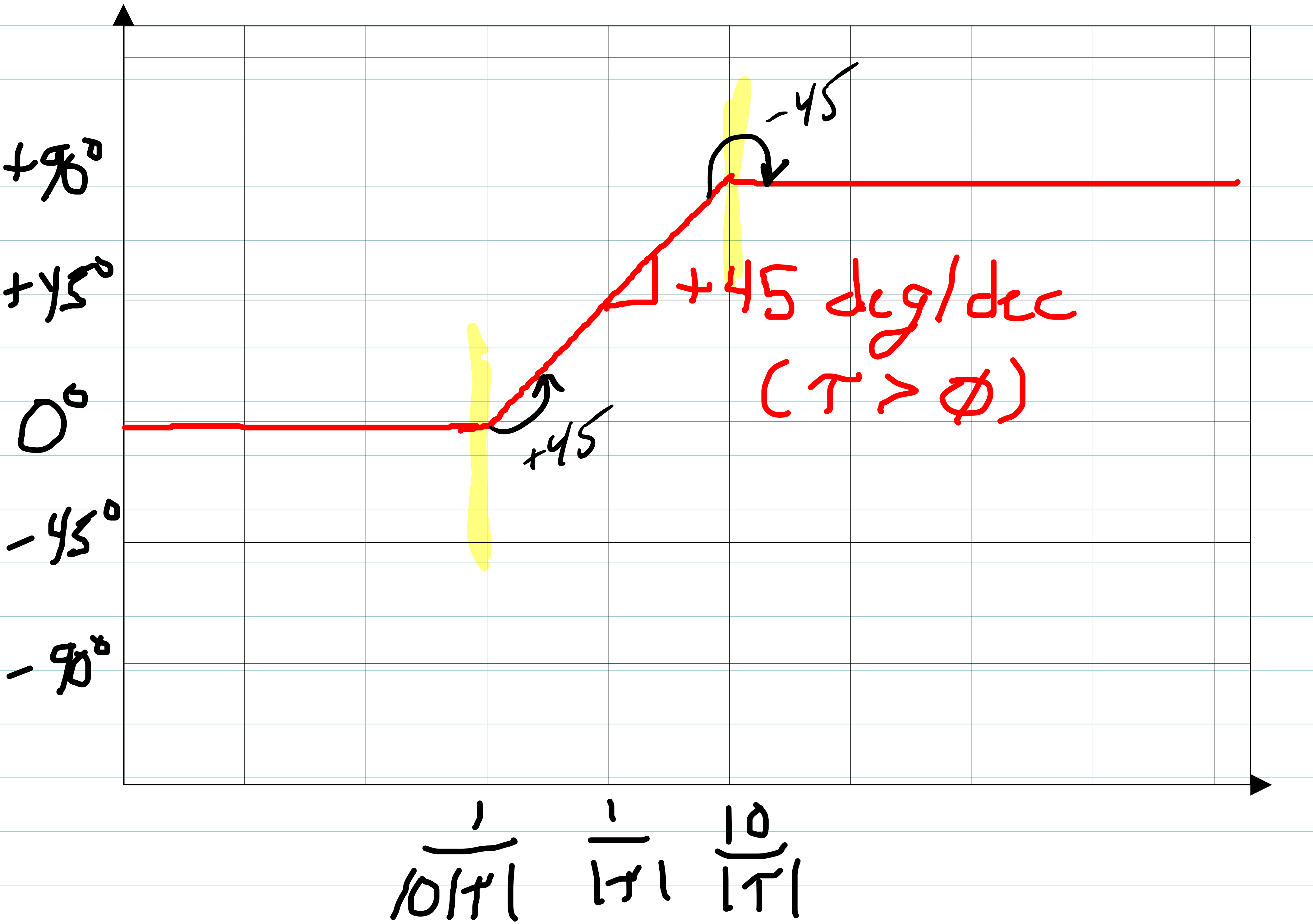
i.e. in band $\frac{1}{10|\tau|} \leq \omega \leq 10/|\tau|$

$\Rightarrow$ Phase is <u>constant</u> outside this band

low freq phase $\approx 0°$

h.f. phase $\approx \pm 90°$ ($+90°$ if $\tau > 0$, $-90°$ if $\tau < 0$

$\Rightarrow$ Phase change is approximate <u>linear</u> across band with slope $\pm 45°/\text{dec}$



+90°
+45°
0°
-45°
-90°
-45
+45
+45 deg/dec
(τ > 0)
1/(10|τ|)
1/|τ|
10/|τ|

Sign of phase change depends on:

$\Rightarrow$ whether factor is pole or zero

$\Rightarrow$ whether factor is RHP ($\tau < \emptyset$) or LHP ($\tau > \emptyset$)

Suppose all factors are LHP, $z_i < \emptyset$ $p_k < \emptyset$

then all $\tau = {}^{-1}/z_i$ or ${}^{-1}/p_k$ are positive.

This is called the "minimum phase" case

Then:

$\Rightarrow$ zeros cause $+90°$ phase change over band $\frac{|z_i|}{10}$ to $10|z_i|$

$\Rightarrow$ poles cause $-90°$ change over $\frac{|p_k|}{10}$ to $10|p_k|$

(Minimum Phase Systems)

<u>Slopes</u> of phase change are $+45°/\text{dec}$ (zeros) or $-45°/\text{dec}$ (poles) in these bands

Note phase changes in minimum phase cases mirror those for magnitude changes:

⇒ zeros cause <u>positive</u> slope changes
⇒ poles cause <u>negative</u> slope changes.

Graphical addition is again straightforward, but requires a little care:

⇒ slopes are nonzero only in a 2 decade band

⇒ bands from different factors may overlap.

Example:

$$G(s) = \frac{10s+1}{s(s+1)(s/10+1)}$$

low freq. phase $-90°$

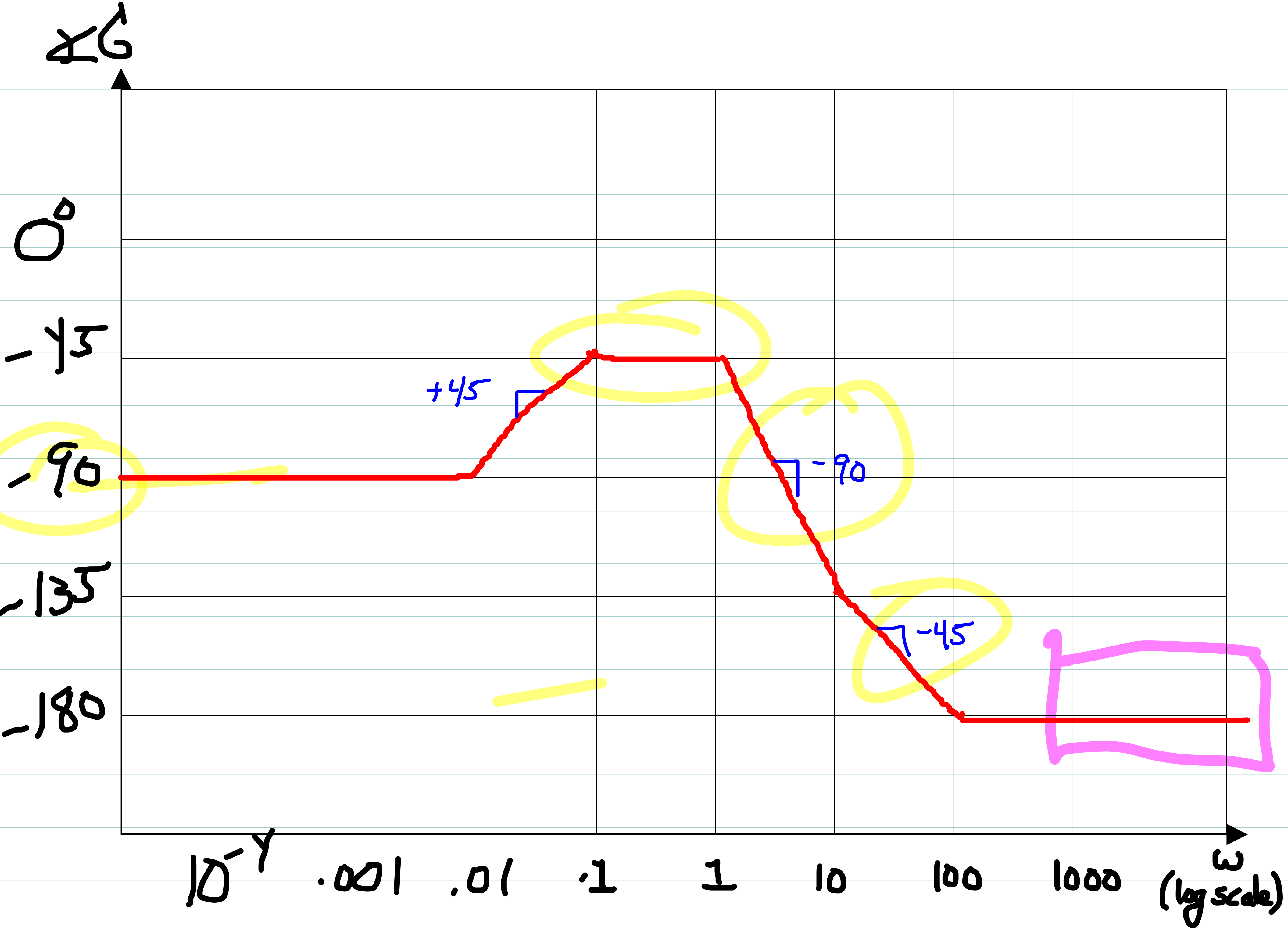
Phase changes:

$+45°$/dec in .01 to 1
$-45°$/dec in .1 to 10
$-45°$/dec in 1 to 100

Net:

$+45°$/dec in .01 to .1
$0°$/dec in .1 to 1
$-90°$/dec in 1 to 10
$-45°$/dec in 10 to 100

constant for $\omega > 100$.


∡G
0°
-45
-90
-135
-180
+45
-90
-45
10^{-4}
.001
.01
.1
1
10
100
1000
ω
(log scale)

Repeated factors

Repeated factors $(1+j\omega T)^{l}$ multiply the phase changes by l, just like magnitudes.

Example:

$$G(s) = \frac{10s+1}{s(s+1)^2}$$

Changes:

+45°/dec in .01 to 1

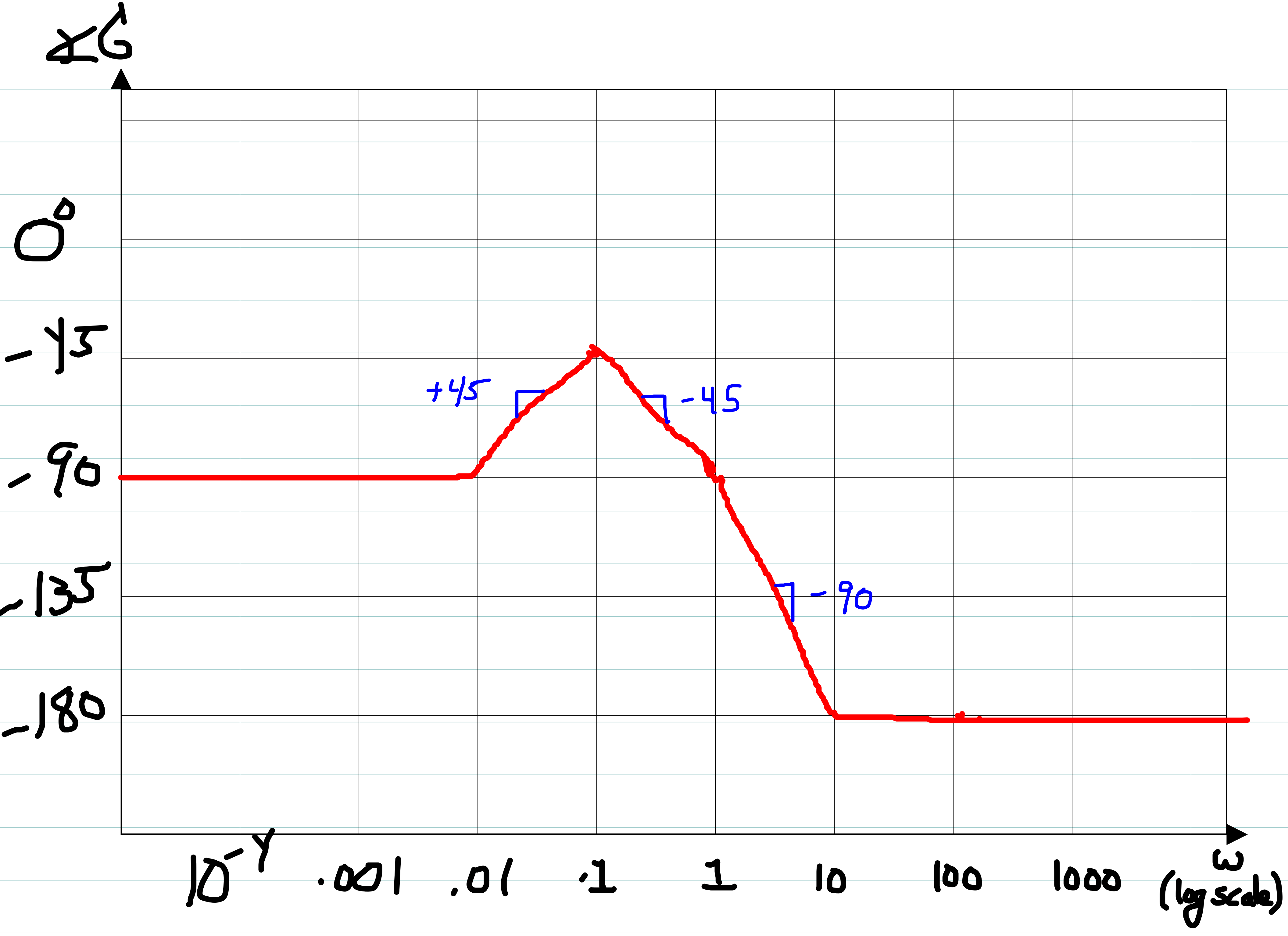
−90°/dec in .1 to 10

Net:

+45°/dec in .01 to .1

−45°/dec in .1 to 1

−90°/dec in 1 to 10



∡G
0°
-45
-90
-135
-180
+45
-45
-90
10⁻⁴
.001
.01
.1
1
10
100
1000
ω
(log scale)

Summary (minimum phase)

$\Rightarrow$ Low freq. phase is $\measuredangle K_B - N\,90^\circ$

$\Rightarrow$ high freq. phase is $\measuredangle K_B - 90^\circ(n-m)$

"relative degree"
#poles − #zeros

$\Rightarrow$ Note low and high freq. phases are constant
(slope is zero).

$\Rightarrow$ Recall typically $n>m$ for a physical system so high freq. phase is typically negative for a minimum phase system.

$\Rightarrow$ zeros cause $+90^\circ$ change at rate of $+45^\circ$/dec in 2 decade band centered at $|z_i|$

$\Rightarrow$ poles cause -90° change at rate of -45°/dec in 2 decade band centered at $|p_k|$.

Can be tricky to accurately sketch phase

=> Overlapping change regions for multiple factors

=> No standard formula for phase change of underdamped factors

=> Helps to 1st make a table of slope changes over frequency ranges as above

=> Generally, straight-line phase sketch is less accurate than magnitude sketch.

=> Still sufficiently accurate to give us a good general idea of phase behavior.

=> We'll use Matlab when greater accuracy is required.

What about phase?

Recall:

$$\measuredangle G(j\omega) = \measuredangle K_B - N\measuredangle(j\omega) + \sum_{i=1}^{m} \measuredangle\left(1-\frac{j\omega}{z_i}\right) - \sum_{K=N+1}^{n} \left(1-\frac{j\omega}{p_K}\right)$$

$$\measuredangle K_B = \begin{cases} \emptyset & K_B > \emptyset \\ -180 & K_B < \emptyset \end{cases} \quad \text{for all } \omega \geq \emptyset$$

$$\measuredangle(j\omega) = 90^\circ \quad \text{for all } \omega \geq \emptyset$$

So, low frequency phase is <u>constant</u> at

$-90N$ if $K_B > \emptyset$

$-180-90N$ if $K_B < \emptyset$

Other poles/zeros will cause "bends" at higher freqs.

Phase response from other poles/zeros

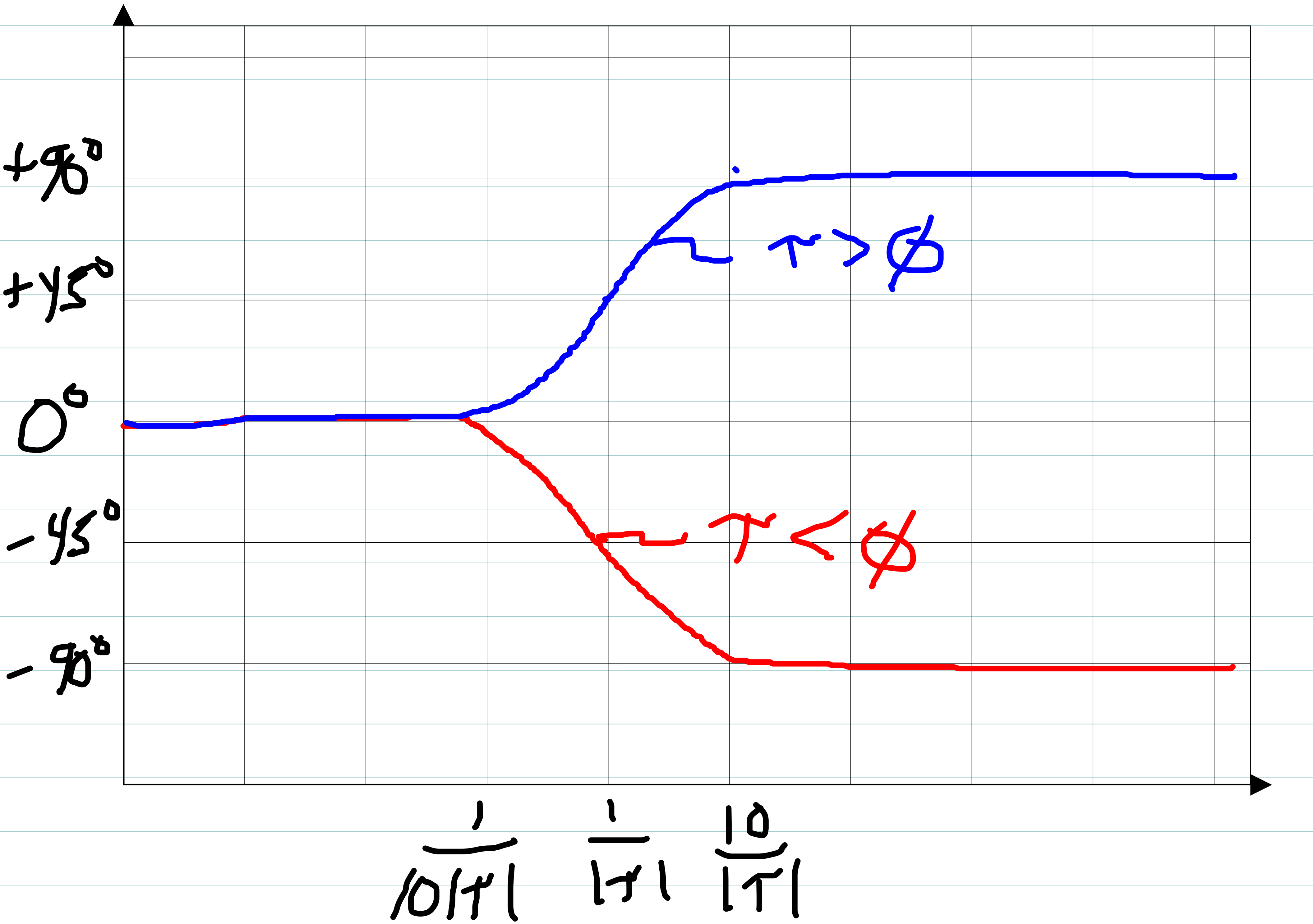
Consider again in generic form $(1+j\omega\tau)$ with $\tau = -1/z_i$ or $\tau = -1/p_k$

$$\angle(1+j\omega\tau) = \tan^{-1}\omega\tau$$

$$= \begin{cases} \phi & \text{if } \omega \ll 1/|\tau| \\ +45^\circ & \text{if } \omega = 1/|\tau| \\ +90^\circ & \text{if } \omega \gg 1/|\tau| \end{cases}$$

above is for $\tau > 0$. If instead $\tau < 0$

$$\angle(1+j\omega\tau) = -\tan^{-1}\omega|\tau| = \begin{cases} \phi & \text{if } \omega \ll 1/|\tau| \\ -45^\circ & \text{if } \omega = 1/|\tau| \\ -90^\circ & \text{if } \omega \gg 1/|\tau| \end{cases}$$


+90°
+45°
0°
-45°
-90°
τ > 0
τ < 0
1/10|τ|
1/|τ|
10/|τ|

Observations

$\Rightarrow$ Phase change due to a single factor occurs in a 2 decade band of frequencies centered at the magnitude corner frequency $1/|\tau|$

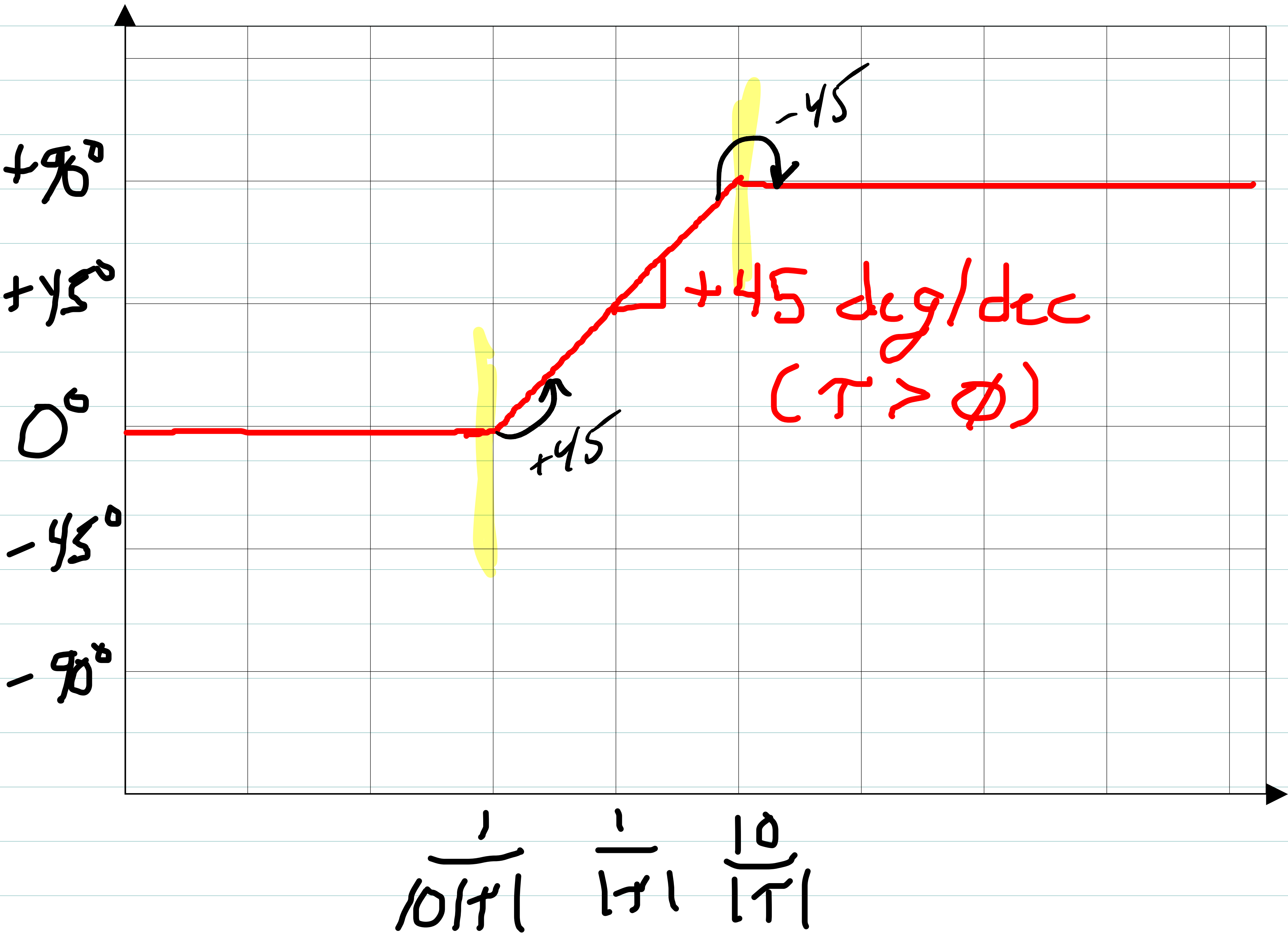
i.e. in band $\frac{1}{10|\tau|} \leq \omega \leq 10/|\tau|$

$\Rightarrow$ Phase is constant outside this band

low freq phase $\approx 0°$

h.f. phase $\approx \pm 90°$ ($+90°$ if $\tau > 0$, $-90°$ if $\tau < 0$

$\Rightarrow$ Phase change is approximate linear across band with slope $\pm 45°/\text{dec}$


+90°
+45°
0°
-45°
-90°
-45
+45
+45 deg/dec
(τ > 0)
1/(10|τ|)
1/|τ|
10/|τ|

Sign of phase change depends on:

⇒ whether factor is pole or zero

⇒ whether factor is RHP ($\tau < \emptyset$) or LHP ($\tau > \emptyset$)

Suppose all factors are LHP, $z_i < \emptyset$ $p_k < \emptyset$

then all $\tau = {-1}/{z_i}$ or ${-1}/{p_k}$ are positive.

This is called the "minimum phase" case

Then:

⇒ zeros cause $+90°$ phase change over band $\frac{|z_i|}{10}$ to $10|z_i|$

⇒ poles cause $-90°$ change over $\frac{|p_k|}{10}$ to $10|p_k|$

(Minimum Phase Systems)

Slopes of phase change are $+45°/\text{dec}$ (zeros) or $-45°/\text{dec}$ (poles) in these bands

Note phase changes in minimum phase cases mirror those for magnitude changes:

$\Rightarrow$ zeros cause positive slope changes

$\Rightarrow$ poles cause negative slope changes.

Graphical addition is again straightforward, but requires a little care:

$\Rightarrow$ slopes are nonzero only in a 2 decade band

$\Rightarrow$ bands from different factors may overlap.

Example:

$$G(s) = \frac{10s+1}{s(s+1)(s/10+1)}$$

low freq. phase $-90°$

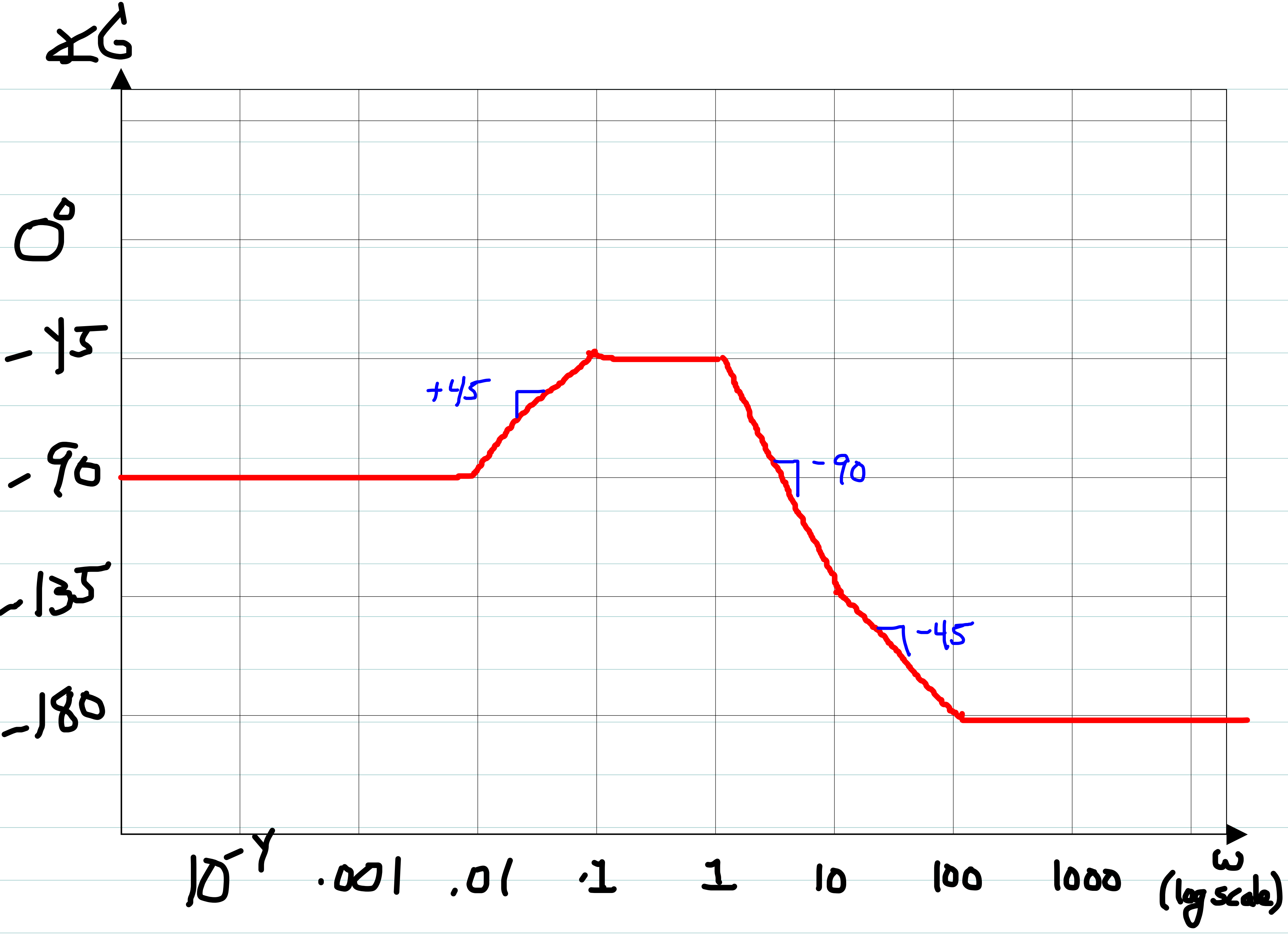
Phase changes:

$+45°$/dec	in	$.01$ to 1
$-45°$/dec	in	$.1$ to 10
$-45°$/dec	in	1 to 100

Net:

$+45°$/dec	in	$.01$ to $.1$
$0°$/dec	in	$.1$ to 1
$-90°$/dec	in	1 to 10
$-45°$/dec	in	10 to 100

Constant for $\omega > 100$.



∡G
0°
-45
-90
-135
-180
+45
-90
-45
10^{-4}
.001
.01
.1
1
10
100
1000
ω
(log scale)

Repeated factors

Repeated factors $(1+j\omega T)^{l}$ multiply the phase changes by l, just like magnitudes.

Example:

$$G(s) = \frac{10s+1}{s(s+1)^2}$$

Changes:

$+45°/\text{dec}$ in $.01$ to 1

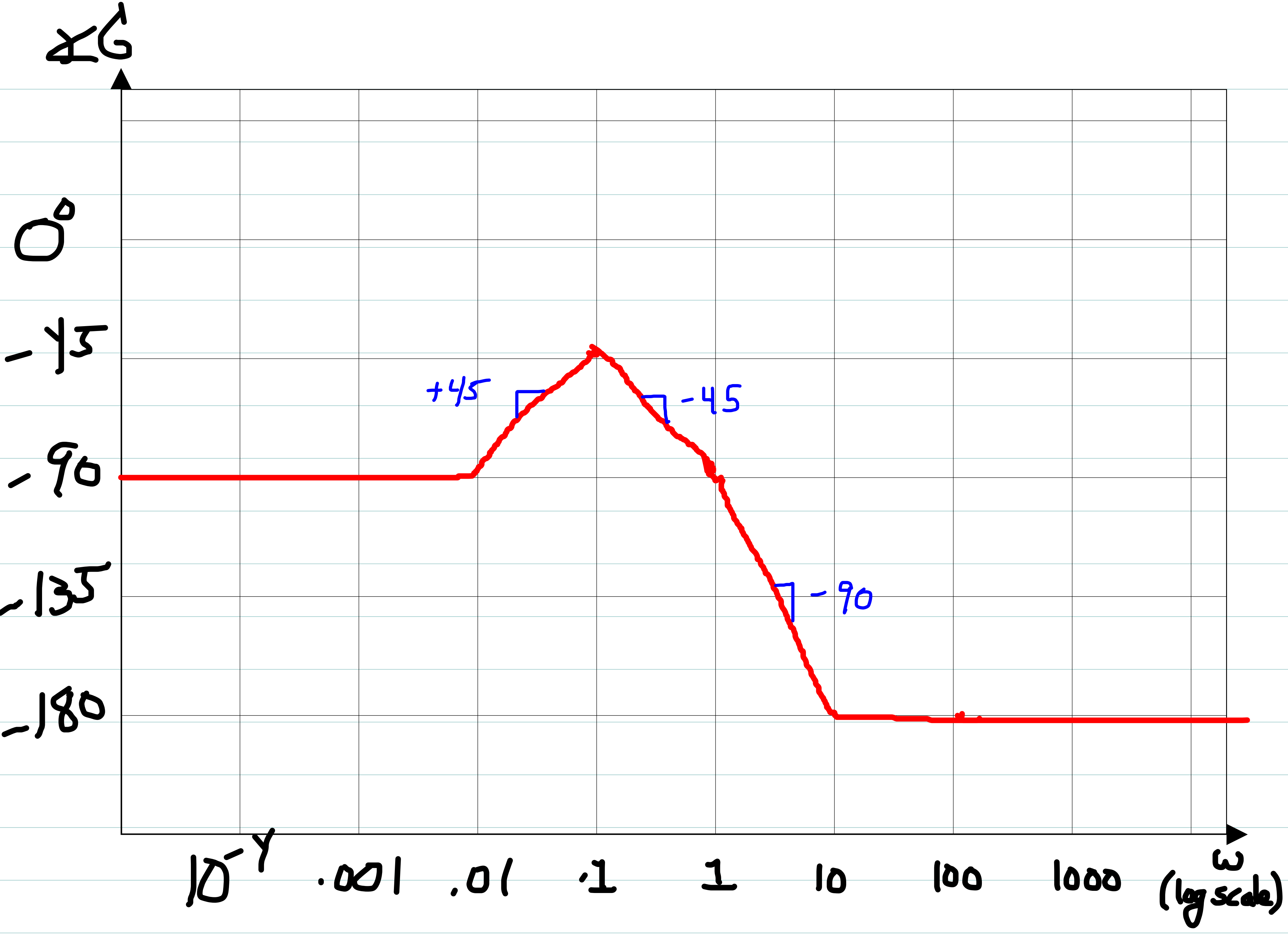
$-90°/\text{dec}$ in $.1$ to 10

Net:

$+45°/\text{dec}$ in $.01$ to $.1$

$-45°/\text{dec}$ in $.1$ to 1

$-90°/\text{dec}$ in 1 to 10



∠G
0°
-45
-90
-135
-180
+45
-45
-90
10^{-4}
.001
.01
.1
1
10
100
1000
ω
(log scale)

Summary (minimum phase)

⇒ Low freq. phase is $\measuredangle K_B - N\,90°$

⇒ high freq. phase is $\measuredangle K_B - 90°(n-m)$

⇒ Note low and high freq. phases are constant (slope is zero).

⇒ Recall typically $n > m$ for a physical system so high freq. phase is typically negative for a minimum phase system.

⇒ zeros cause $+90°$ change at rate of $+45°/\text{dec}$ in 2 decade band centered at $|z_i|$

⇒ poles cause $-90°$ change at rate of $-45°/\text{dec}$ in 2 decade band centered at $|p_k|$.

Can be tricky to accurately sketch phase

⇒ Overlapping change regions for multiple factors

⇒ No standard formula for phase change of underdamped factors

⇒ Helps to 1st make a table of slope changes over frequency ranges as above

⇒ Generally, straight-line phase sketch is less accurate than magnitude sketch.

⇒ Still sufficiently accurate to give us a good general idea of phase behavior.

⇒ We'll use Matlab when greater accuracy is required.

Non-minimum phase systems

If any poles or zeros of $G(s)$ in RHP, the system is "Non-minimum phase"

Corresponds to $\tau < \phi$ in phase analysis and $\measuredangle(1+j\omega\tau) = -\tan^{-1}\omega/|\tau|$.

$\Rightarrow$ Phase response is negative of that seen above

$\Rightarrow$ In particular, zeros cause -90 deg phase change in 2 decade band around corner freq.

poles cause $+90$ deg change

Opposite of minimum phase behavior, but

$\boxed{\text{Corner freqs unchanged}}$ ($|z_i|$ or $|p_k|$)

Example:

$$G(s) = \frac{(10s+1)}{(1-s)}$$

Min phase zero: $+45°$/dec change in .01 to 1

Nonmin phase pole: $+45°$/dec change in .1 to 10

Net: $+45°$/dec in .01 to .1
$+90°$/dec in .1 to 1
$+45°$/dec in 1 to 10.

Note: h.f. phase is $+180°$ here. Above rule for h.f. phase in min phase systems does not apply if $G(s)$ has RHP poles or zeros

Underdamped factors

$$(s^2+2\zeta\omega_n s+\omega_n^2) \Rightarrow \left[\left(\frac{s}{\omega_n}\right)^2+2\zeta\left(\frac{s}{\omega_n}\right)+1\right] \text{ in Bode form.}$$

How do we draw magnitude response when $G(s)$ contains these factors?

$\Rightarrow$ If $\frac{\sqrt{2}}{2} \leq \zeta \leq 1$, we can well approximate the response as a repeated pole at $-\omega_n$ (it isn't really, but it's a good approx to sketch this way).

$\Rightarrow$ If $0 \leq \zeta < \frac{\sqrt{2}}{2}$ a more substantial correction is needed...

$\Rightarrow$ To illustrate, suppose

$$G(s) = \frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2} = \frac{1}{\left[(s/\omega_n)^2+2\zeta\left(\frac{s}{\omega_n}\right)+1\right]}$$

Let's find $|G(j\omega)|$ here

$$|G(j\omega)| = \left|\left(\frac{j\omega}{\omega_n}\right)^2 + 2\xi\left(\frac{j\omega}{\omega_n}\right) + 1\right|^{-1}$$

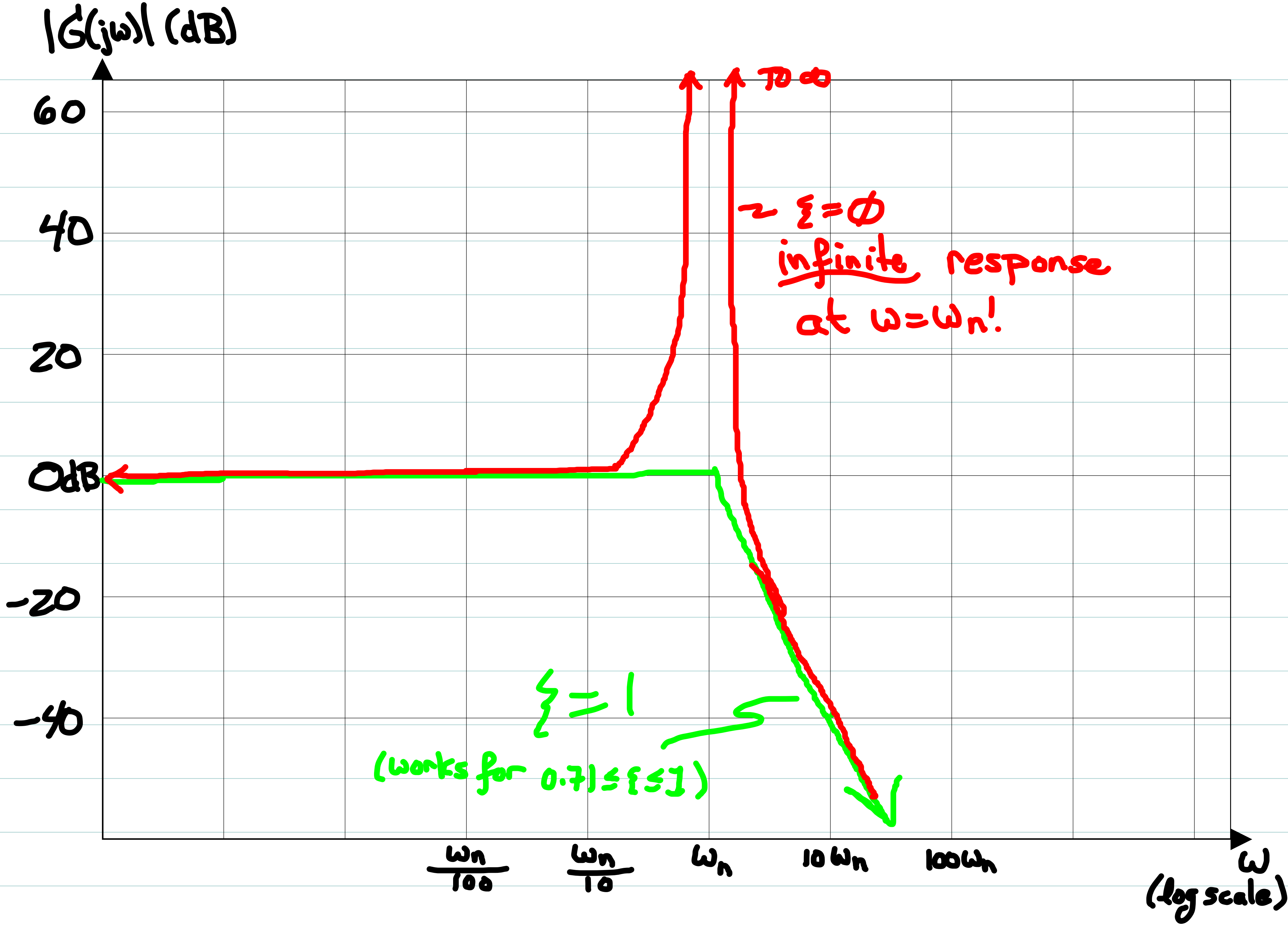
$$= \left[\sqrt{\left(1-\left(\frac{\omega}{\omega_n}\right)^2\right)^2 + 4\xi^2\left(\frac{\omega}{\omega_n}\right)^2}\,\right]^{-1}$$

Which is ugly, so why bother?

Consider if $\xi=\emptyset$, then

$$|G(j\omega)| = \frac{1}{\left|1-\left(\frac{\omega}{\omega_n}\right)^2\right|} \approx \begin{cases} 1 & \text{if } \omega \ll \omega_n \\ \left|\frac{\omega}{\omega_n}\right|^{-2} & \omega \gg \omega_n \end{cases}$$

so that $|G(j\omega)|\Big|_{\omega=\omega_n} = \infty$!!! Definitely something going on!



|G(jw)| (dB)
60
40
20
0dB
-20
-40
TO ∞
← ξ = 0
infinite response
at ω = ωn!
ξ = 1
(works for 0.71 ≤ ξ ≤ 1)
ωn/100
ωn/10
ωn
10 ωn
100ωn
ω
(log scale)

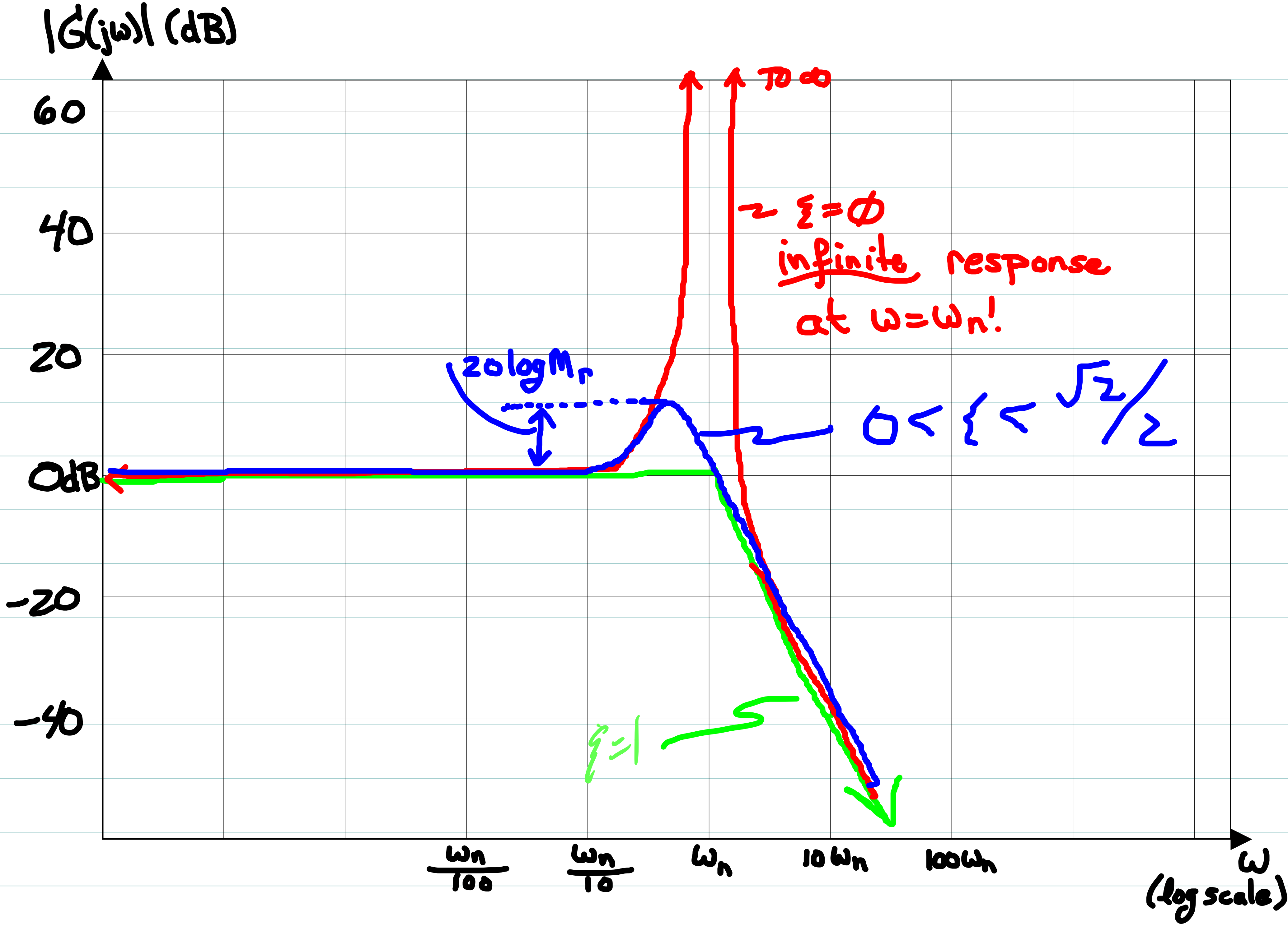
When $0 < \zeta < \sqrt{2}/2$, a similar "peaking" phenomenon occurs, but peak height is finite:

for
$$G(s) = \left[\left(\frac{s}{\omega_n}\right)^2 + 2\zeta\left(\frac{s}{\omega_n}\right) + 1\right]^{-1}$$

$\max\limits_{\omega \geq 0} |G(j\omega)|$ occurs at:

$$\omega_r = \omega_n\sqrt{1-2\zeta^2}$$

and $|G(j\omega_r)| = \dfrac{1}{2\zeta\sqrt{1-\zeta^2}} \triangleq M_r$



|G(jw)| (dB)
60
40
20
0dB
-20
-40
TO ∞
ξ=0
infinite response
at ω=ωn!
20logMr
0<ξ<√2/2
ξ=1
ωn/100
ωn/10
ωn
10ωn
100ωn
ω
(log scale)

This is the phenomenon of <u>resonance</u>

An ideal (no zeros) underdamped 2^{nd} order system with $0 \le \zeta < \sqrt{2}/2$ will exhibit output amplitudes significantly greater than the input amplitude when input frequency is <u>close</u> to the natural frequency ω_n.

The largest amplitude ratio will occur at the

<u>resonant frequency</u> $\boxed{\omega_r = \omega_n\sqrt{1-2\zeta^2}} < \omega_n$

and the maximal amplitude ratio (maximal resonance) is

$$\boxed{M_r = \frac{1}{2\zeta\sqrt{1-\zeta^2}}}$$

Notes:

1.) Height of peak on diagram is M_r in dB, i.e. $20\log M_r$

2.) When 2nd order factor in TF with other factors, the peak is $20\log M_r$ above whatever magnitude the plot would otherwise have at ω_r. That is, M_r is a relative offset to plot, not absolute.

3.) For small ζ, say $0 < \zeta \lesssim 1/10$

$$\omega_r \approx \omega_n \quad \text{and} \quad M_r \approx 1/(2\zeta)$$

So $20\log M_r \approx -[6 + 20\log \zeta]$ is a good approximation

i.e. at $\zeta = 1/10$, $20\log M_r \approx +14$ dB

Example

$$G(s) = \left[\frac{(10s+1)}{s\left(\left(\frac{s}{10}\right)^2 + 0.2\left(\frac{s}{10}\right) + 1\right)}\right]$$

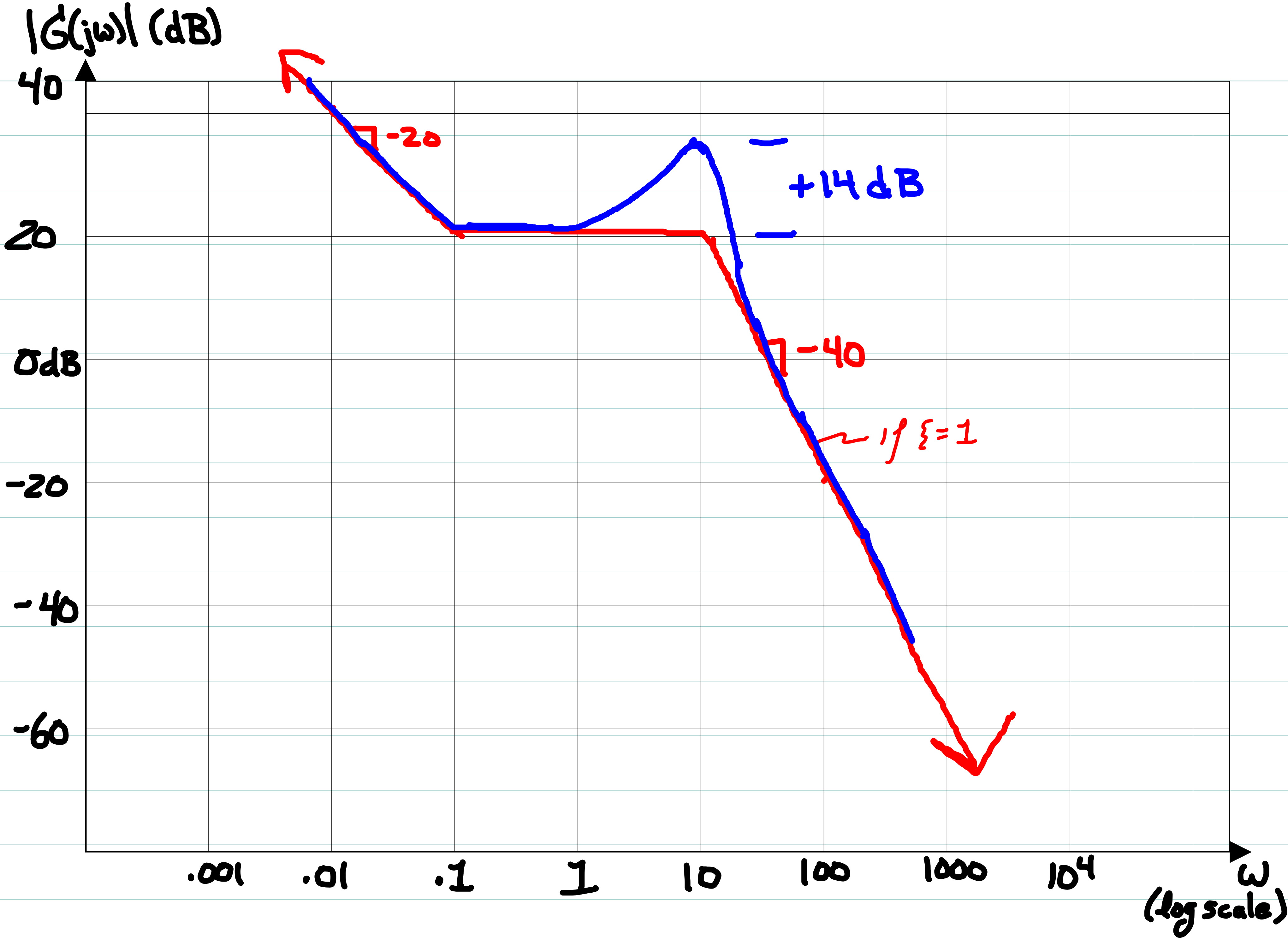
Same as example above, except:

$K_B = 1$ (instead of 10)

$\xi = 0.1$ (instead of $\xi = 1$)

$\omega_n = 10$

$\Rightarrow$ Expect resonant peak of height +14dB near $\omega = 10$.



|G(jw)| (dB)
40
20
0dB
-20
-40
-60
-20
-40
+14 dB
if ξ=1
.001
.01
.1
1
10
100
1000
10^4
ω
(log scale)

2^{nd} order min phase factors - Phase

$$\left[\left(\frac{s}{\omega_n}\right)^2 + 2\zeta\left(\frac{s}{\omega_n}\right)s + 1\right]^{\pm 1}$$

$\Rightarrow$ Like magnitudes, can sketch as repeated real factor ($\zeta = 1$) for $\sqrt{2}/2 \le \zeta \le 1$

$\Rightarrow$ for $0 \le \zeta < \frac{\sqrt{2}}{2}$, a more significant correction is needed

$\Rightarrow$ when $\zeta = 0$, phase changes <u>discontinuously</u> by $\pm 180°$ at $\omega = \omega_n$

<u>Example:</u>

$$G(s) = \frac{\left(s/10 + 1\right)}{\left[s^2 + 2\zeta s + 1\right]}$$

If $\zeta = 1$:

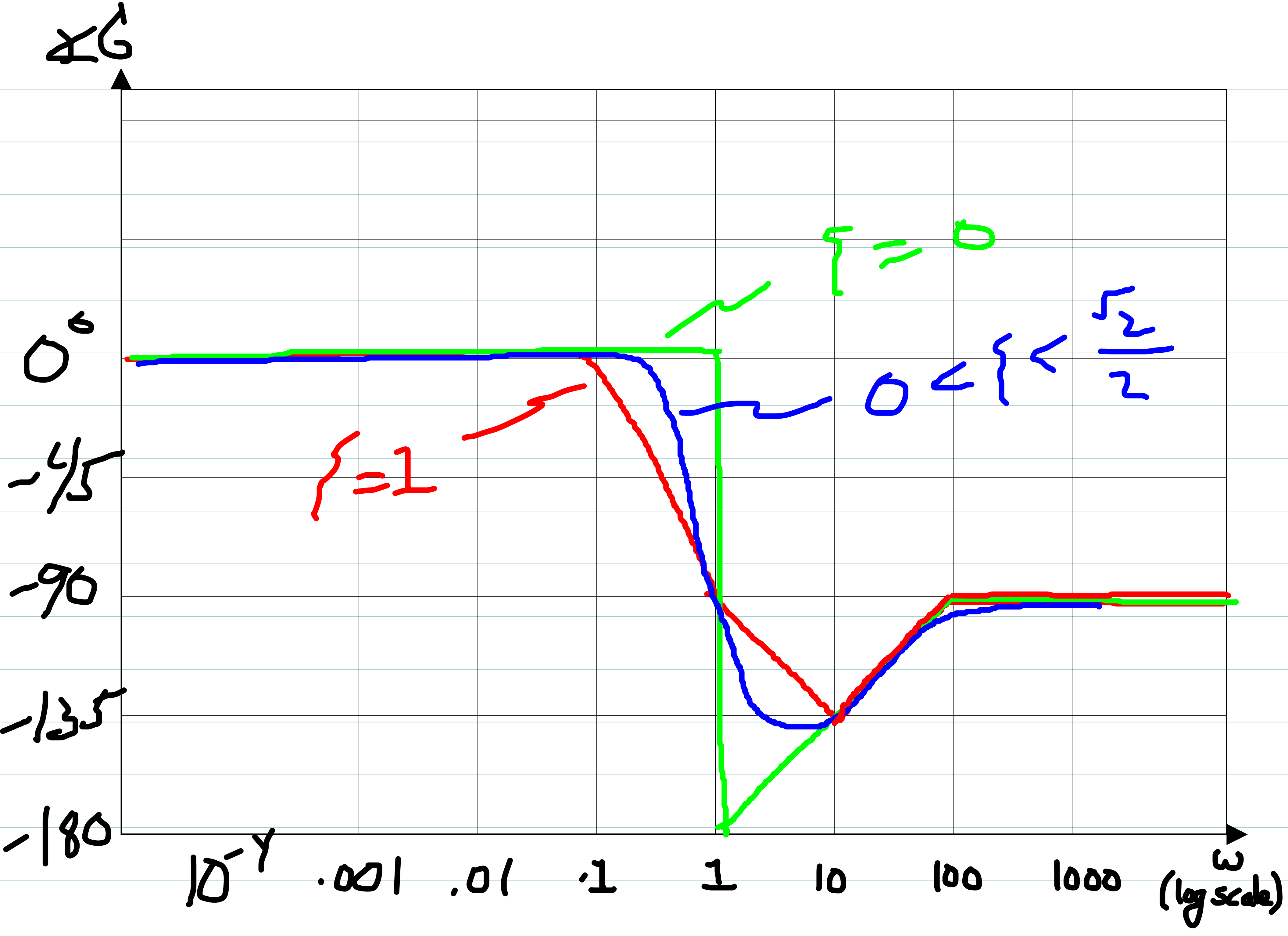
Change of $-90°/\text{dec}$ in $.1$ to 10
Change of $+45°/\text{dec}$ in 1 to 100

Net is $-90°/\text{dec}$ from $.1$ to 1
$-45°/\text{dec}$ from 1 to 10
$+45°/\text{dec}$ from 10 to 100

If $\zeta = 0$

Change of $+45°/\text{dec}$ in 1 to 100

$-180°$ drop at $\omega = \omega_n = 1$



∡G
0°
-45
-90
-135
-180
10^-4
.001
.01
.1
1
10
100
1000
ω (log scale)
ζ = 0
0 < ζ < √2/2
ζ = 1

Notes: (2nd order phase, small ξ)

- Unlike magnitude, no useful simple formula to quantify "steepness" of phase drop for small ξ.

- Usually sketch something in between the $\xi = 1$ and $\xi = 0$ limits

- Necessarily qualitative - will use Matlab when precise analysis is needed.

- Note generally that we expect to see steep phase drops near frequencies where magnitude diagram shows resonant peaks!

Polar Plots

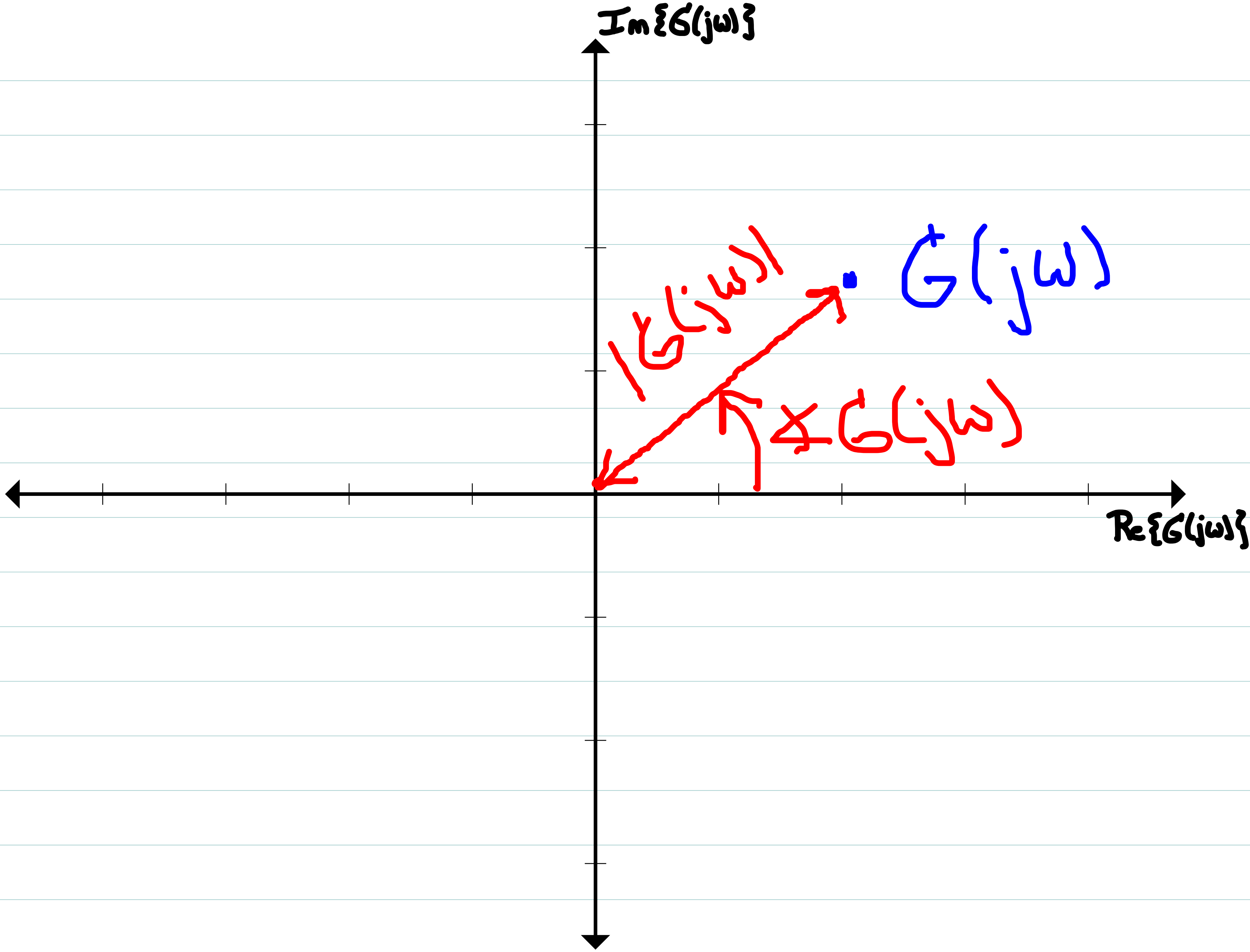
⇒ A different way of showing the properties of $G(j\omega)$

⇒ Bode plots $|G(j\omega)|$ and $\measuredangle G(j\omega)$ vs. ω, using logarithmic scales for $0 \le \omega < \infty$

⇒ Polar shows $G(j\omega)$ as points on complex plane as ω VARIES from $\emptyset$ to ∞ using actual (non-logarithmic) scales

⇒ Learn to sketch polar from Bode

⇒ We are aiming for something qualitatively correct, but will deliberately distort scales to make certain critical features readily apparent.


Im{G(jω)}
|G(jω)|
G(jω)
∡G(jω)
Re{G(jω)}

⇒ For each $\omega \in [0, \infty)$, $G(j\omega)$ is a <u>different</u> point on complex plane

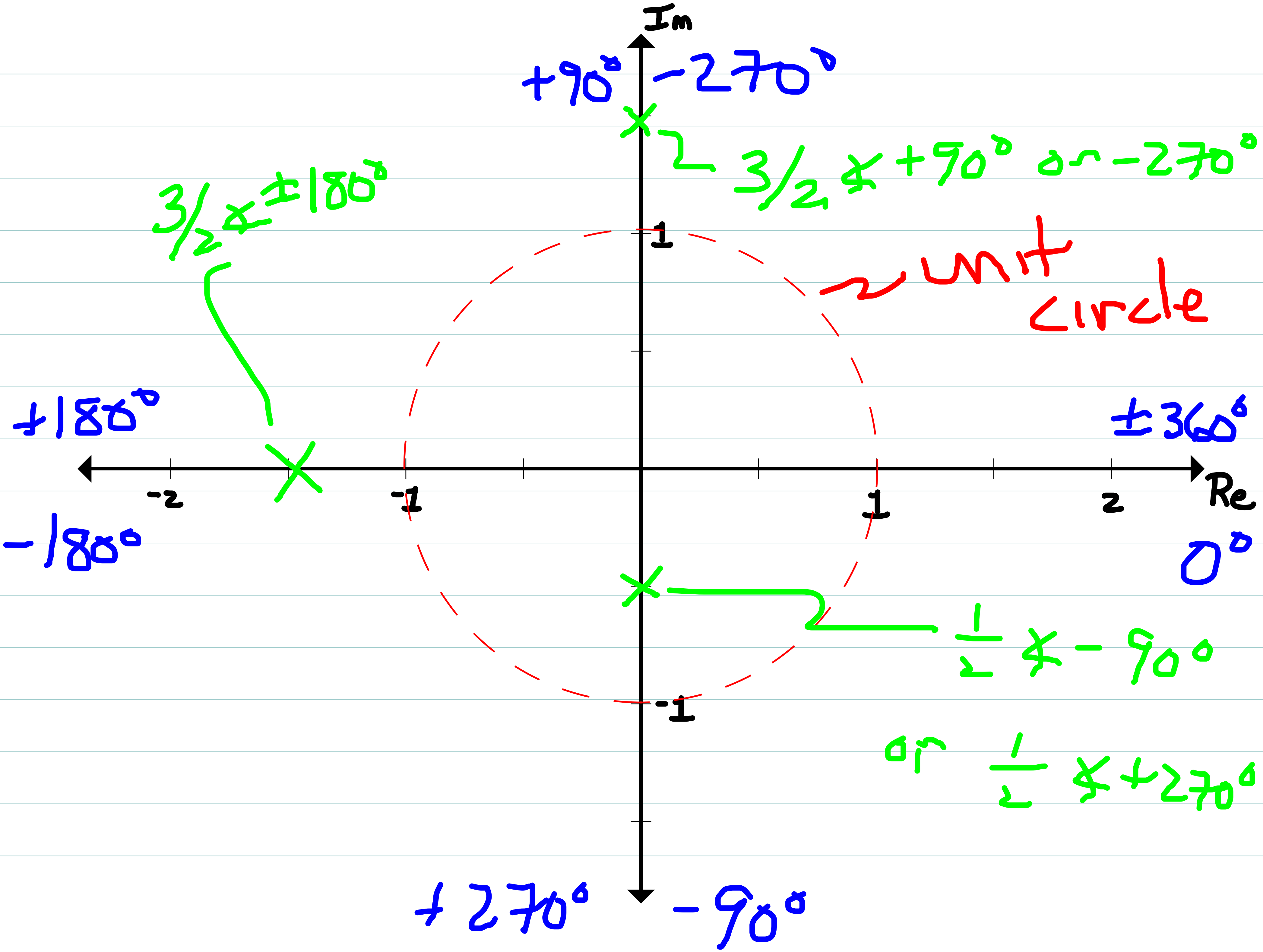
⇒ As ω varies from $\emptyset$ to ∞, these points will trace out a <u>curve</u> on complex plane.

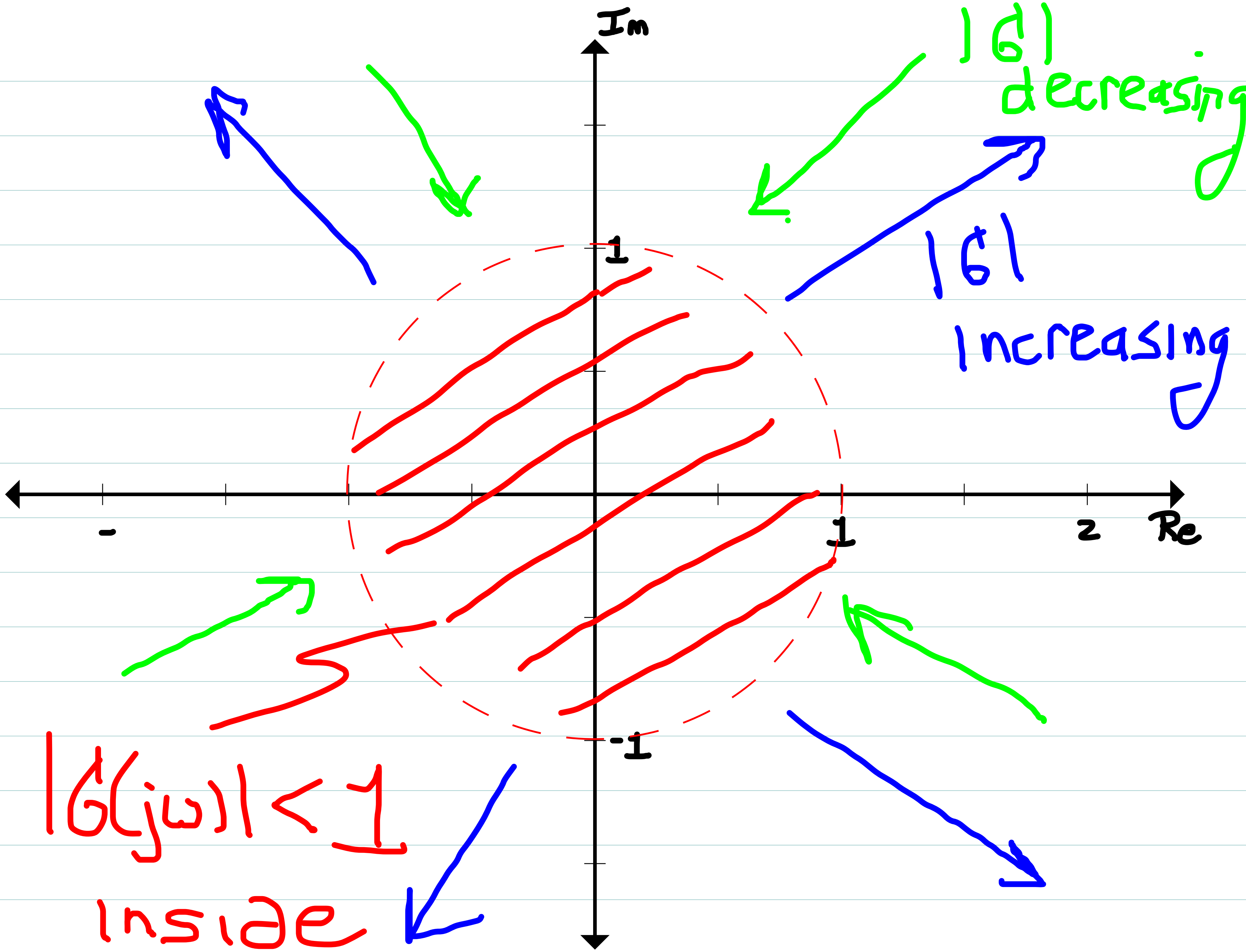
⇒ Bode diagrams show us the <u>polar coordinates</u> of the points $G(j\omega)$ for each ω

⇒ To map from Bode to polar

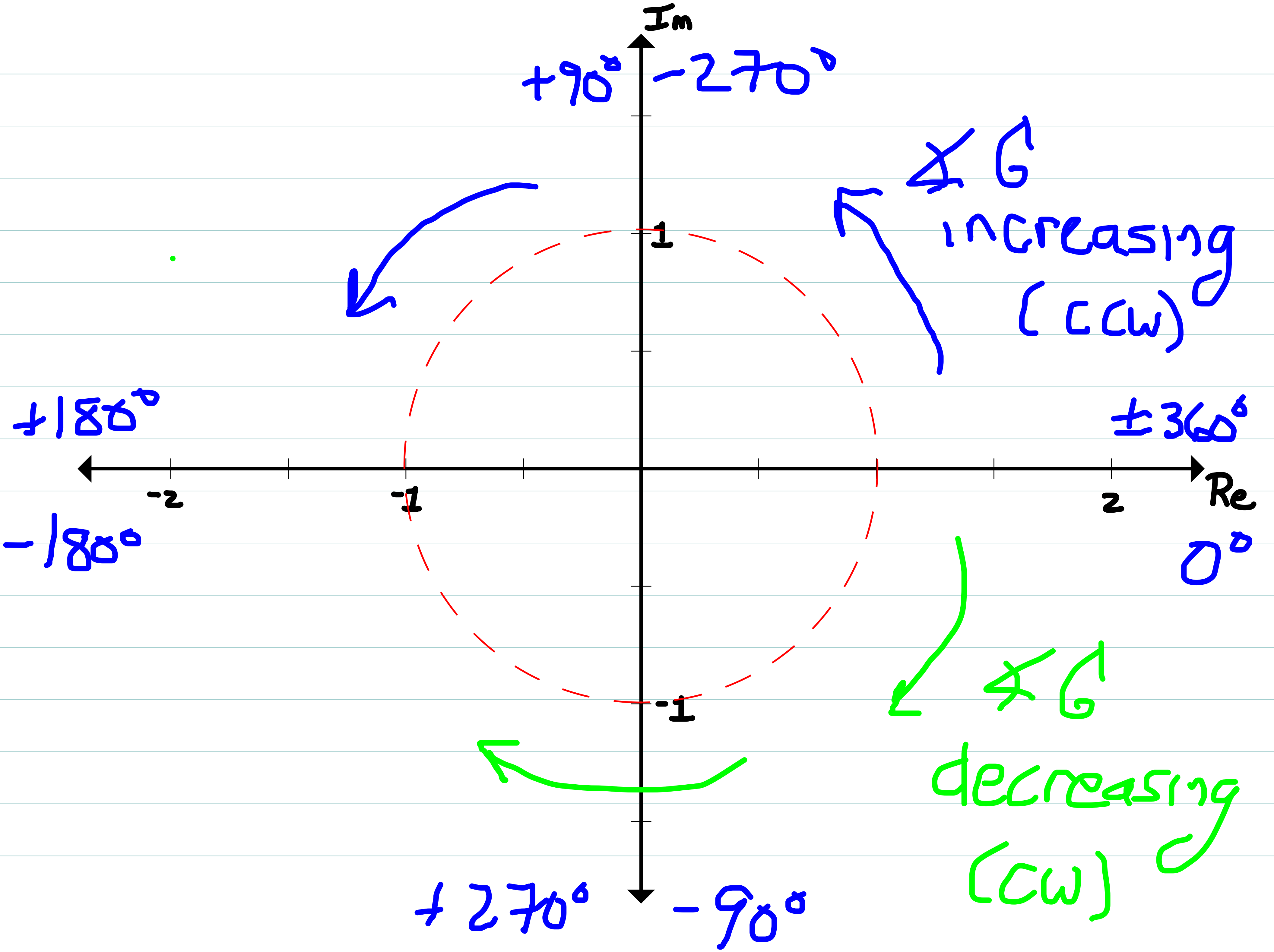
1.) Remember to convert magnitudes from dB back to actual.

2.) Remember angle convention for complex numbers.


Im
Re
+90° -270°
$\frac{3}{2}\measuredangle +90°$ or $-270°$
$\frac{3}{2}\measuredangle \pm 180°$
unit circle
+180°
-180°
±360°
0°
-2
-1
1
2
-1
$\frac{1}{2}\measuredangle -90°$
or $\frac{1}{2}\measuredangle +270°$
+270°
-90°



Im
Re
1
-1
1
2
-
|G| decreasing
|G| Increasing
|G(jω)| < 1
inside



Im
+90°
-270°
1
∡G
increasing
(CCW)
+180°
±360°
-2
-1
2
Re
-180°
0°
-1
∡G
decreasing
(CW)
+270°
-90°

A simple Example

$$G(s) = \frac{K_B}{(1+Ts)} \qquad T > \emptyset \text{ (min phase)} \qquad K_B > 1$$

Alway start by thinking about low/high freq. limiting behavior:

for $\omega << \frac{1}{|T|}$: Mag slope =

Phase =

for $\omega >> \frac{1}{|T|}$: Mag slope =

Phase =

A simple Example

$$G(s) = \frac{K_B}{(1+Ts)} \qquad T > \emptyset \text{ (min phase)} \quad K_B > 1$$

Always start by thinking about low/high freq. limiting behavior:

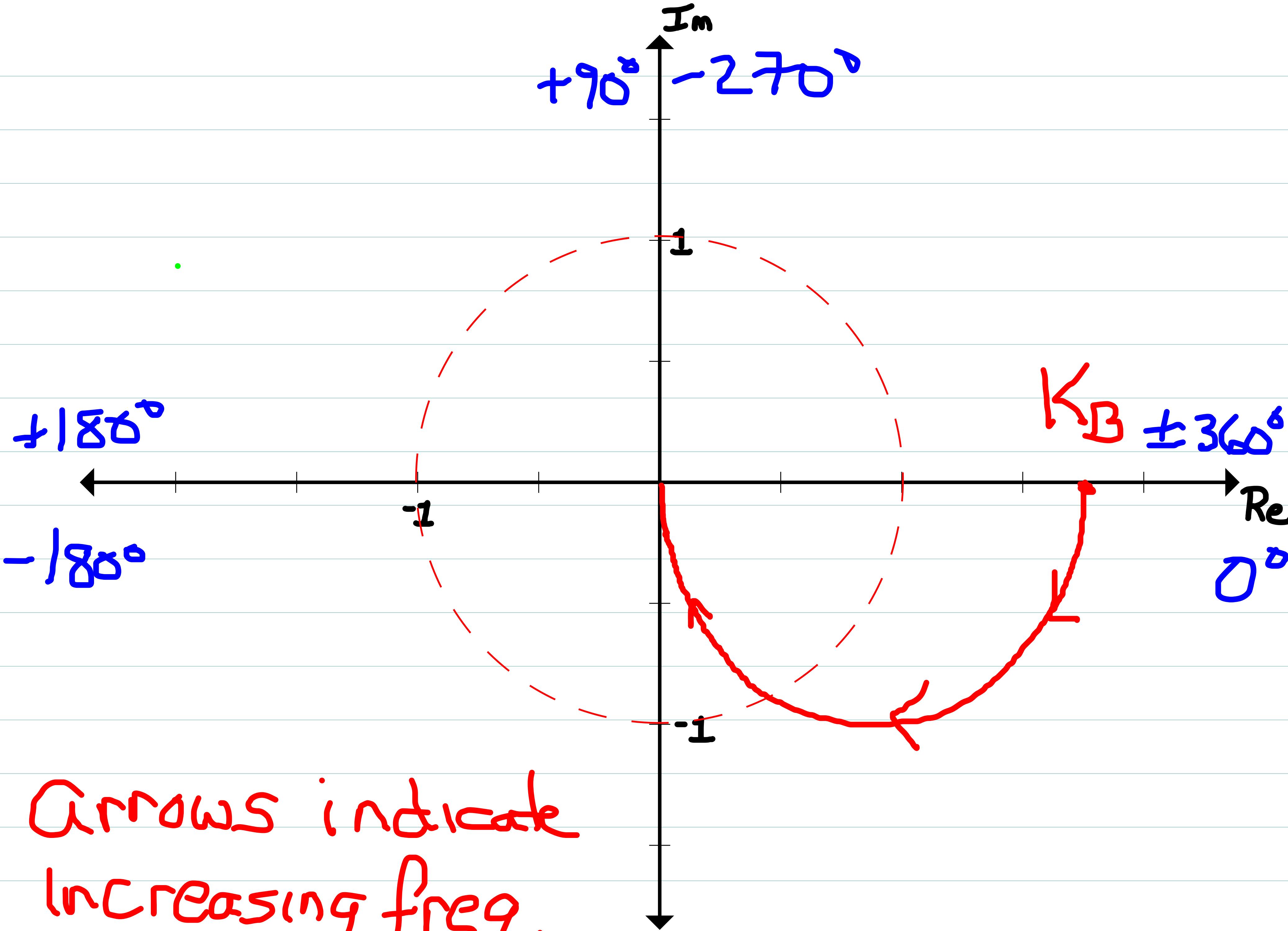
for $\omega \ll \frac{1}{|T|}$: Mag slope = 0 dB/dec (constant)

Phase = 0° (constant)

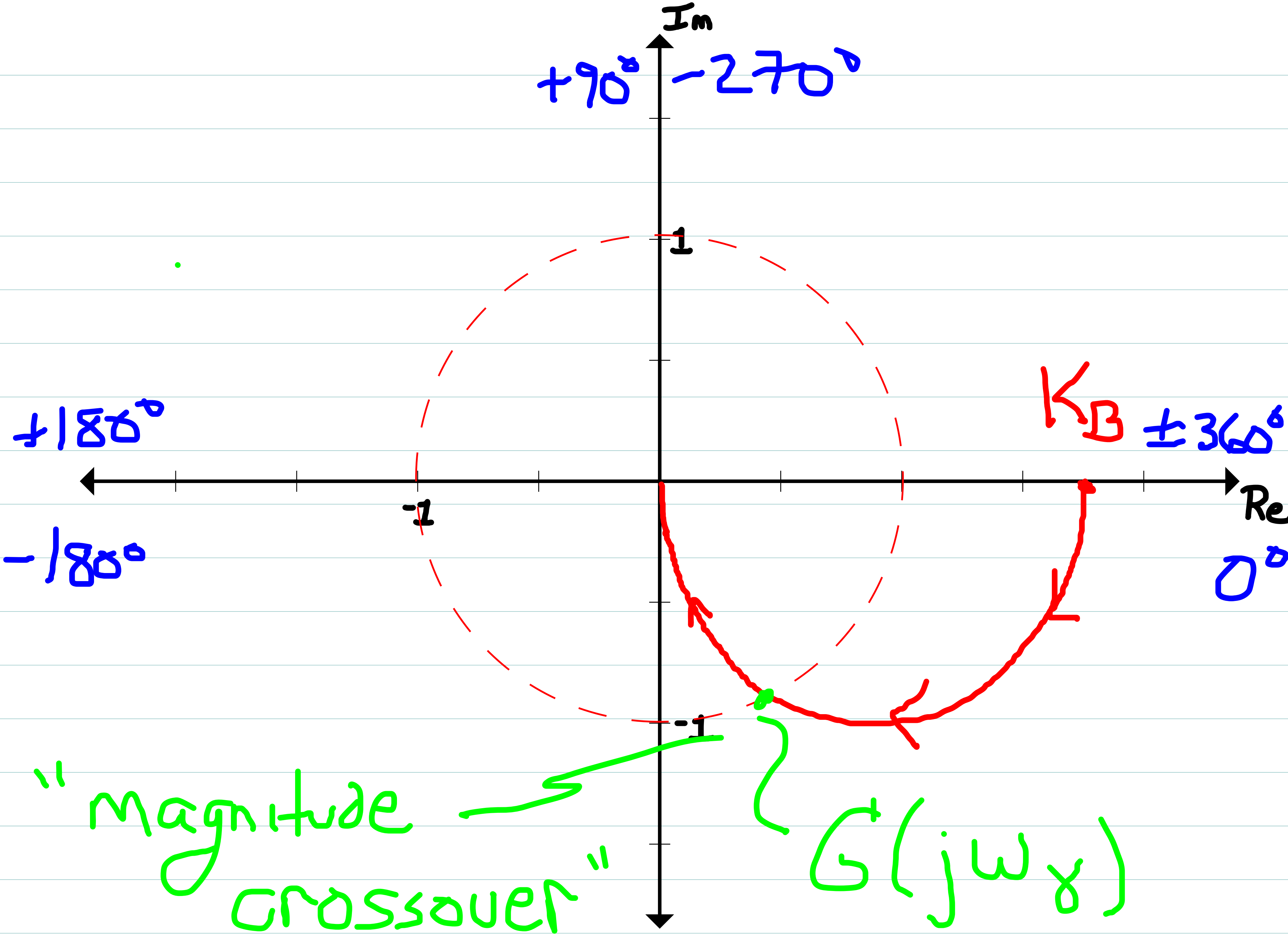
for $\omega \gg \frac{1}{|T|}$: Mag slope = -20 dB/dec

Phase = -90°

Low freq. magnitude is $|K_B| > 1$, high freq. magnitude is $\emptyset$: $\lim_{\omega \to \infty} |G(j\omega)| = \emptyset$



Arrows indicate increasing freq.



Im
Re
+90°
-270°
+180°
-180°
±360°
0°
1
-1
-1
-1
K_B
"magnitude crossover"
$G(j\omega_\gamma)$

Magnitude Crossover

"Magnitude crossover" occurs where polar plot "punctures" the unit circle

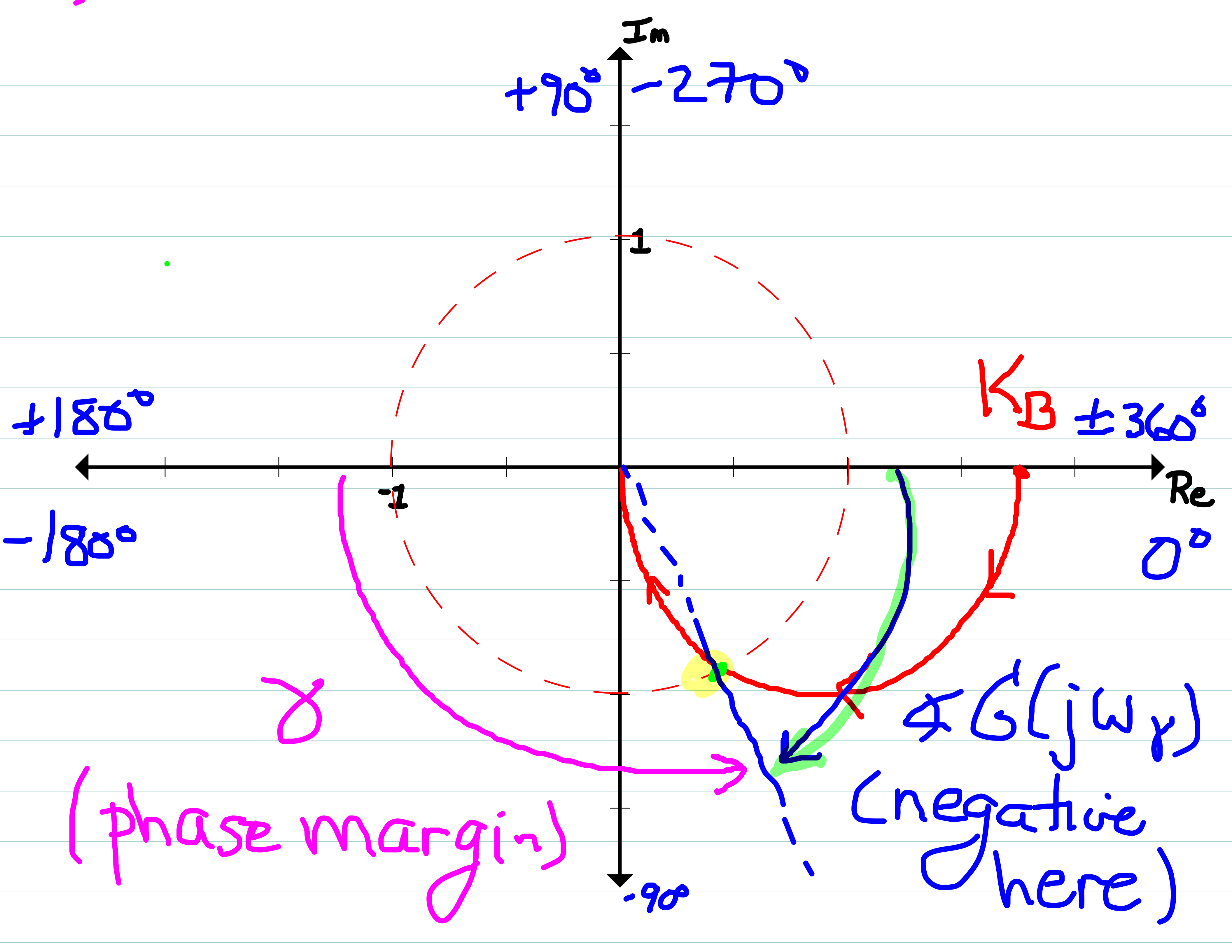
$|G(j\omega)| = 1$ at this point.

The frequency at which this occurs is the "magnitude crossover freq", termed ω_γ

Easily seen on Bode: $\boxed{\omega_\gamma \text{ is the frequency where}}$
$$|G(j\omega)| = \underline{\emptyset\ dB}$$

Note: depending on the system there may be one, many, or no magnitude xover freq.

Important quantity: $\measuredangle G(j\omega_\gamma)$: phase at magnitude xover



Im
+90°
-270°
1
+180°
-180°
-1
Re
±360°
0°
-90°
γ
(phase margin)
∡G(jω_γ)
(negative here)

Phase Margin

The **phase margin** is the angle around the unit circle from -1 to magnitude crossover point, measured positive **counter-clockwise** from -1 (or, equiv, CW from mag xover to -1)

The phase margin angle, γ, is expressed in deg (although later it will be convenient to express in rad).

Assuming we write $\measuredangle G(j\omega_\gamma)$ in range $[0°, -360°]$ an expression for γ is:

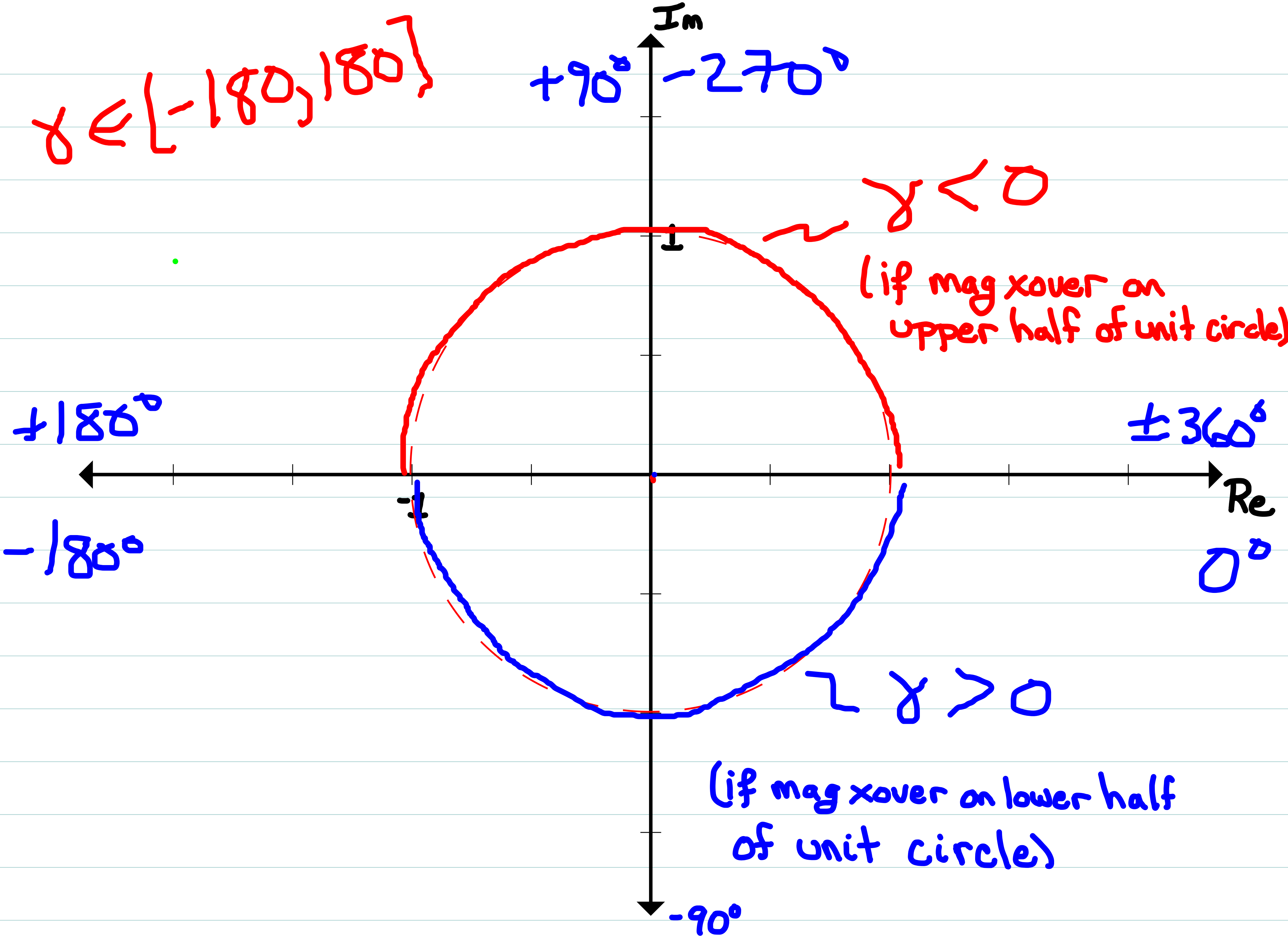
$\gamma \in [-180°, 180°]$

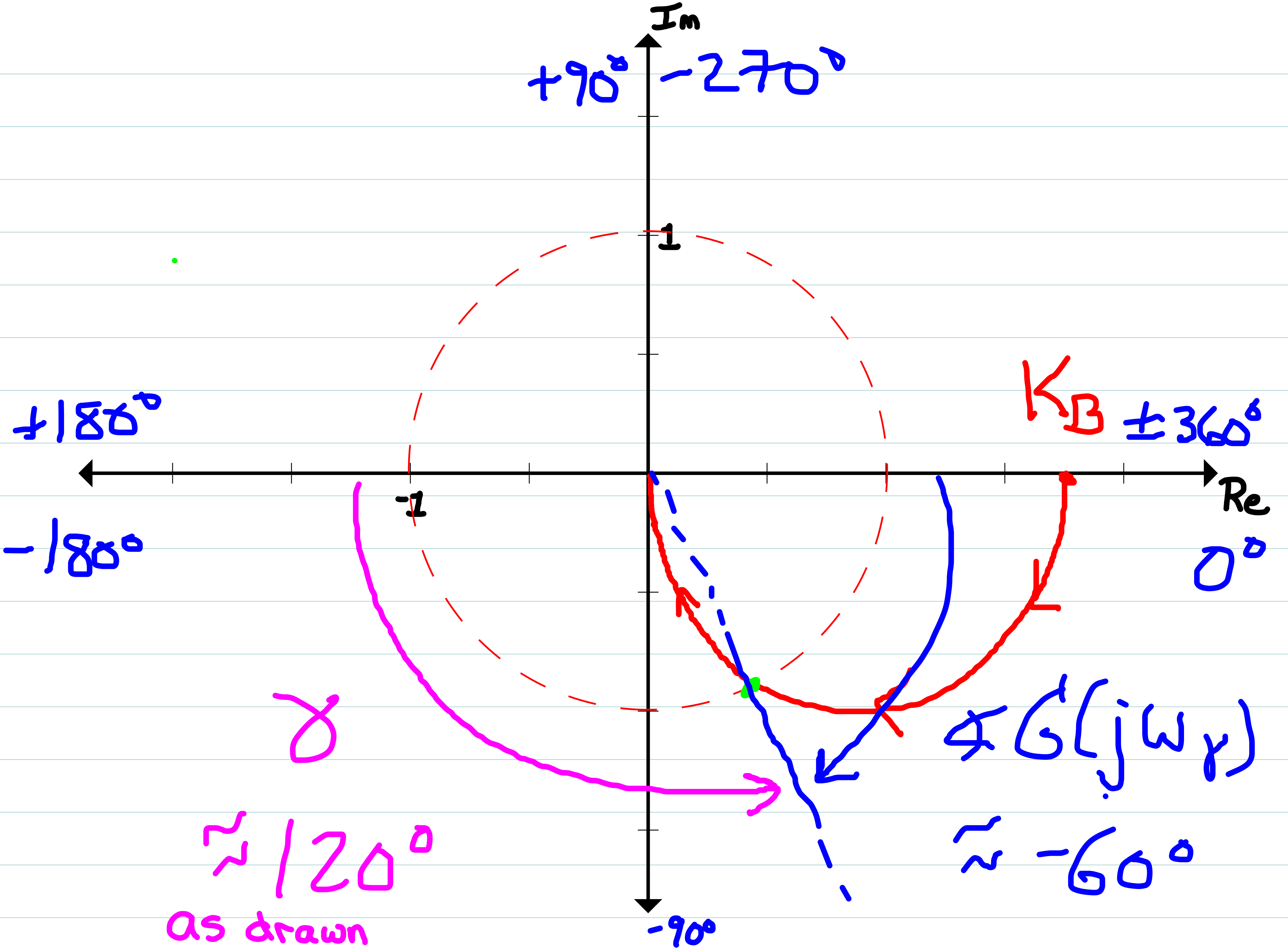
$$\gamma = 180° + \measuredangle G(j\omega_\gamma)$$

$\Rightarrow \gamma > 0$ if $\measuredangle G(j\omega_\gamma) > -180°$

<u>Note</u>: Matlab will usually try to wrap phase plot $\measuredangle G(j\omega)$ so that $\measuredangle G(j\omega_\gamma)$ is in this range. Sometimes it doesn't. You can always manually add or subtract a multiple of $360°$ to get $\measuredangle G(j\omega_\gamma)$ in this range.

$\gamma \in [-180, 180]$




Im
+90°
-270°
1
+180°
-180°
-1
K_B
±360°
Re
0°
γ
≈ 120°
as drawn
∡G(jω_γ)
≈ -60°
-90°

Another Example

$$G(s) = \frac{K_B}{(1+\tau s)^3} \qquad K_B > 1 \qquad \tau > 0$$

Low freq mag: Constant at K_B

Low freq. phase: Constant at $0°$

High freq. mag slope:

High freq. phase:

Another Example

$$G(s) = \frac{K_B}{(1+\tau s)^3} \qquad K_B > 1 \quad \tau > \emptyset$$

Low freq mag: Constant at K_B

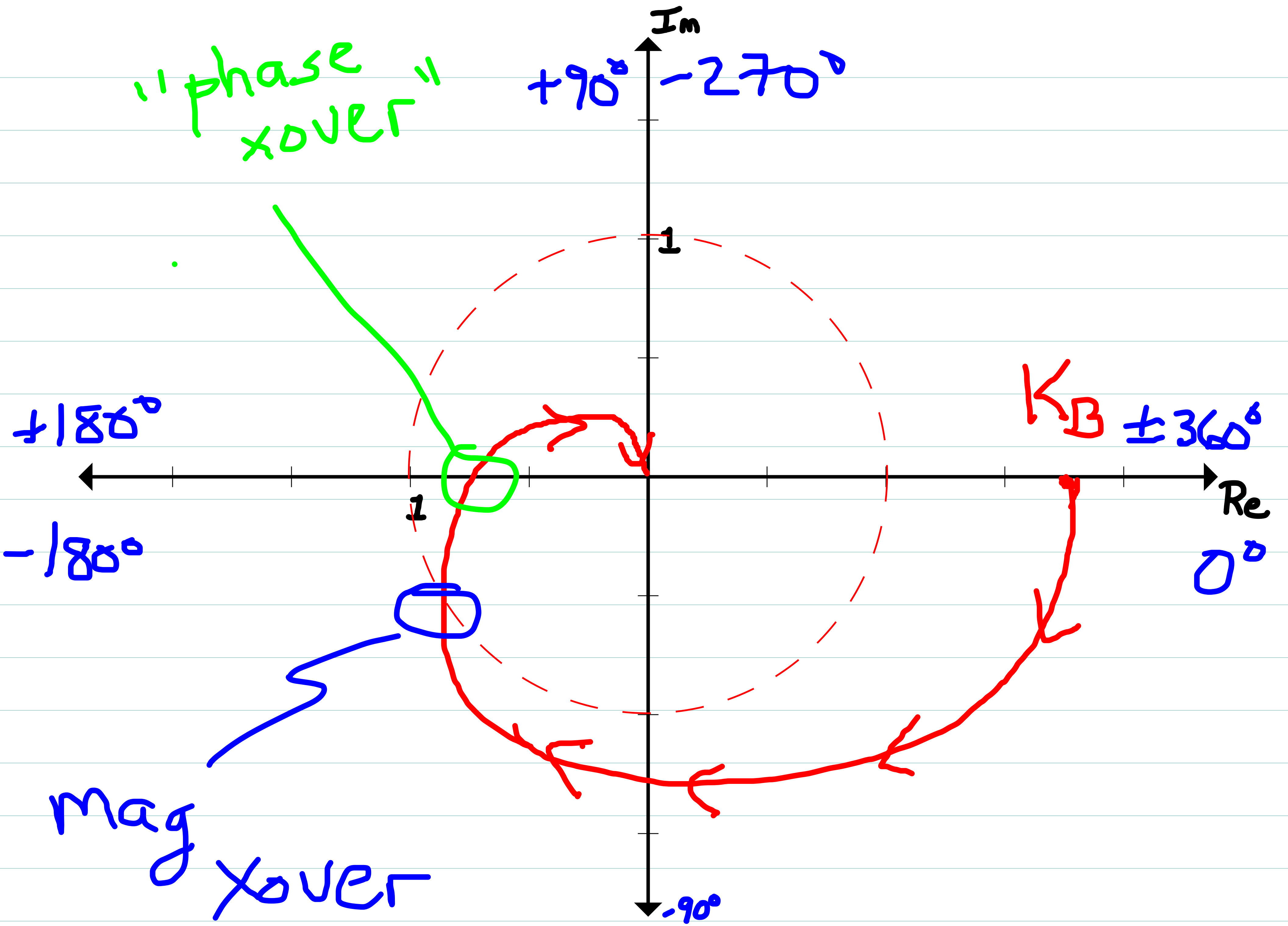
Low freq. phase: Constant at $0°$

High freq. mag slope: -60 dB/dec

High freq. phase: $-270°$

Recall: negative high freq. slope means

$$|G(j\omega)| \to \emptyset \quad \text{as} \quad \omega \to \infty$$



"phase xover"
Im
+90°
-270°
1
+180°
-180°
1
Re
K_B
±360°
0°
-90°
mag xover

Phase Crossover

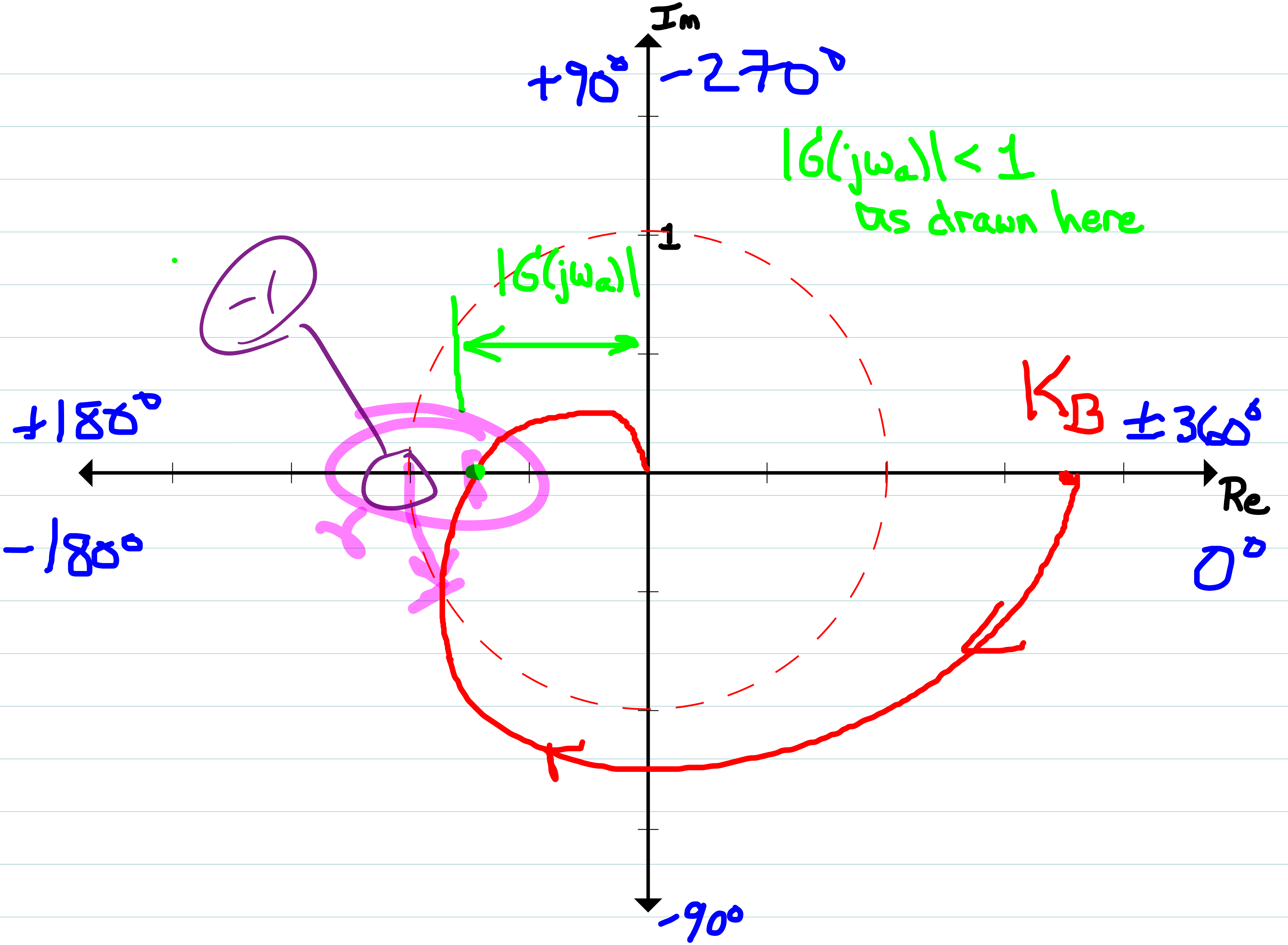
The "phase crossover" of a polar plot is the point where the plot crosses through the negative real axis.

This corresponds to the point where $\measuredangle G(j\omega) = -180°$

Again, easily seen from Bode phase diagrams: call ω_a "phase crossover freq." the value of ω for which $\measuredangle G(j\omega) = -180°$.

Note: May be one, none, or many ω_a depending on system.

Important quantity: $|G(j\omega_a)|$ magnitude at phase xover frequency



Im
+90°
-270°
|G(jωa)| < 1
as drawn here
1
|G(jωa)|
-1
+180°
-180°
KB
±360°
Re
0°
-90°

Gain Margin

The gain margin, a, is defined as:

$$\boxed{a = \frac{1}{|G(j\omega_a)|}}$$

Gain margin is commonly expressed in dB:

$$a_{dB} = 20 \log a$$

$$= -|G(j\omega_a)|_{dB}$$

So gain margin in dB is negative of Bode magnitude at phase crossover freq.

Meaning of Gain and phase margins

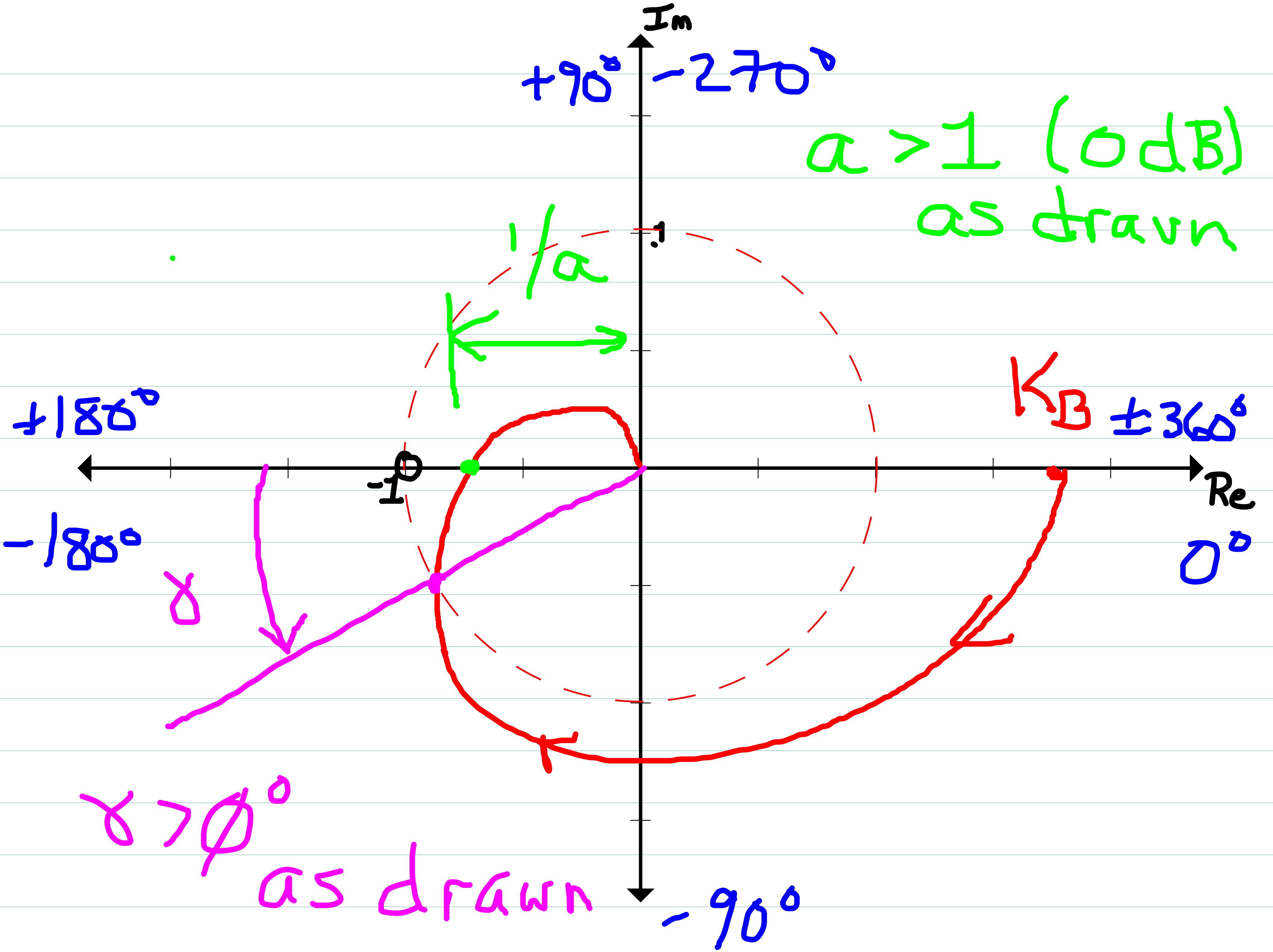
a, γ measure how close polar plot comes to point $-1+\emptyset j$ ("-1 point") in complex plane. Recall $-1+\emptyset j = 1 \angle -180°$

Two "pseudo-orthogonal" directions

→ a measures distance to -1 along real Axis as a ratio $1/|G(j\omega_a)|$

⇒ γ measures distance to -1 as an angle around unit circle.

<u>Note</u>: $a > 1$ ($a > \emptyset$ dB) means phase crossover occurs <u>inside</u> unit circle. $a < 1$ ($a < \emptyset$ dB) means phase crossover is outside unit circle

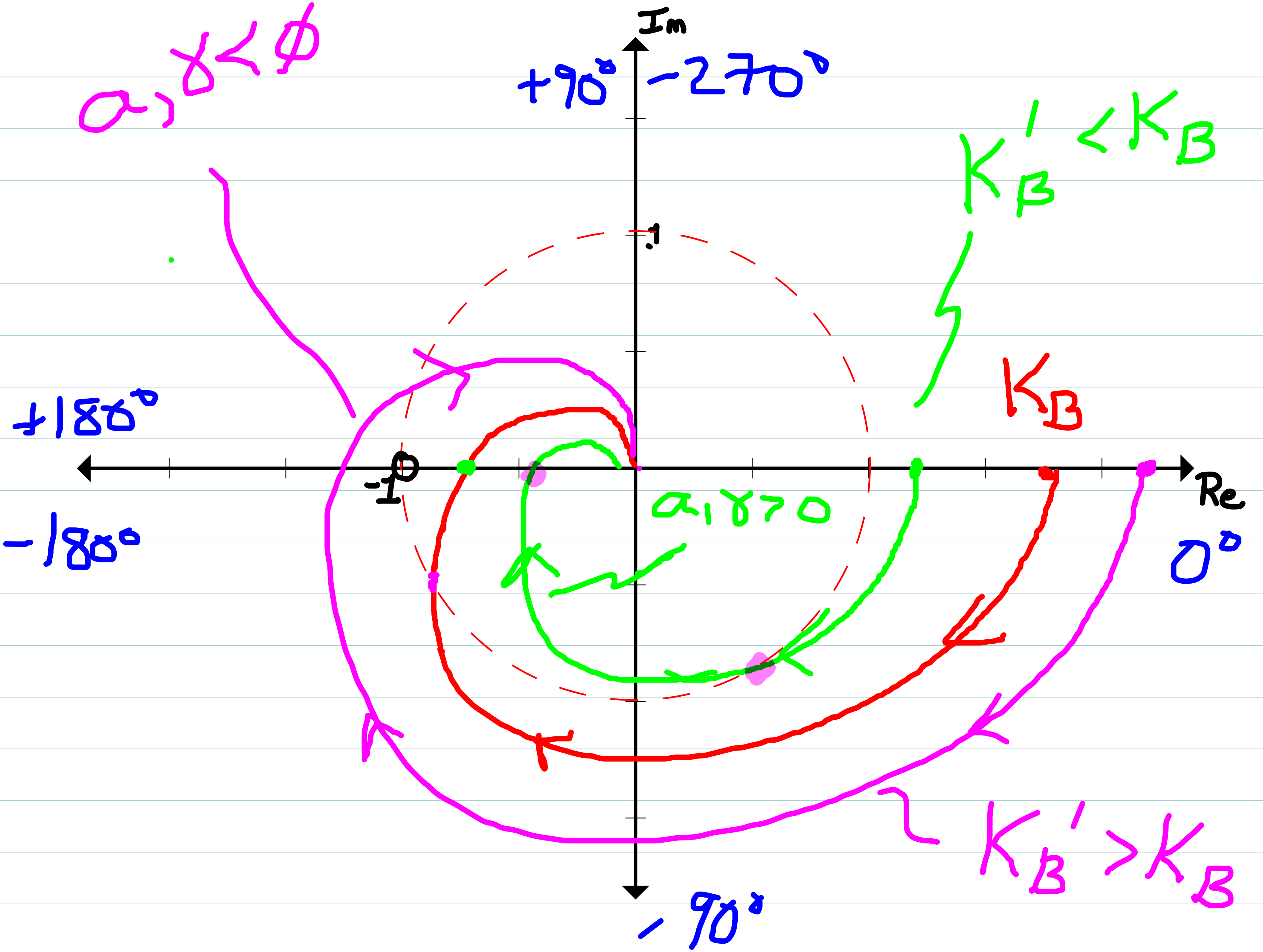


Im
+90°
-270°
a > 1 (0dB)
as drawn
1/a
1
+180°
-1
K_B
±360°
Re
-180°
0°
γ
γ > 0°
as drawn
-90°

Effect of Gain Changes

Increasing or decreasing K_B uniformly expands or contracts polar plot about the origin

$\Rightarrow$ Will generally change crossovers and margins



a, γ < 0
Im
+90°
-270°
K_B' < K_B
1
K_B
+180°
-1
-180°
Re
0°
a, γ > 0
-90°
K_B' > K_B

Effect of Zeros

Since they affect magnitude and phase, zeros will change shape of polar plot.

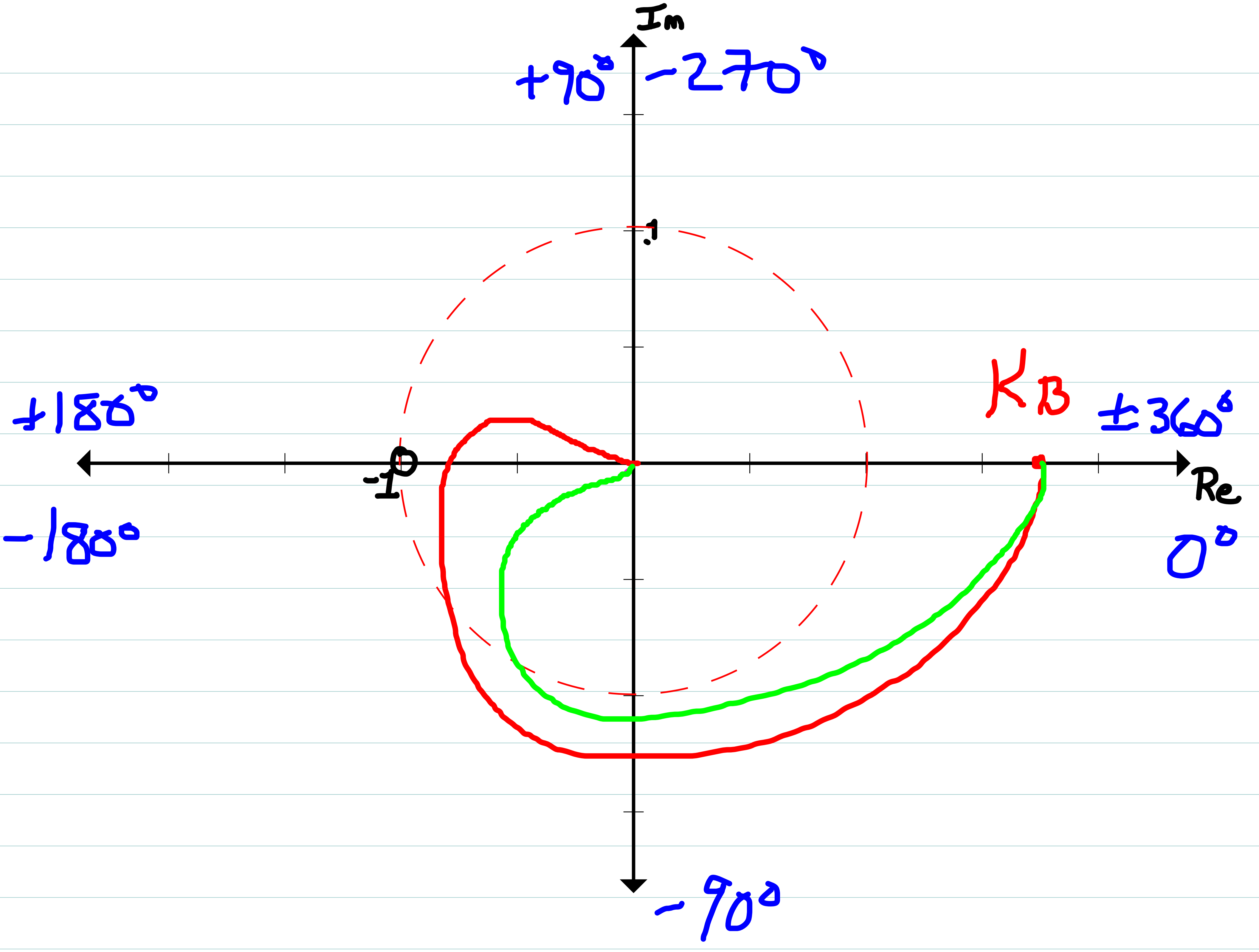
Example:

$$G(s) = K_B \frac{(T_1 s + 1)}{(T_2 s + 1)^3} \qquad K_B > 1 \quad T_1, T_2 > \emptyset$$

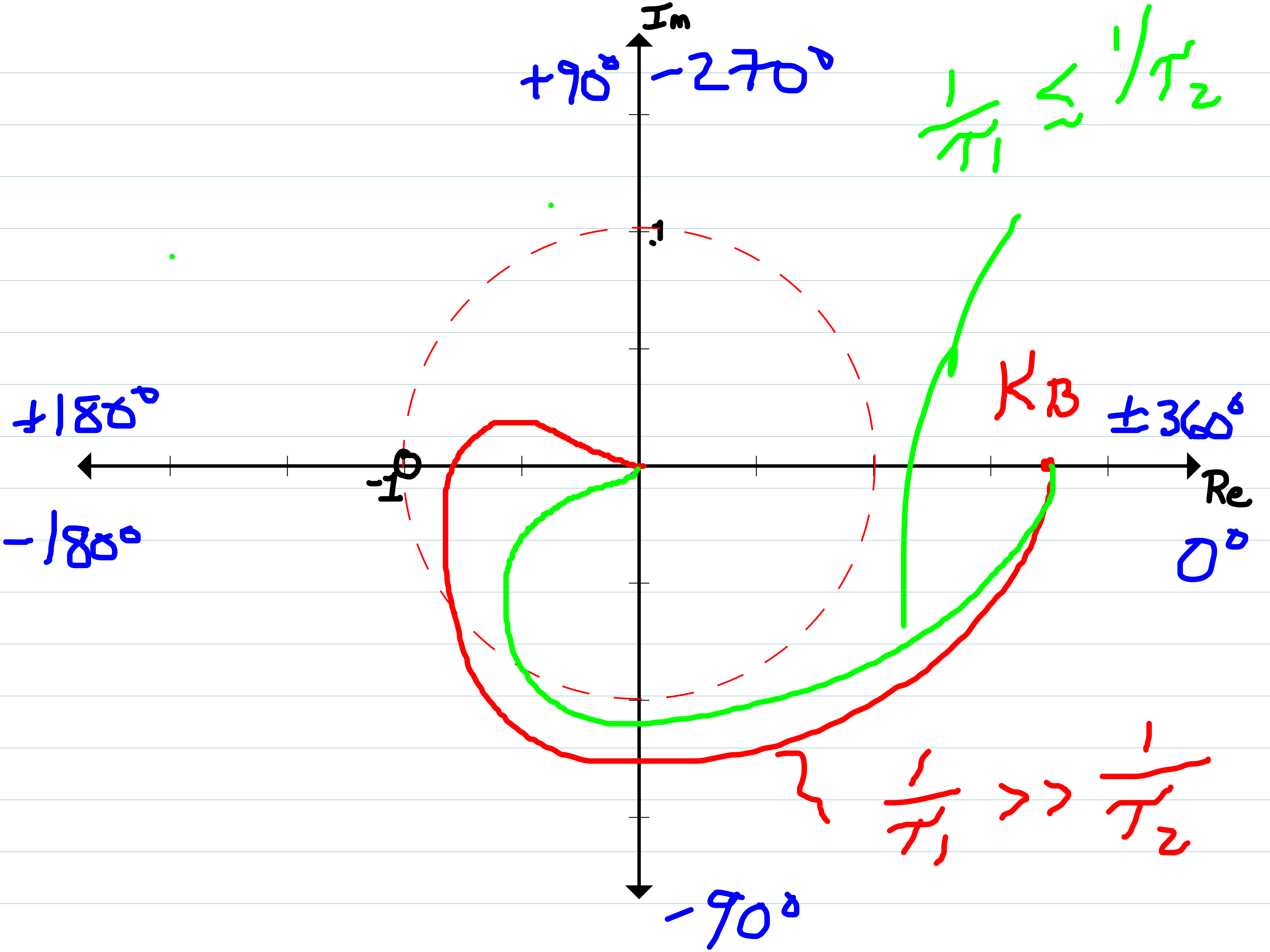
high freq phase: $-180°$ here (Why??)

But this limit may be asymptotically approached from above or below as $\omega \to \infty$

This difference can have a profound impact on shape of plot. Need to check Bode for accuracy, but can often "reason it out" for simple cases.

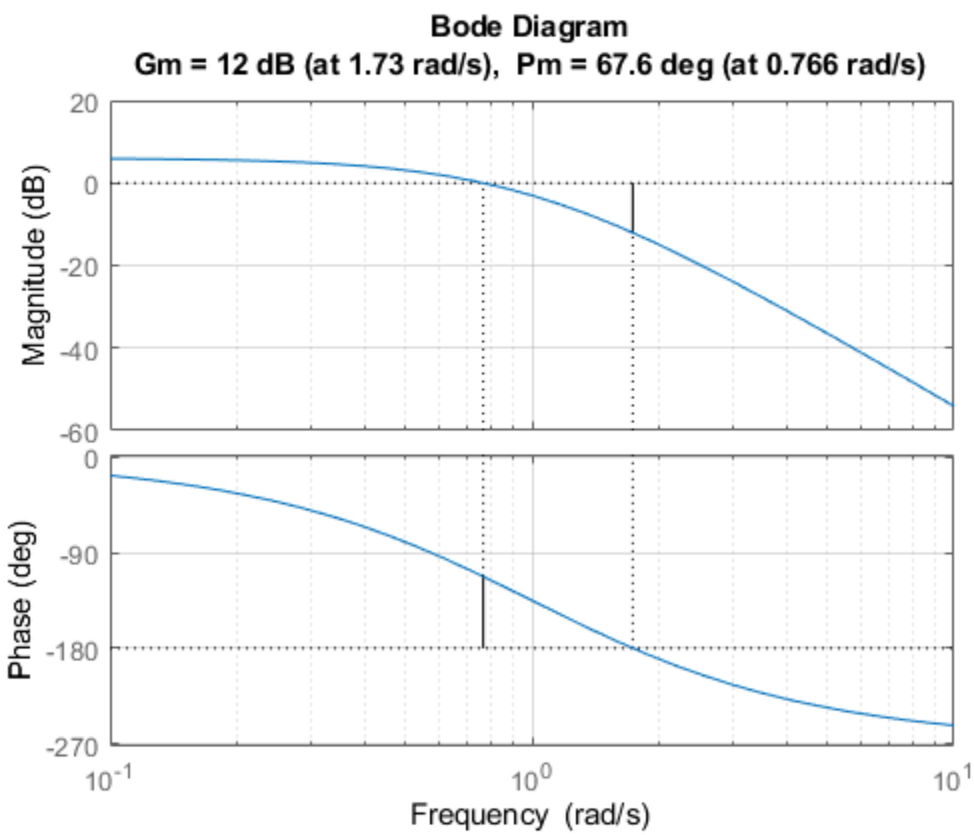


Im
+90°
-270°
1
+180°
-180°
-1
KB
±360°
0°
Re
-90°

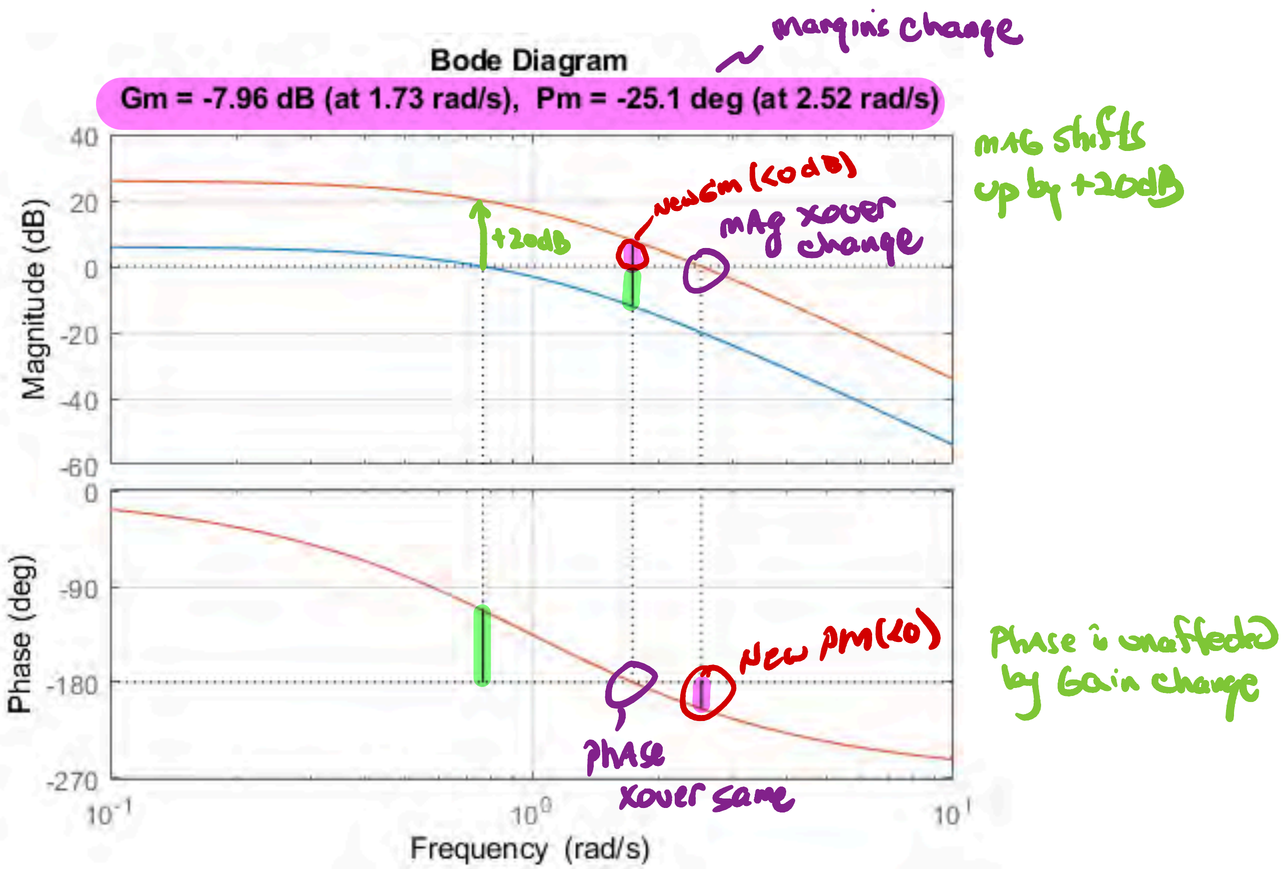


Im
Re
+90°
-270°
±180°
-180°
±360°
0°
-90°
1
-1
K_B
$1/\tau_1 \lesssim 1/\tau_2$
$1/\tau_1 >> 1/\tau_2$

$$G(s) = \frac{2}{(s+1)^3}$$


Bode Diagram
Gm = 12 dB (at 1.73 rad/s), Pm = 67.6 deg (at 0.766 rad/s)
Magnitude (dB)
20
0
-20
-40
-60
Phase (deg)
0
-90
-180
-270
10^-1
10^0
10^1
Frequency (rad/s)

$$G'(s) = \frac{20}{(s+1)^3}$$ (K_B increased by factor of 10)



Poles at origin

Poles at origin will introduce a unique feature to a polar plot.

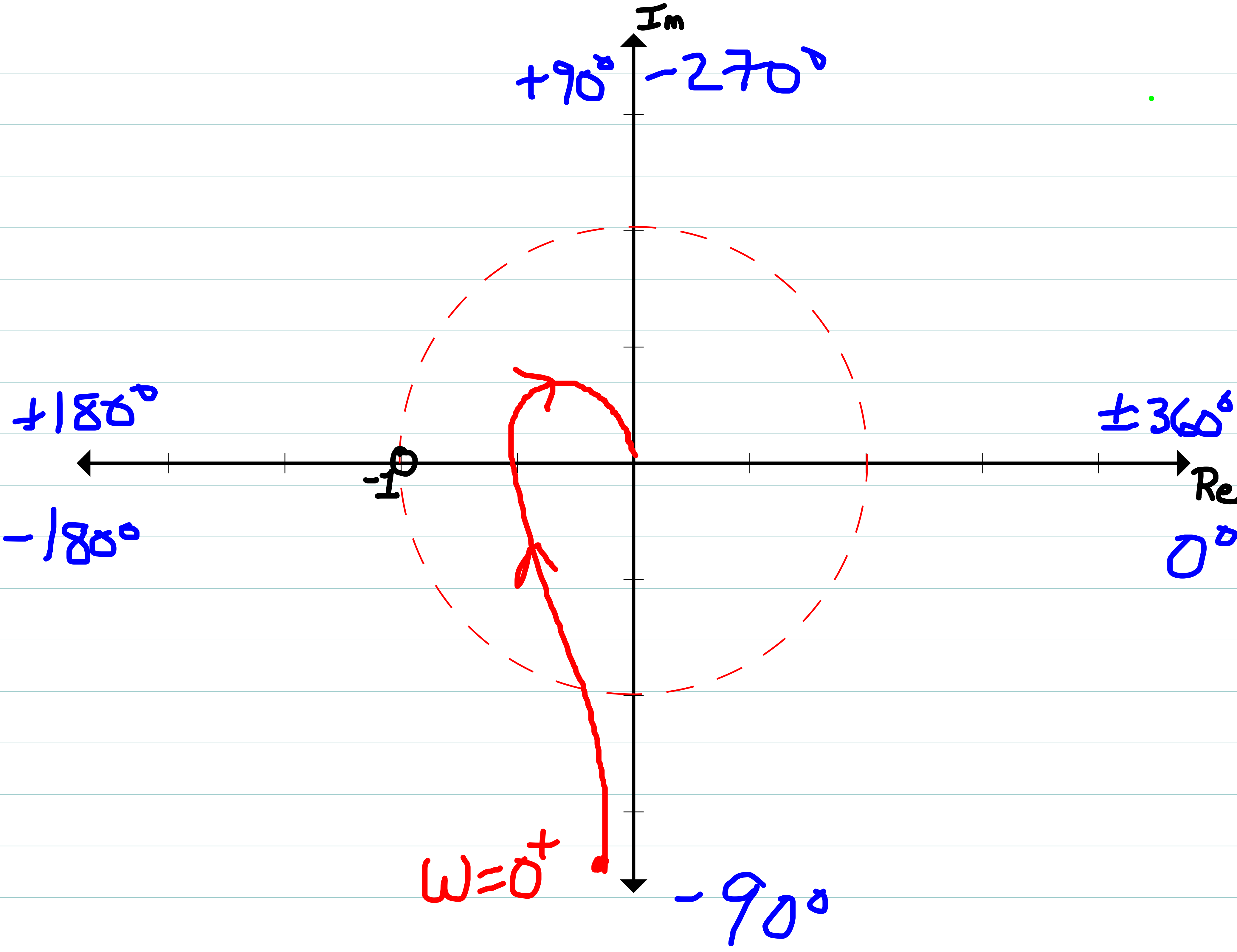
$$\lim_{\omega \to 0} \measuredangle G(j\omega) = \measuredangle K_B - N\,90^\circ$$

and $\displaystyle\lim_{\omega \to \phi} |G(j\omega)| = \infty$ in these cases

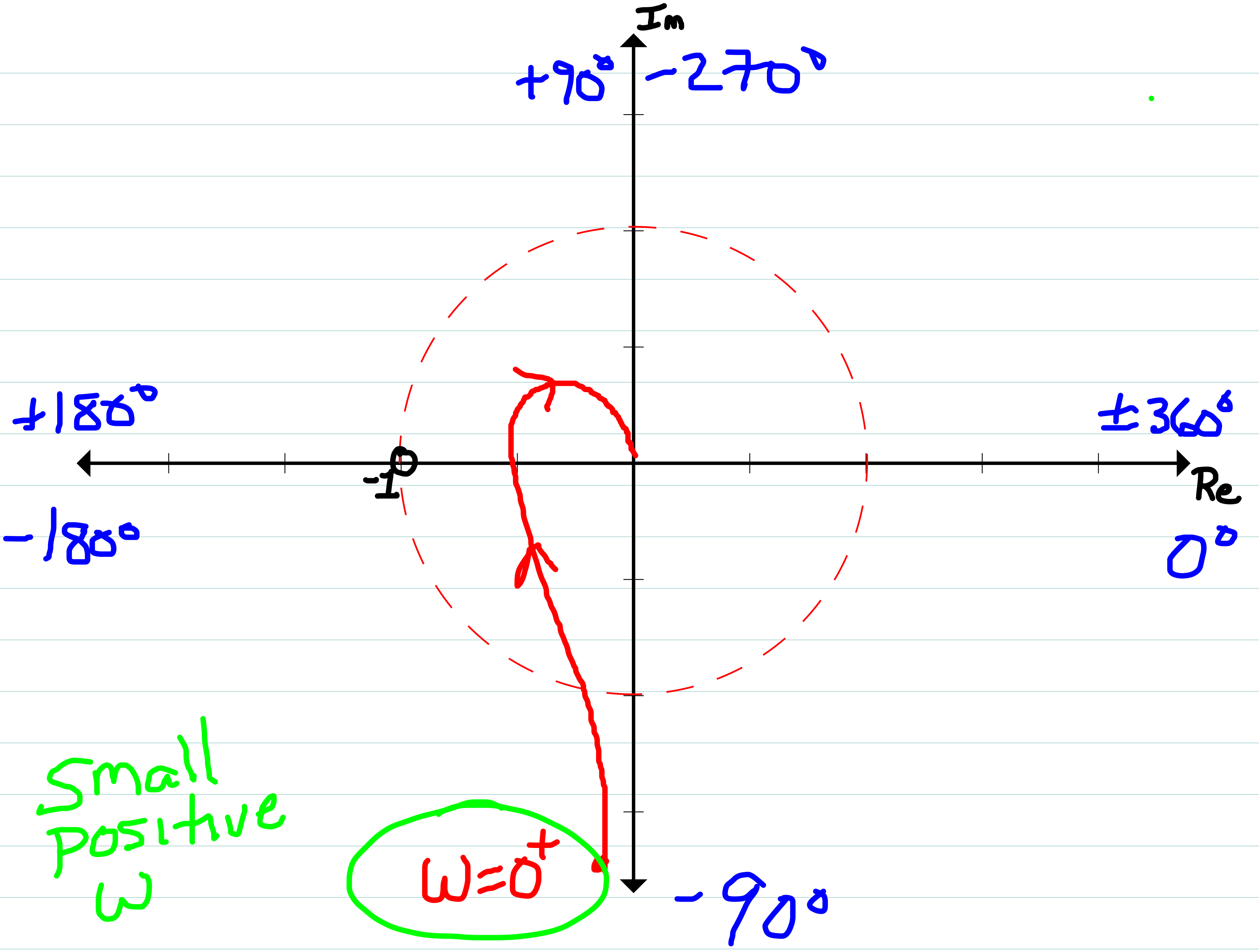
$\Rightarrow$ Polar plot will exhibit a "tail" along one of the coordinate axes.

Example:

$$G(s) = \frac{K_B}{s(\tau s+1)^2} \qquad \tau, K_B > \phi$$



Im
+90°
-270°
+180°
-180°
±360°
0°
-1
Re
-90°
ω=0+


Im
+90°
-270°
+180°
±360°
-180°
0°
-1
Re
Small
Positive
ω
ω=0+
-90°

Note: Which side of a coordinate axis the tail lies on is sometimes important.

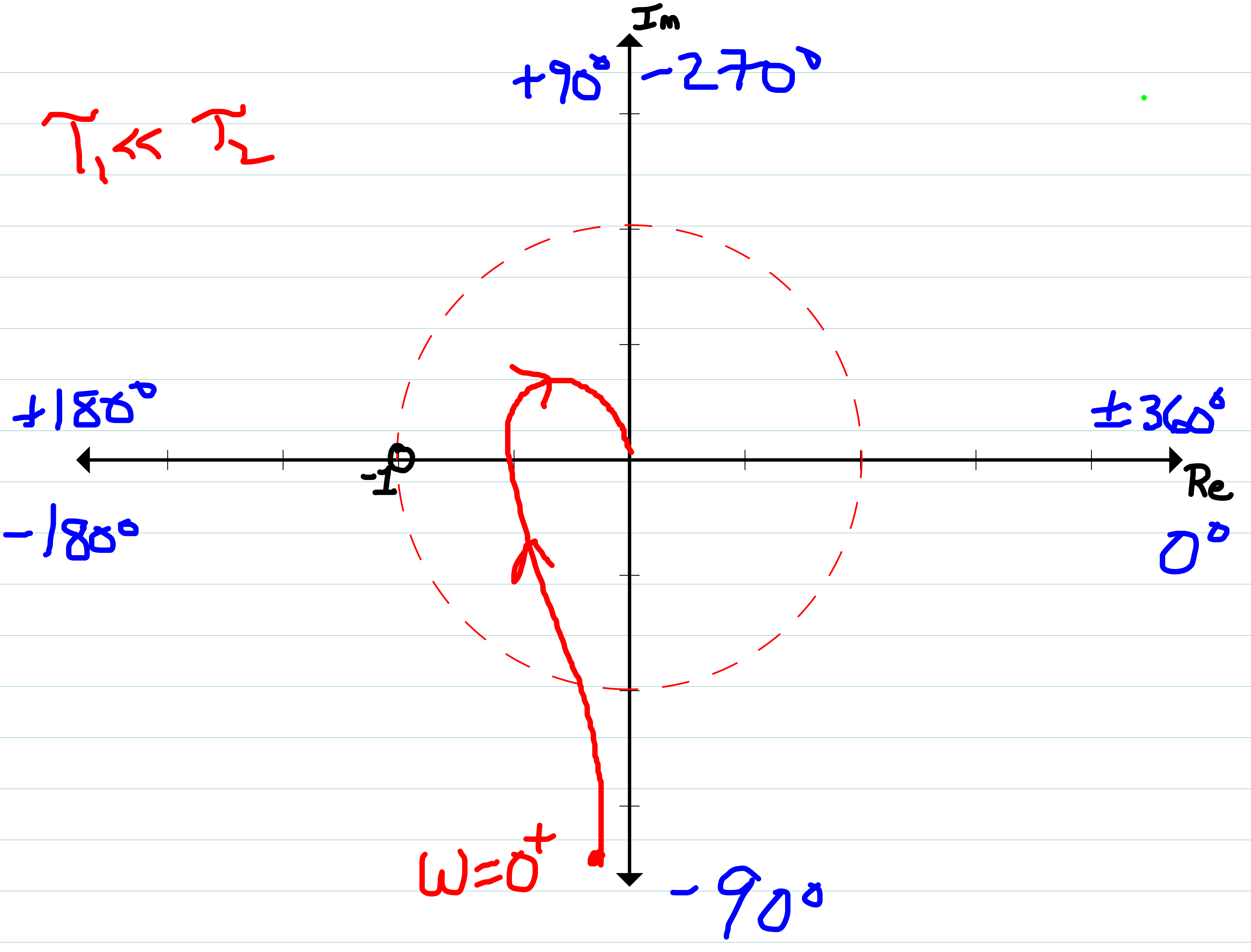
$\Rightarrow$ Determined by asymptotic behavior of phase as $\omega \to \phi$.

Example:

$$G(s) = K_B\left[\frac{(T_1 s+1)}{s(T_2 s+1)^3}\right]$$

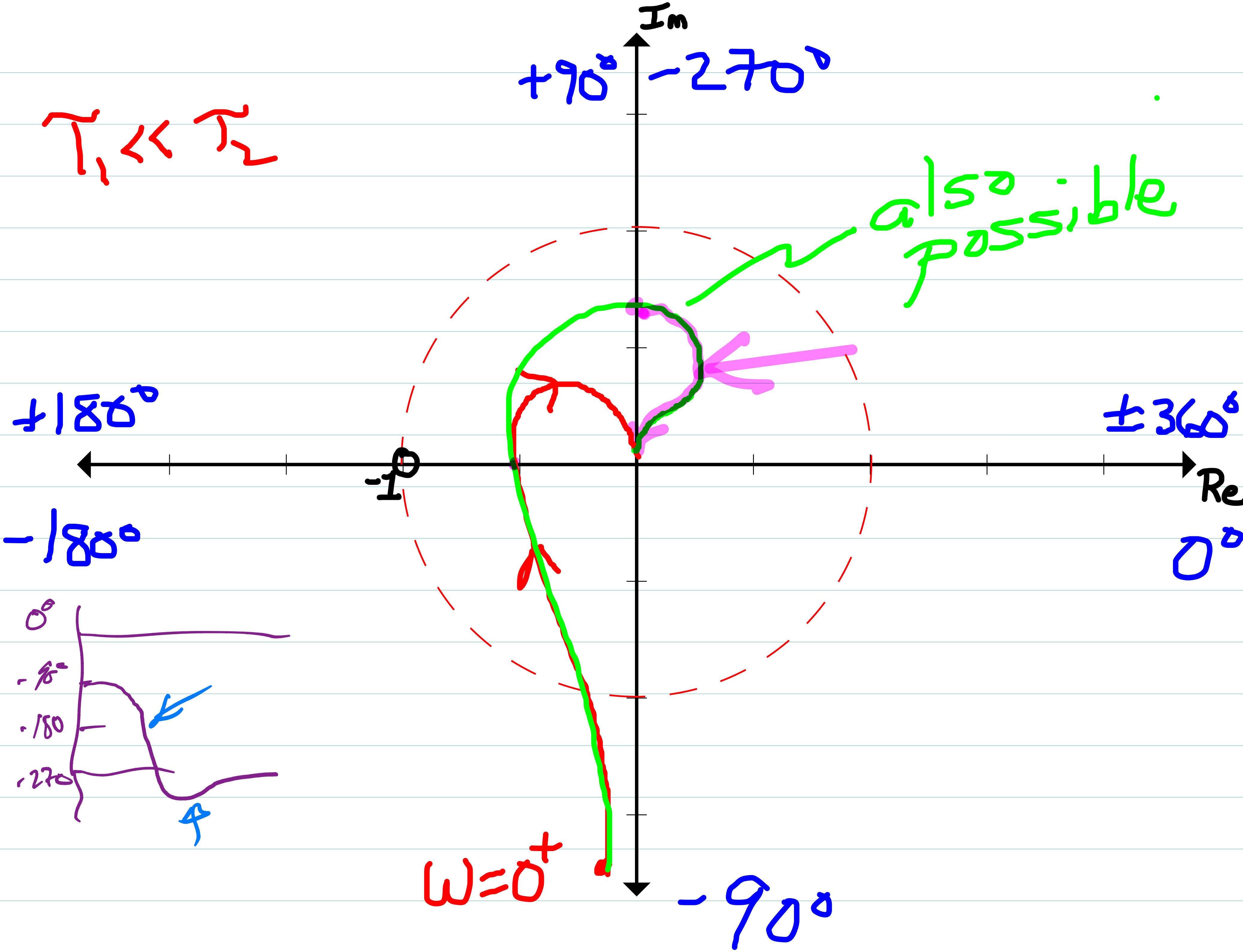
if $T_1 \ll T_2$ (so $\frac{1}{T_1} \gg \frac{1}{T_2}$) then as $\omega \to \phi$ phase approaches $-90°$ *from below* (equivalently, phase is decreasing as ω increases from ϕ).

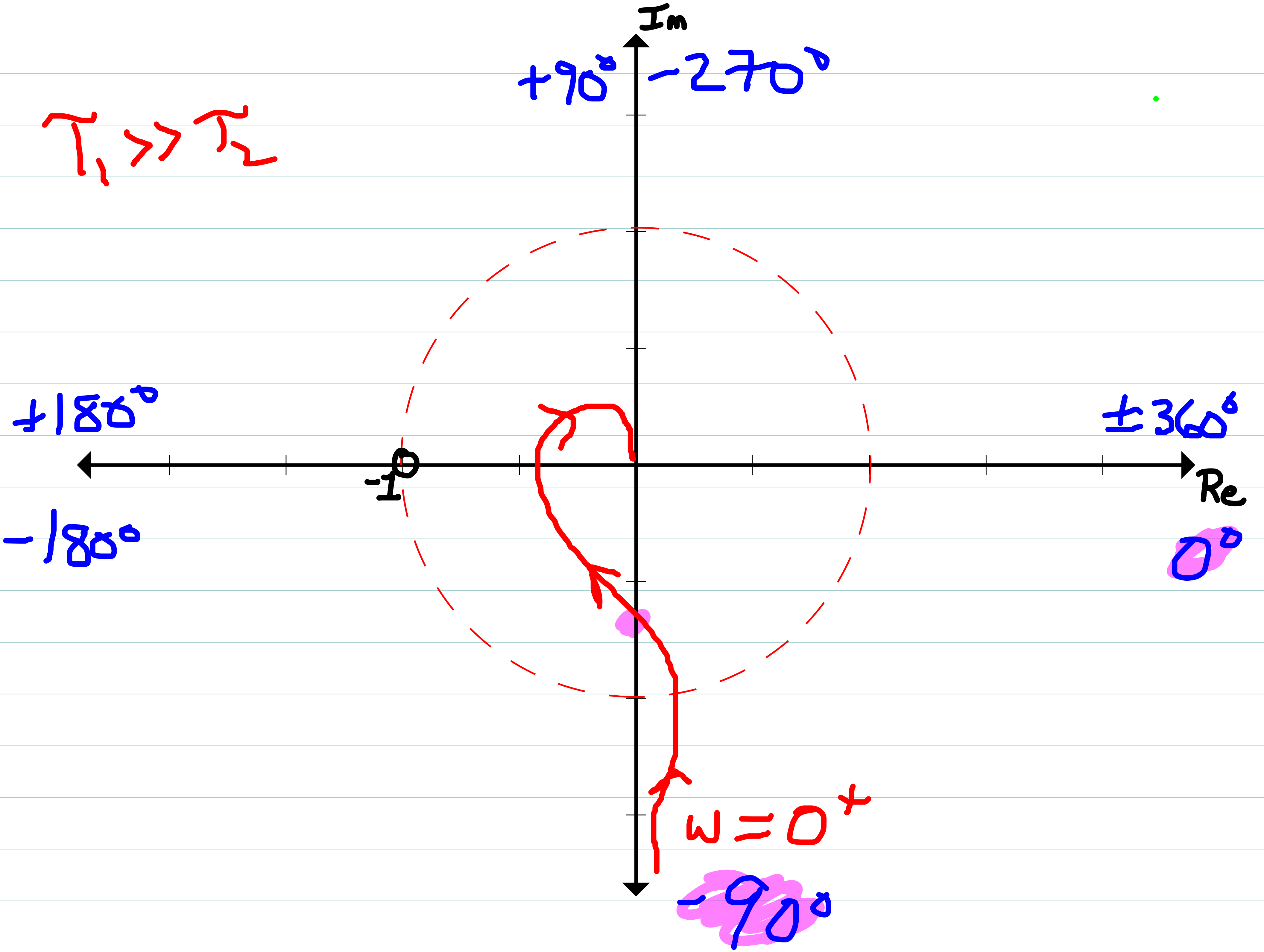
Conversely, if $T_1 \gg T_2$, phase approaches $-90°$ *from above* as $\omega \to 0$.



$T_1 \ll T_2$
Im
$+90°$
$-270°$
$+180°$
$-180°$
$\pm 360°$
$0°$
-1
Re
$\omega = 0^+$
$-90°$

$T_1 << T_2$



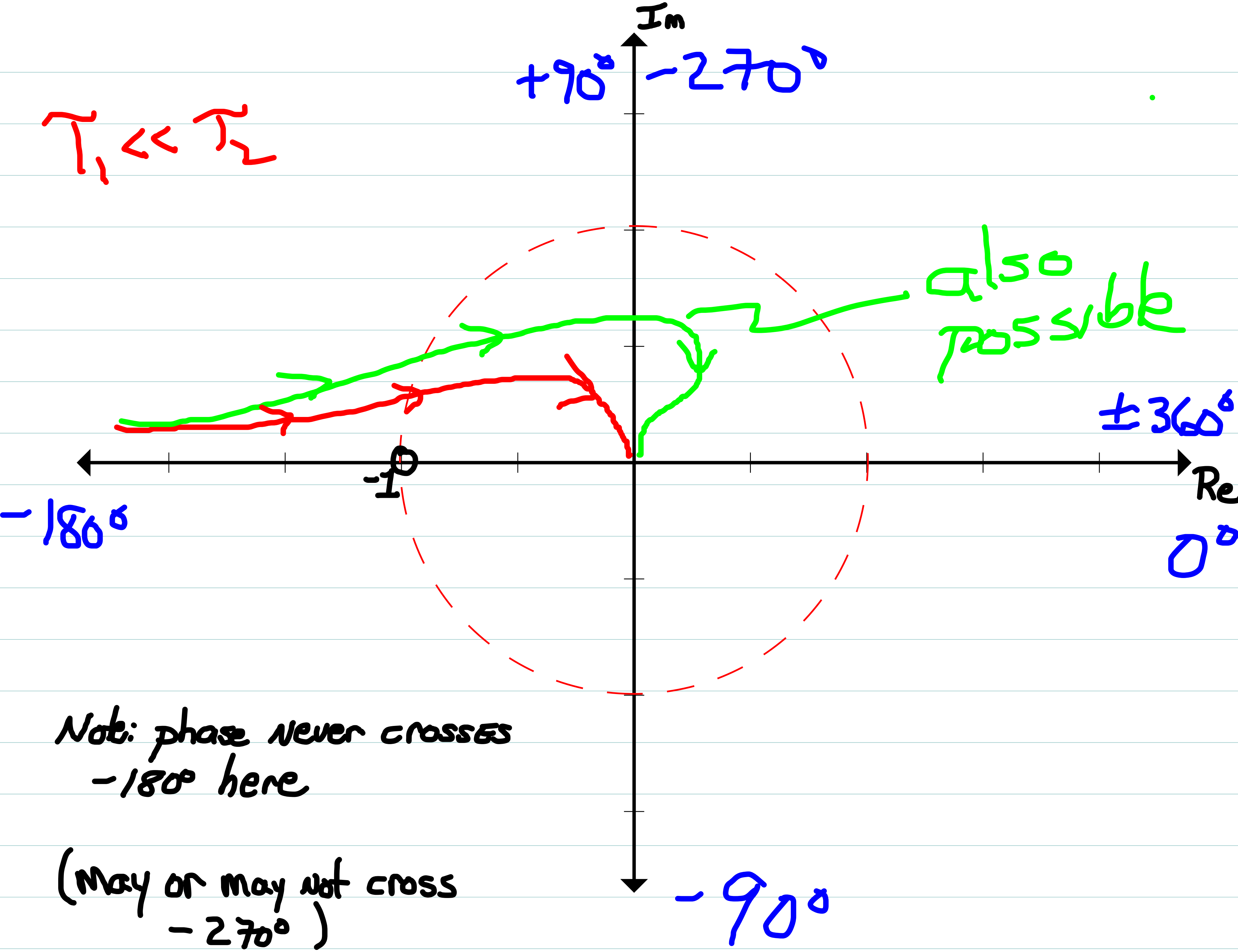


$T_1 >> T_2$
Im
+90°
-270°
+180°
-180°
±360°
0°
-1
Re
$\omega = 0^+$
-90°

Add'l poles at origin change the coordinate axis the tail lies along.

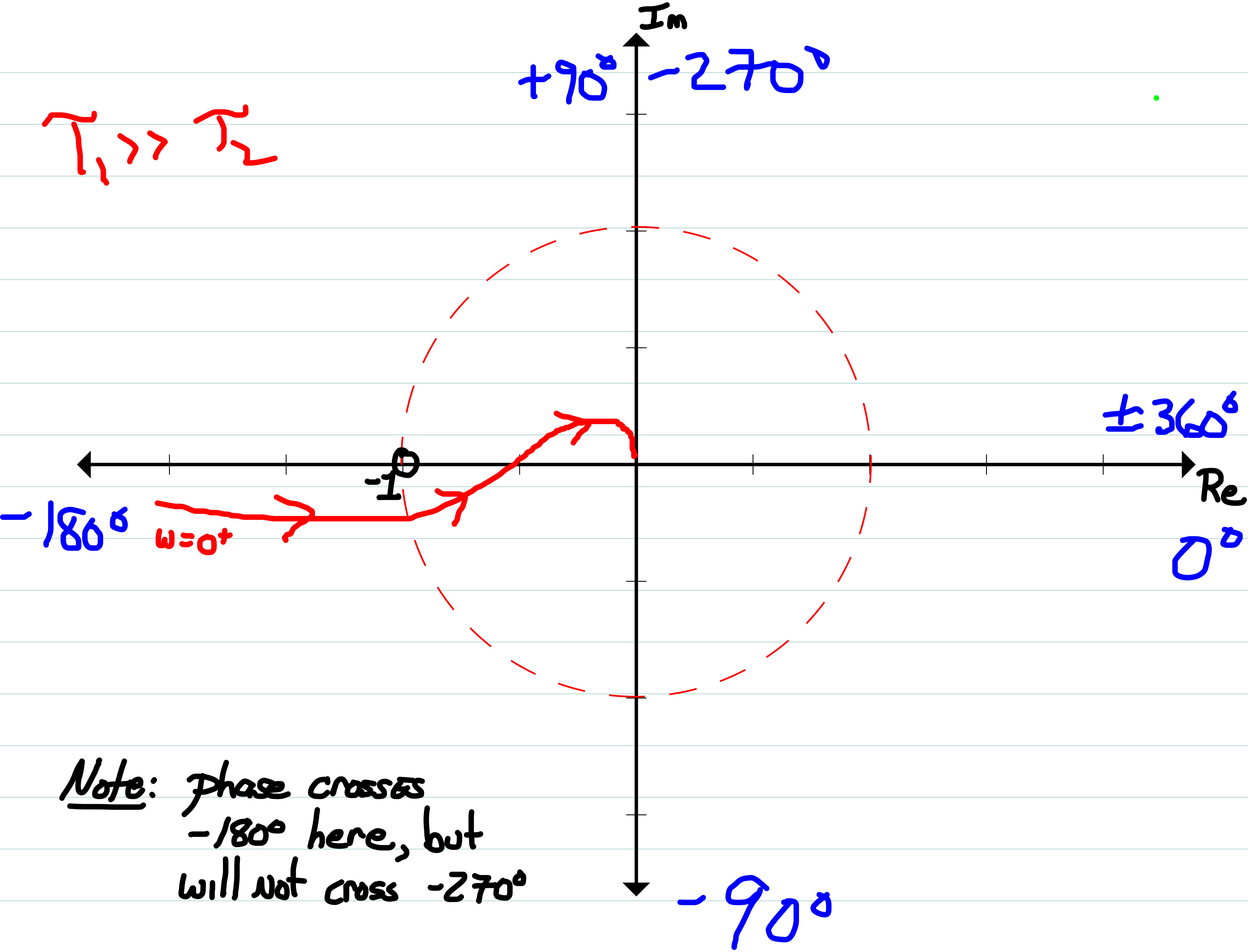
Example:

$$G(s) = K_B \left[\frac{T_1 s + 1}{s^2 (T_2 s + 1)^2} \right]$$



Im
+90°
−270°
$T_1 << T_2$
also possible
±360°
−1
−180°
Re
0°
Note: phase never crosses
−180° here
(may or may not cross
−270°)
−90°

$T_1 >> T_2$



<u>Note</u>: Phase crosses −180° here, but will not cross −270°

A more complicated example

$$G(s) = K_B \left[\frac{(T_2 s+1)^2}{s^2 (T_1 s+1)(T_3 s+1)^2} \right]$$

With $T_1 \gg T_2 \gg T_3 > \phi$ $\left(\frac{1}{T_1} \ll \frac{1}{T_2} \ll \frac{1}{T_3} \right)$

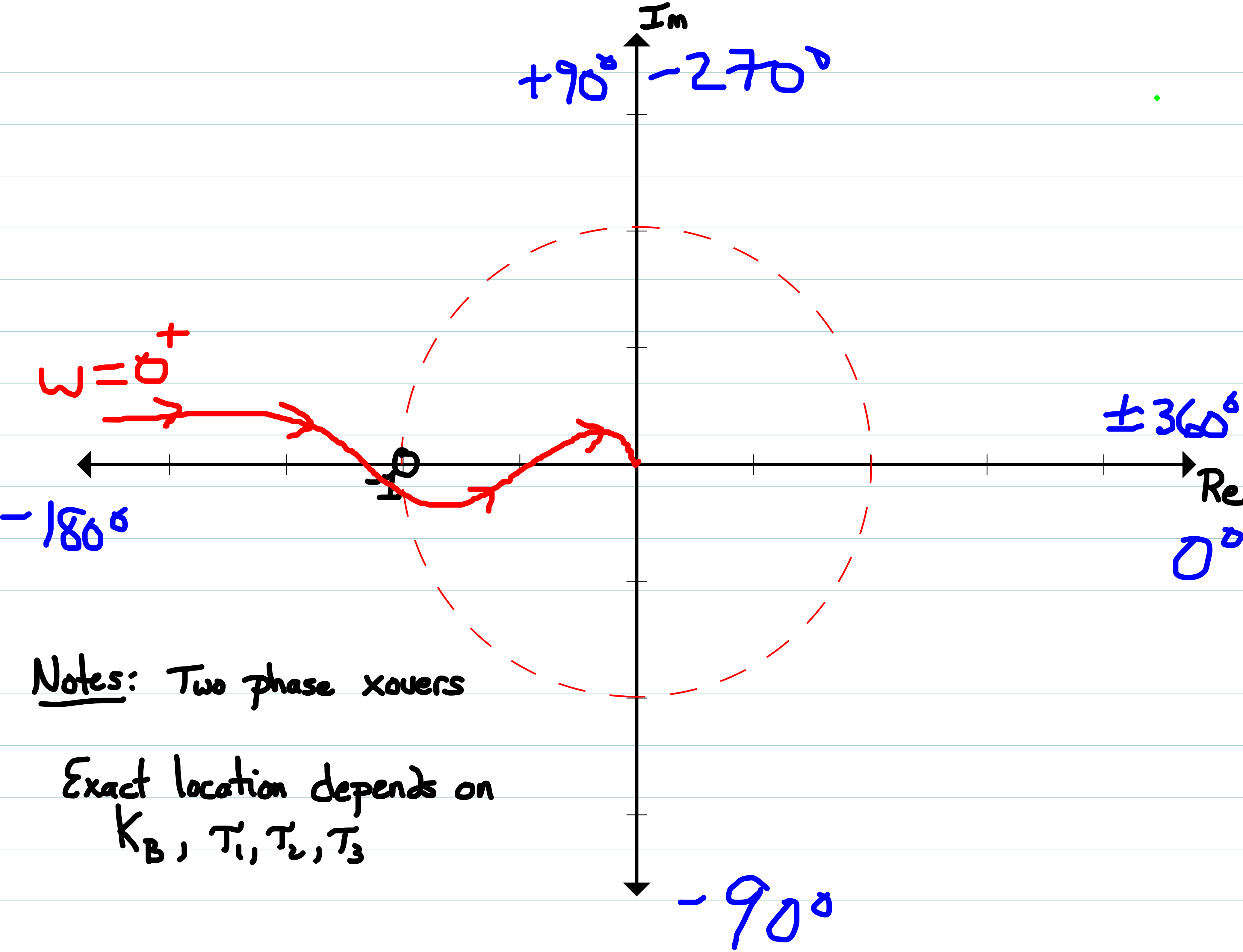
Low freq phase: $-180°$

high freq phase: $-270°$

phase initially decreases from pole at $-1/T_1$

Then increases due to double zero at $-1/T_2$

Then falls again due to double pole at $-1/T_3$



Notes: Two phase xovers

Exact location depends on K_B, T_1, T_2, T_3

Log magnitude - Angle Diagram (Nichols plot)

⇒ Plot $|G(j\omega)|_{dB}$ vs. $\measuredangle G(j\omega)$ as ω varies from $\emptyset$ to ∞

⇒ Angle in deg is horizontal axis

⇒ Magnitude in dB is vertical axis

⇒ Plot usually centered so "origin" corresponds to $-180°$ in phase, $\emptyset$ dB in magnitude

Log magnitude - Angle Diagram (Nichols plot)

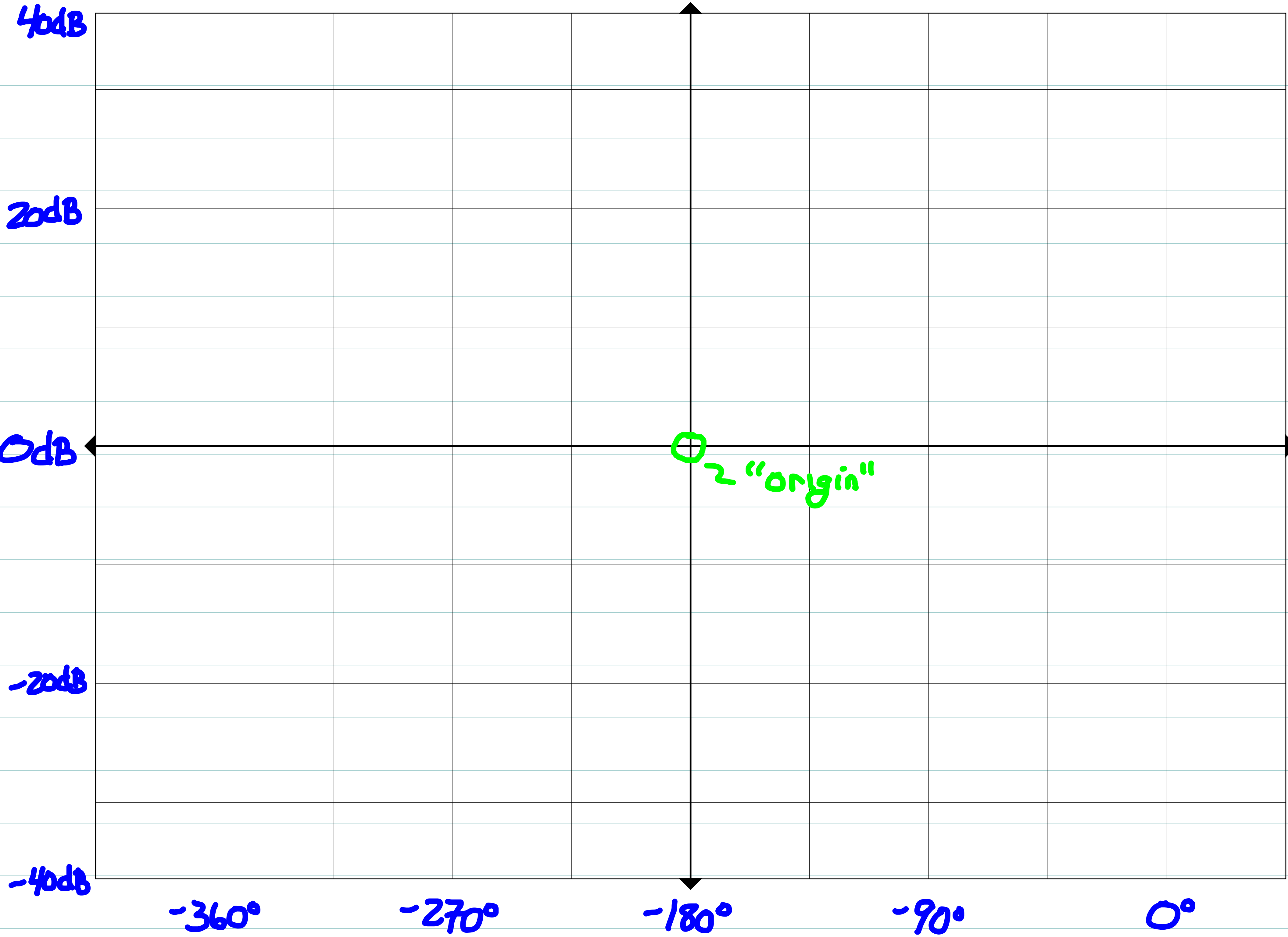
⇒ Plot $|G(j\omega)|_{dB}$ vs. $\measuredangle G(j\omega)$ as ω varies from ϕ to ∞

⇒ Angle in deg is horizontal Axis

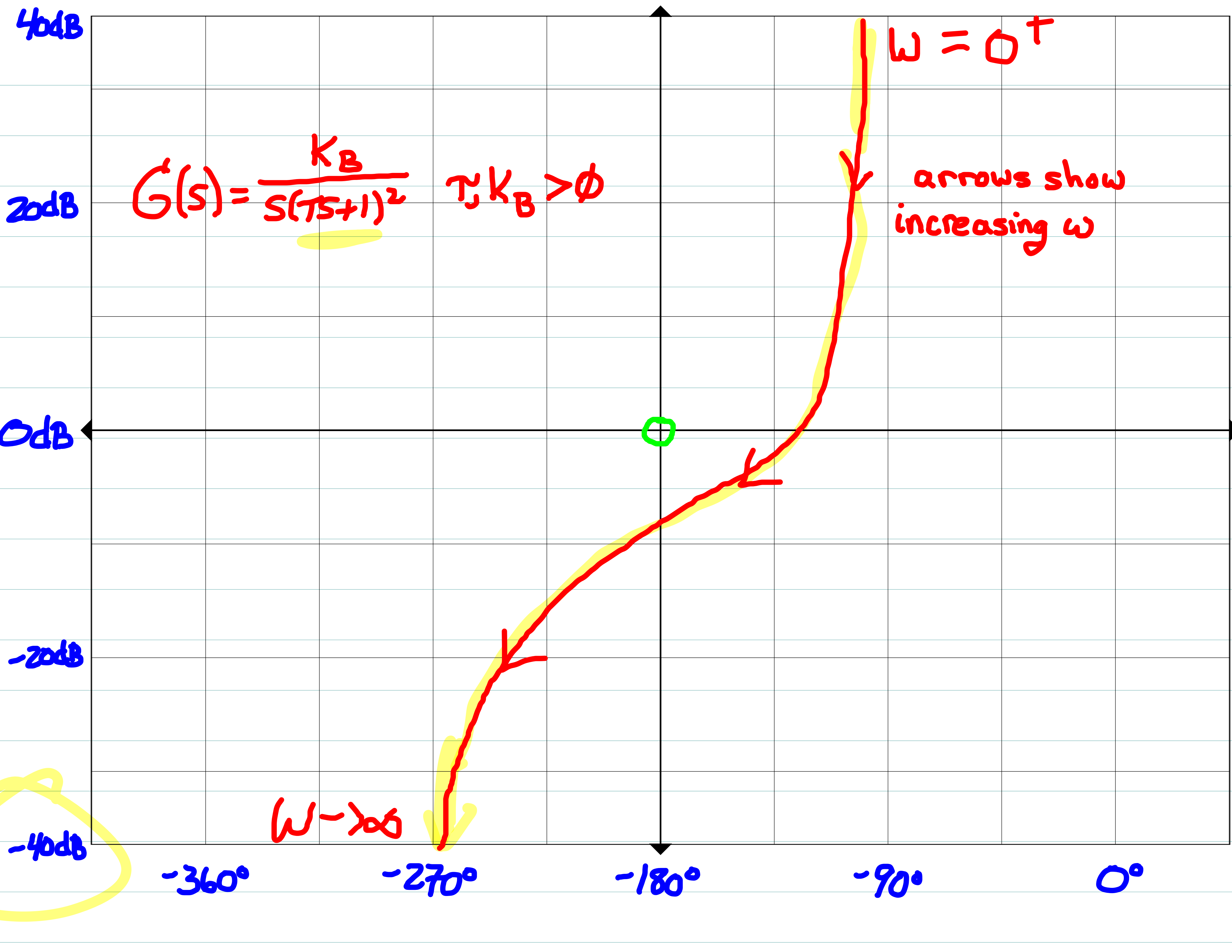
⇒ Magnitude in dB is vertical Axis

⇒ Plot usually centered so "origin" corresponds to $-180°$ in phase, ϕ dB in magnitude

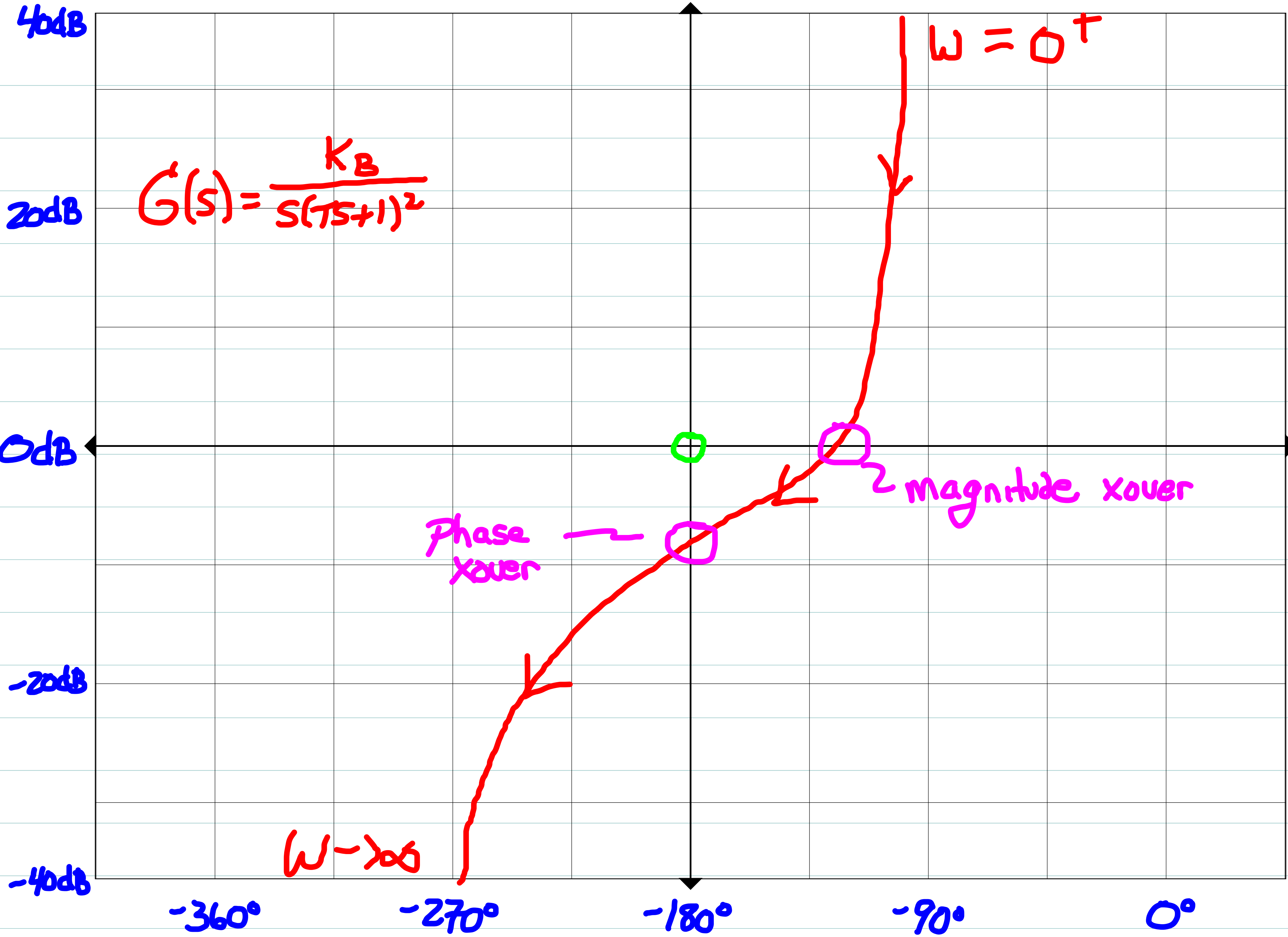
⇒ "Origin" of plot corresponds to -1 point on polar diagram



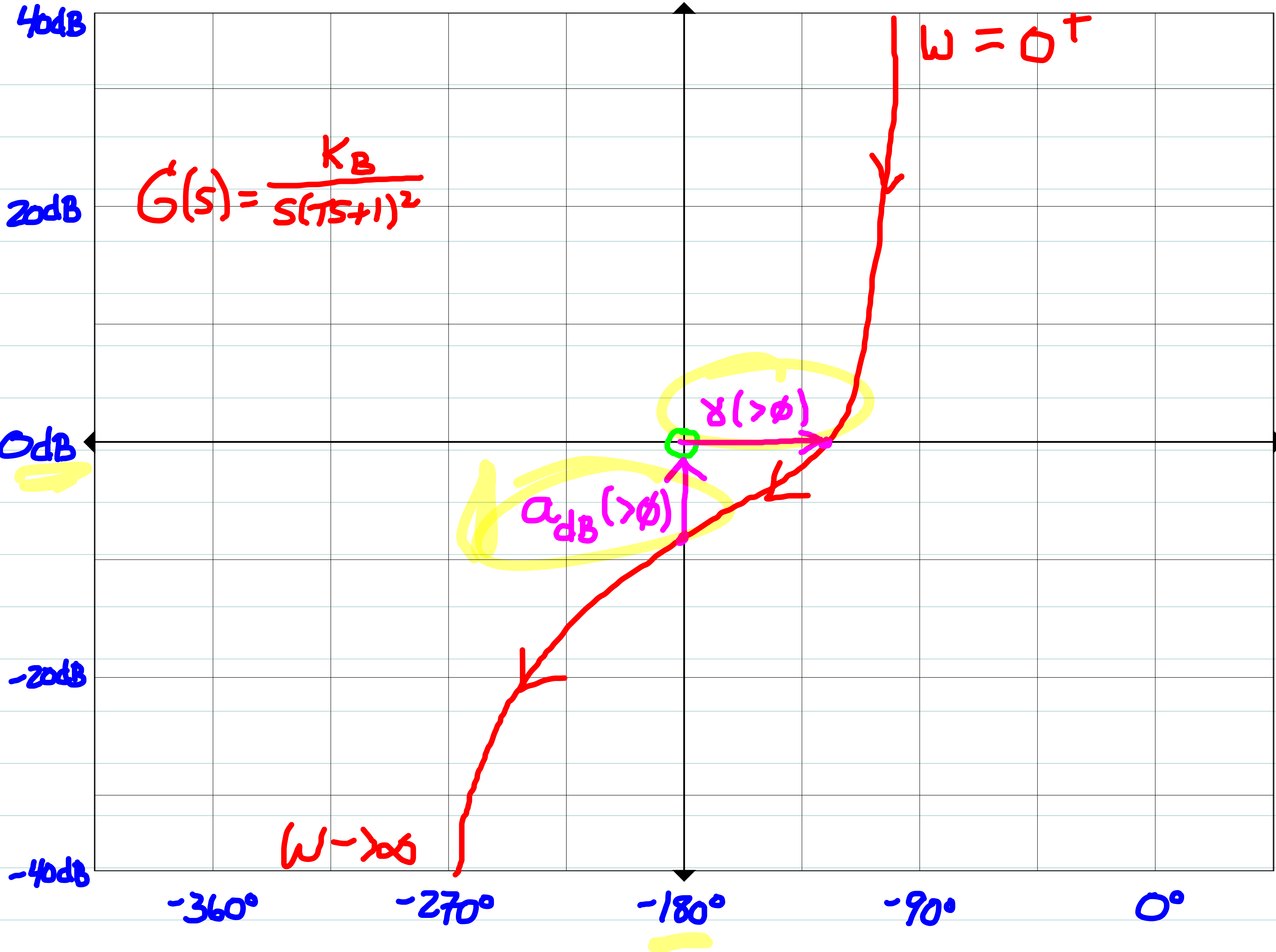
40dB
20dB
0dB
-20dB
-40dB
-360°
-270°
-180°
-90°
0°
"origin"



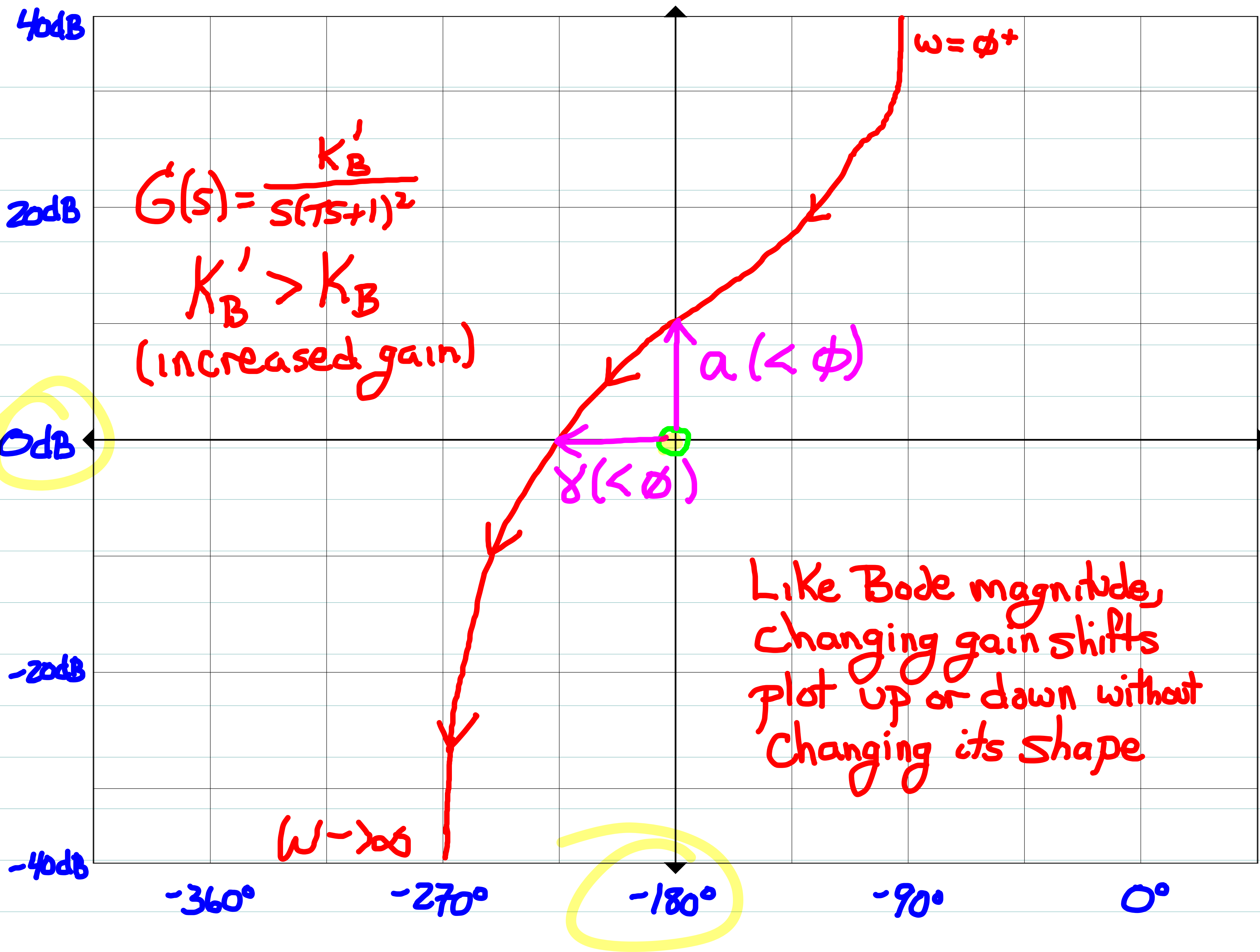
40dB
20dB
0dB
-20dB
-40dB
-360°
-270°
-180°
-90°
0°
$G(s) = \frac{K_B}{s(\tau s+1)^2}$ $\tau, K_B > \phi$
$\omega = 0^+$
arrows show increasing ω
$\omega \to \infty$



40dB
20dB
0dB
-20dB
-40dB
-360°
-270°
-180°
-90°
0°
$G(s) = \frac{K_B}{s(Ts+1)^2}$
$\omega = 0^+$
$\omega \to \infty$
magnitude xover
Phase xover



40dB
20dB
0dB
-20dB
-40dB
-360°
-270°
-180°
-90°
0°
$G(s) = \frac{K_B}{s(Ts+1)^2}$
$\omega = 0^+$
$\omega \to \infty$
$\gamma (>\phi)$
$a_{dB} (>\phi)$


40dB
20dB
0dB
-20dB
-40dB
-360°
-270°
-180°
-90°
0°
$\omega = \phi^+$
$G(s) = \dfrac{K_B'}{s(Ts+1)^2}$
$K_B' > K_B$
(increased gain)
$a(<\phi)$
$\gamma(<\phi)$
$\omega \to \infty$
Like Bode magnitude, changing gain shifts plot up or down without changing its shape

⇒ Primary use is to easily see margins, measured along orthogonal AXES relative to "origin"

⇒ Phase margin measured along horizontal Axis to magnitude crossover point

⇒ <u>positive</u> if crossing is to <u>right</u> of "origin" negative otherwise

⇒ Gain margin (in dB) measured along vertical Axis to phase crossover point

⇒ <u>Positive</u> if crossing is <u>below</u> "origin" negative otherwise.

⇒ Primary use is to easily see margins, measured along orthogonal AXES relative to "origin"

⇒ Phase margin measured along horizontal Axis to magnitude crossover point

⇒ <u>positive</u> if crossing is to <u>right</u> of "origin" negative otherwise

⇒ Gain margin (in dB) measured along vertical Axis to phase crossover point

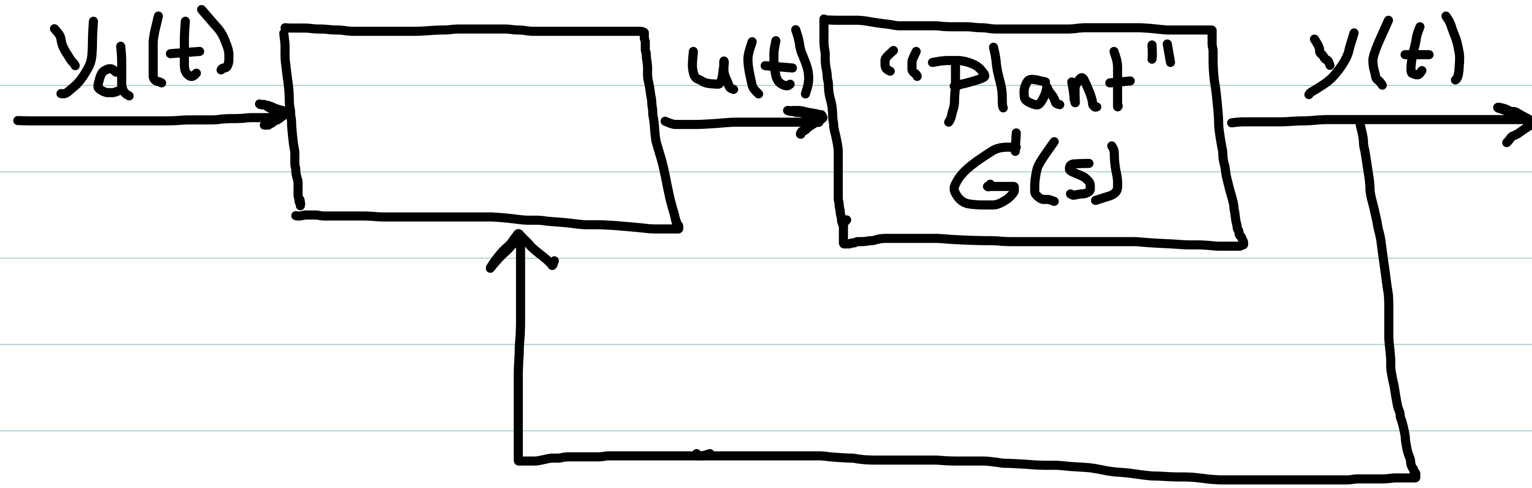
⇒ <u>Positive</u> if crossing is <u>below</u> "origin" negative otherwise.

⇒ <u>Why</u> is proximity of polar/Nichols to −1 so important??

Feedback Control (finally!)

⇒ Automatically generate inputs $u(t)$ so that output $y(t)$ tracks "desired output" $y_d(t)$ as closely as possible

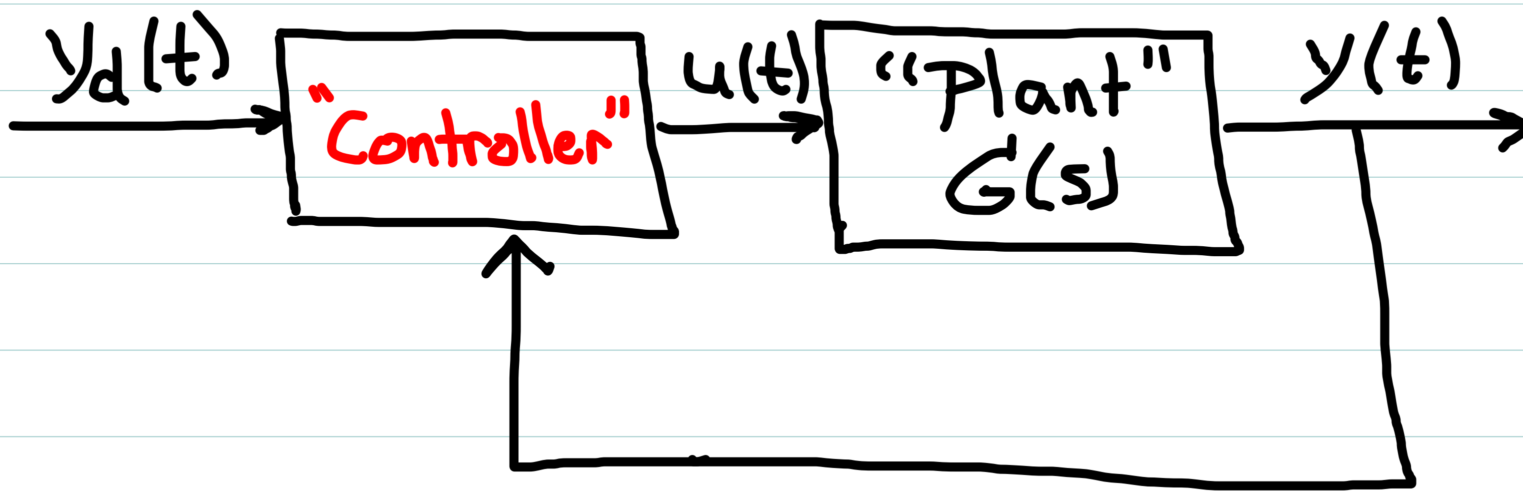
⇒ Input determined in real-time by continually comparing $y(t)$ with $y_d(t)$



Feedback Control (finally!)

⇒ Automatically generate inputs $u(t)$ so that output $y(t)$ tracks "desired output" $y_d(t)$ as closely as possible

⇒ Input determined in real-time by continually comparing $y(t)$ with $y_d(t)$



Feedback Controllers

=> The controller is a device that we design to compute $u(t)$ from $y_d(t)$ and $y(t)$, to satisfy specified constraints.

=> The relationship between $y_d(t)$, $y(t)$ and $u(t)$ is known as the "control law". This is a mathematical algorithm for computing $u(t)$.

=> for example:

$$u(t) = K\left[y_d(t) - y(t)\right]$$

In this control law, $u(t)$ is proportional to the difference between $y_d(t)$ and $y(t)$.

=> Controllers are implemented as programs (usually in C/C++) on a digital computer onboard the vehicle.

Control Laws

⇒ Control laws can be any mathematical function of $y(t)$ and $y_d(t)$, including differential equations

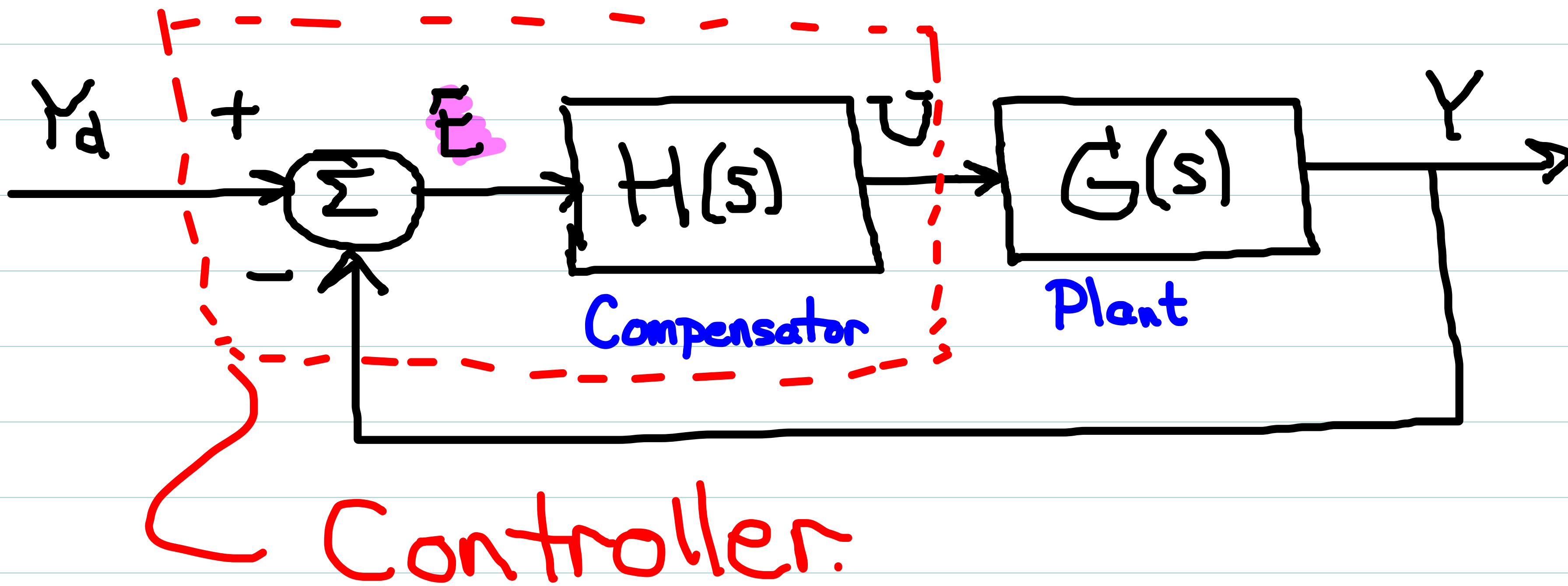
⇒ For example: $H(s) = \dfrac{\beta_1 s + \beta_0}{s + \alpha_0}$

$$\dot{u}(t) + \alpha_0 u(t) = \beta_1 \frac{d}{dt}\Big[\overbrace{y_d(t) - y(t)}^{e(t)}\Big] + \beta_0\Big[\overbrace{y_d(t) - y(t)}^{e(t)}\Big]$$

⇒ In such cases, we can model the operation of the controller in the same transfer function framework used to model the physical system being controlled.

⇒ The "standard servo loop" is a systematic framework for analyzing these control strategies.

Standard Servo Loop



Action of controller is:

$$U(s) = H(s)E(s)$$

"Tracking error"

where $E(s) = Y_d(s) - Y(s) \Rightarrow \boxed{e(t) = y_d(t) - y(t)}$

however finding $u(t)$ from $e(t)$ requires solving differential equation corresponding to $H(s)$.

Controller Design

$$U(s) = H(s)E(s)$$

$H(s)$ is a new transfer function that we design

It has no physical basis, we create it to solve the control problem for a particular physical system $G(s)$.

There is no unique specification of $H(s)$ for a specific $G(s)$. Many different design tradeoffs which do not have a unique soln.

Guiding principle: Use the simplest $H(s)$ (fewest poles + zeros) which will provide desired performance.

Servo Loop Analysis

$$U(s) = H(s)\left[Y_d(s) - Y(s)\right]$$

$$Y(s) = G(s)\,U(s)$$

Circular! Y depends on U, but U depends on Y.

Very tricky to "untangle" the circularity using the governing diff'l eq'ns for G, H.

Laplace makes it easy!

$$Y = GU = GHE = GH(Y_d - Y)$$

$$\Rightarrow (1+GH)Y = GHY_d$$

or

$$Y(s) = \underbrace{\boxed{\left[\frac{G(s)H(s)}{1+G(s)H(s)}\right]}}_{T(s)}\,Y_d(s)$$

T(s) "closed-loop" TF

Loop Transfer Functions

Define

$$L(s) = G(s)H(s)$$ "open-loop" TF

then

$$\boxed{T(s) = \frac{L(s)}{1+L(s)}}$$ "closed-loop" TF

and

$$Y(s) = L(s)E(s)$$ open-loop dynamics

$$Y(s) = T(s)Y_d(s)$$ closed-loop dynamics

$T(s)$ gives us direct information about system performance

$$Y_d \rightarrow \boxed{T(s)} \xrightarrow{Y}$$

$L(s)$ is an important intermediate quantity in analysis + design

Another Useful relationship

$$E(s) = Y_d(s) - Y(s) = Y_d(s) - T(s)\,Y_d(s)$$

$$= \underbrace{[1 - T(s)]}_{S(s)}\,Y_d(s)$$

$S(s)$: "Sensitivity" TF

Note that:

$$S(s) = 1 - T(s) = 1 - \frac{L(s)}{1+L(s)}$$

So:
$$S(s) = \frac{1}{1+L(s)}$$

Thus:

$$S(s) = 1 - T(s) = \frac{1}{1+L(s)}$$

are equivalent, although we will primarily work with the second form.

Final Important Relationship

$U(s) = H(s)E(s)$ [TF model of control law]

$$= \underbrace{\left[H(s)S(s)\right]}_{R(s)} Y_d(s)$$

$$R(s) = H(s)S(s) = \frac{H(s)}{1+L(s)}$$

Used to predict control signals which will be generated under ideal conditions

⇒ $y_d(t)$ known perfectly for all $t \geq 0$

⇒ perfect model of system (no errors in model, no disturbances)

$R(s)$ used only theoretically. $H(s)$ is used for actual implementation.

Example:

Suppose $G(s) = \dfrac{2(s+1)}{s+3}$ $\quad H(s) = \dfrac{K}{s}$

Then $L = GH = \dfrac{2K(s+1)}{s(s+3)}$

$$T = \frac{L}{1+L} = \frac{2K(s+1)}{s(s+3)+2K(s+1)} = \frac{2K(s+1)}{s^2+(3+2K)s+2K}$$

$$S = \frac{1}{1+L} = \frac{s(s+3)}{s^2+(3+2K)s+2K}$$

$$R = \frac{H}{1+L} = \frac{K(s+3)}{s^2+(3+2K)s+2K}$$

Three ~~Derived~~ TFs for Feedback Loops

Given $G(s)$ and $H(s)$, we can derive $\mathcal{R}(s)$, $\mathcal{S}(s)$, $\mathcal{T}(s)$ so that:

$$Y_d(s) \longrightarrow \boxed{\mathcal{T}(s)} \longrightarrow Y(s) \qquad\qquad \mathcal{T}(s) = \frac{L(s)}{1+L(s)}$$

$$Y_d(s) \longrightarrow \boxed{\mathcal{S}(s)} \longrightarrow E(s) \qquad\qquad \mathcal{S}(s) = \frac{1}{1+L(s)}$$

$$Y_d(s) \longrightarrow \boxed{\mathcal{R}(s)} \longrightarrow U(s) \qquad\qquad \mathcal{R}(s) = \frac{H(s)}{1+L(s)}$$

$\Rightarrow$ Each of these derived TFs can be analyzed using the same tools developed for $G(s)$.

Uses of derived TF:

$\Rightarrow$ $T(s)$ tells us about actual response of controlled system for specific $y_d(t)$

$$\boxed{Y(s) = T(s)Y_d(s)}$$

$\Rightarrow$ $S(s)$ tells us about tracking accuracy for specific $y_d(t)$

$$\boxed{E(s) = S(s)Y_d(s)}$$

$\Rightarrow$ $R(s)$ tells us about required input for specific $y_d(t)$:

$$\boxed{U(s) = R(s)Y_d(s)}$$

Note: all 3 of these TF have the same denominator, hence same poles!!!

Example use of loop TF:

Suppose $y_d(t) = A\,\mathbb{1}(t)$ (step of magnitude A)

Then:

$$y(t) = A \times \{\text{step response of } T(s)\}$$

$$u(t) = A \times \{\text{step response of } R(s)\}$$

$$e(t) = A \times \{\text{step response of } S(s)\}$$

Note in particular here that:

$$e_{ss}(t) =$$

Example use of loop TF:

Suppose $y_d(t) = A\mathbb{1}(t)$ (step of magnitude A)

Then:

$$y(t) = A \times \{\text{step response of } T(s)\}$$

$$u(t) = A \times \{\text{step response of } R(s)\}$$

$$e(t) = A \times \{\text{step response of } S(s)\}$$

Note in particular here that:

$$e_{ss}(t) = A\,S(0) \qquad \text{(constant)}$$

Thus generally we'd like to make sure $S(0) = 0$ (or at least very small).

Example Application: Tracking Ability

A good feedback loop needs to ensure $|e_{ss}(t)|$ small for a wide variety of $y_d(t)$.

Suppose $y_d(t) = A$ (constant)

Then (assuming all poles of $\mathcal{S}(s)$ at least stable)

$$e_{ss}(t) = A\,\mathcal{S}(0)$$

So good tracking requires $|\mathcal{S}(0)|$ small.

Ideally, $\mathcal{S}(0) = 0 \Rightarrow e_{ss}(t) = 0$ "perfect tracking"

and this is often a basic design <u>requirement</u>.

Tracking (cont)

Suppose more generally $y_d(t) = A\cos\omega t$

then $e_{ss}(t) = A|S(j\omega)|\cos(\omega t + \measuredangle S(j\omega))$

and in particular $|e_{ss}(t)| \le A|S(j\omega)|$

So we want $|S(j\omega)| \ll 1$ for a wide range of frequencies ω (including $\omega = \phi$)

$\Rightarrow$ Want Bode magnitude diagram $|S(j\omega)| \ll \phi\,dB$ for a large range of ω (including ϕ).

$\Rightarrow$ We will show feedback loops with good tracking properties place constraints on design process, which often conflict with other requirements (stability + performance).

Bandwidth

Define ω_B to be largest ω for which

$$|S(j\omega)| \leq -3\,dB \quad \text{for all } \omega \in [0, \omega_B]$$

this is the (tracking) <u>bandwidth</u> of the system.

$\Rightarrow$ We want designs with <u>high bandwidth</u>.

<u>Note</u>: $-3\,dB$ is an arbitrary boundary between acceptable and poor tracking. Realistic performance constraints are typically much tighter:

$$|S(j\omega)| \leq -20\,dB \quad (\leq 10\% \text{ worst case error})$$

or

$$|S(j\omega)| \leq -40\,dB \quad (\leq 1\% \text{ worst case error})$$

Example Application: Utility of $R(s)$

⇒ $R(s)$ lets us theoretically predict the $u(t)$ which will be generated under ideal circumstances given a specified $y_d(t)$.

$$u(t) = \mathcal{L}^{-1}\{R(s)Y_d(s)\}$$

⇒ Primary quantity of interest is $\max_{t \geq 0} |u(t)|$

⇒ Quantifies maximum control effort required.

⇒ Real actuators have limits $|u(t)| \leq u_{max}$

⇒ Must ensure our control strategy does not "saturate" the actuators, i.e. $\max_t |u(t)| \leq u_{max}$

Saturation

Saturation of actuators, i.e. $|u(t)| > u_{max}$ for some $t \geq 0$, may produce performance degradation or even instability even when the poles of $R(s)$ are "good."

Unfortunately, no simple design guidelines for $H(s)$ which ensure saturation does not occur.

Some degree of design iteration typically required

Advanced (graduate level) techniques do exist to incorporate actuator limits into the design process.

Closed-loop poles

Transient

=> Performance of controlled system (settling time, steady-state, overshoot, etc) depends on poles of $T(s)$

=> ($R(s)$ and $S(s)$ have same poles!!)

=> Where are these poles??

=> Determined by denominator of $T(s)$

=> ($R(s)$ and $S(s)$ have same denominator)

=> Denom of all 3 derived TF is:

$$1+L(s)$$

Charactistic Equation

Poles of $T(s)$, $R(s)$, $S(s)$ are at values of $s \in \mathbb{C}$ such that

(CE) $$\boxed{1 + L(s) = 0}$$ "Characteristic equation" of feedback system

We need sol'ns of this equation to be in "good" locations of complex plane.

Will identify required properties for $L(s)$ so this is true, then work backwards to determine required properties of $H(s)$.

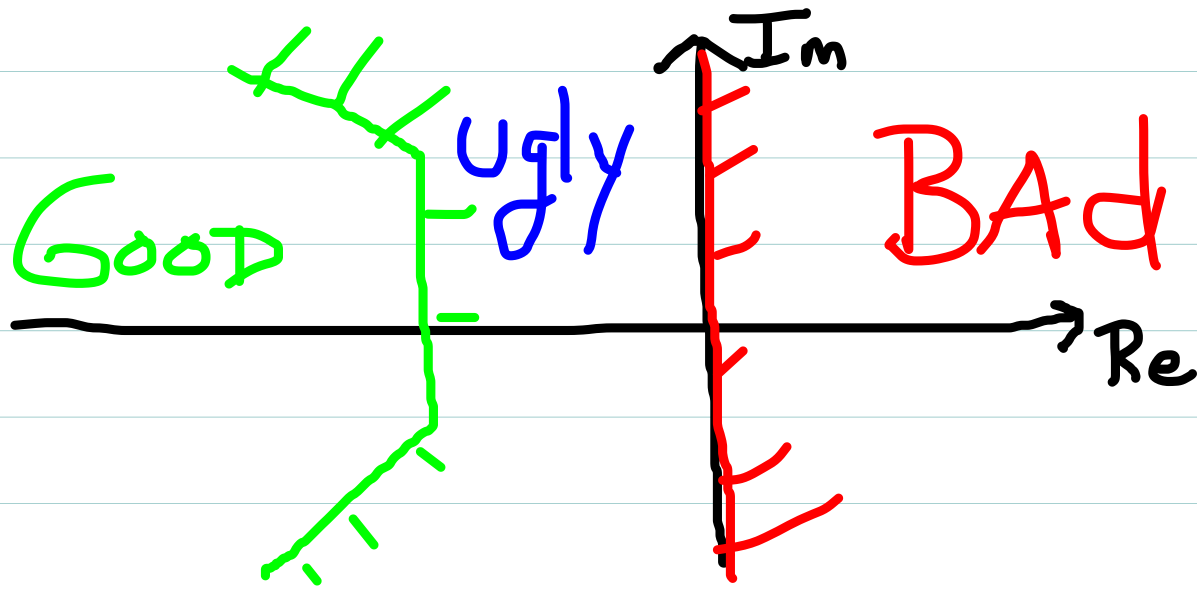
(recall: $L(s) = G(s)H(s)$).

Fundamental Consideration: Closed-loop Stability

$\boxed{\text{Most basic design consideration:}}$!!

Closed-loop poles should be "good", and certainly <u>must be stable</u>.

Thus, sol'ns of $\boxed{\text{CE: } \quad 1+L(s)=\emptyset}$ must be in left half of complex plane, preferably in "good region" (far from imag Axis, relatively close to or on the real Axis).



Note: $1+L(s)=0$ is a polynomial equation

Suppose

$$G(s)=\frac{2}{s^2} \qquad H(s)=\frac{K(s-z)}{(s-p)}$$ K, z, p design choices.

Then

$$L(s)=G(s)H(s)=\frac{2K(s-z)}{s^2(s-p)}$$

and

$$\boxed{1+L(s)=0=1+\frac{2K(s-z)}{s^2(s-p)}}$$

SAME

Equivalently:

$$s^2(s-p)+2K(s-z)=0$$

or

$$\boxed{s^3-ps^2+2Ks-2Kz=0}$$

Three <u>Derived</u> TFs for Feedback Loops

Given $G(s)$ and $H(s)$, we can derive $R(s)$, $S(s)$, $T(s)$ so that:

$$Y_d(s) \longrightarrow \boxed{T(s)} \xrightarrow{Y(s)} \qquad T(s) = \frac{L(s)}{1+L(s)}$$

$$\xrightarrow{Y_d(s)} \boxed{S(s)} \xrightarrow{E(s)} \qquad S(s) = \frac{1}{1+L(s)}$$

$$\xrightarrow{Y_d(s)} \boxed{R(s)} \xrightarrow{U(s)} \qquad R(s) = \frac{H(s)}{1+L(s)}$$

$\Rightarrow$ Each of these <u>derived</u> TFs can be analyzed using the same tools developed for $G(s)$.

Uses of derived TF:

⇒ $T(s)$ tells us about actual response of controlled system for specific $y_d(t)$

$$\boxed{Y(s) = T(s)\,Y_d(s)}$$

⇒ $S(s)$ tells us about tracking accuracy for specific $y_d(t)$

$$\boxed{E(s) = S(s)\,Y_d(s)}$$

⇒ $R(s)$ tells us about required input for specific $y_d(t)$:

$$\boxed{U(s) = R(s)\,Y_d(s)}$$

Note: all 3 of these TF have the same denominator, hence same poles!!!

Example use of loop TF:

Suppose $y_d(t) = A\mathbb{1}(t)$ (step of magnitude A)

Then:

$$y(t) = A \times \{\text{step response of } T(s)$$

$$u(t) = A \times \{\text{step response of } R(s)\}$$

$$e(t) = A \times \{\text{step response of } S(s)\}$$

Note in particular here that:

$$e_{ss}(t) =$$

Example use of loop TF:

Suppose $y_d(t) = A\mathbb{1}(t)$ (step of magnitude A)

Then:

$$y(t) = A \times \{\text{step response of } T(s)\}$$

$$u(t) = A \times \{\text{step response of } R(s)\}$$

$$e(t) = A \times \{\text{step response of } \mathcal{S}(s)\}$$

Note in particular here that:

$$e_{ss}(t) = A\,\mathcal{S}(0) \qquad \text{(constant)}$$

Thus generally we'd like to make sure $\mathcal{S}(0) = 0$ (or at least very small).

Example Application: Tracking Ability

A good feedback loop needs to ensure $|e_{ss}(t)|$ small for a wide variety of $y_d(t)$.

Suppose $y_d(t) = A$ (constant)

Then (assuming all poles of $\mathcal{S}(s)$ at least stable)

$$e_{ss}(t) = A\,\mathcal{S}(0)$$

So good tracking requires $|\mathcal{S}(0)|$ small.

Ideally, $\mathcal{S}(0) = 0 \Rightarrow e_{ss}(t) = 0$ "perfect tracking"

and this is often a basic design <u>requirement</u>.

Tracking (cont)

Suppose more generally $y_d(t) = A\cos\omega t$

then $e_{ss}(t) = A|S(j\omega)|\cos(\omega t + \measuredangle S(j\omega))$

and in particular $|e_{ss}(t)| \leq A|S(j\omega)|$

So we want $|S(j\omega)| << 1$ for a wide range of frequencies ω (including $\omega = \emptyset$)

$\Rightarrow$ Want Bode magnitude diagram $|S(j\omega)| << \emptyset$ dB for a large range of ω (including $\emptyset$).

$\Rightarrow$ We will show feedback loops with good tracking properties place constraints on design process, which often conflict with other requirements (stability + performance).

Bandwidth

Define ω_B to be largest ω for which

$$|S(j\omega)| \leq -3\,dB \quad \text{for all } \omega \in [0, \omega_B]$$

this is the (tracking) <u>bandwidth</u> of the system.

$\Rightarrow$ We want designs with <u>high bandwidth</u>.

<u>Note</u>: $-3\,dB$ is an arbitrary boundary between acceptable and poor tracking. Realistic performance constraints are typically much tighter:

$$|S(j\omega)| \leq -20\,dB \quad (\leq 10\% \text{ worst case error})$$

or

$$|S(j\omega)| \leq -40\,dB \quad (\leq 1\% \text{ worst case error})$$

Example Application: Utility of $R(s)$

⇒ $R(s)$ lets us theoretically predict the $u(t)$ which will be generated under ideal circumstances given a specified $y_d(t)$.

$$u(t) = \mathcal{L}^{-1}\{R(s)Y_d(s)\}$$

⇒ Primary quantity of interest is $\max_{t \geq 0} |u(t)|$

⇒ Quantifies maximum control effort required.

⇒ Real actuators have limits $|u(t)| \leq u_{max}$

⇒ Must ensure our control strategy does not "saturate" the actuators, i.e. $\max_{t} |u(t)| \leq u_{max}$

Saturation

Saturation of actuators, i.e. $|u(t)| > u_{max}$ for some $t \geq 0$, may produce performance degradation or even instability even when the poles of $R(s)$ are "good."

Unfortunately, no simple design guidelines for $H(s)$ which ensure saturation does not occur.

Some degree of design iteration typically required

Advanced (graduate level) techniques do exist to incorporate actuator limits into the design process.

Closed-loop poles

=> Performance of controlled system (settling time, steady-state, overshoot, etc) depends on poles of $T(s)$

=> ($R(s)$ and $S(s)$ have same poles!!)

=> Where are these poles??

=> Determined by denominator of $T(s)$

=> ($R(s)$ and $S(s)$ have same denominator)

=> Denom of all 3 derived TF is:

$$1+L(s)$$

Charactistic Equation

Poles of $T(s), R(s), S(s)$ are at values of $s \in \mathbb{C}$ such that

(CE) $\boxed{1 + L(s) = \emptyset}$ "Chaderistic equation" of feedback system

We need sol'ns of this equation to be in "good" locations of complex plane.

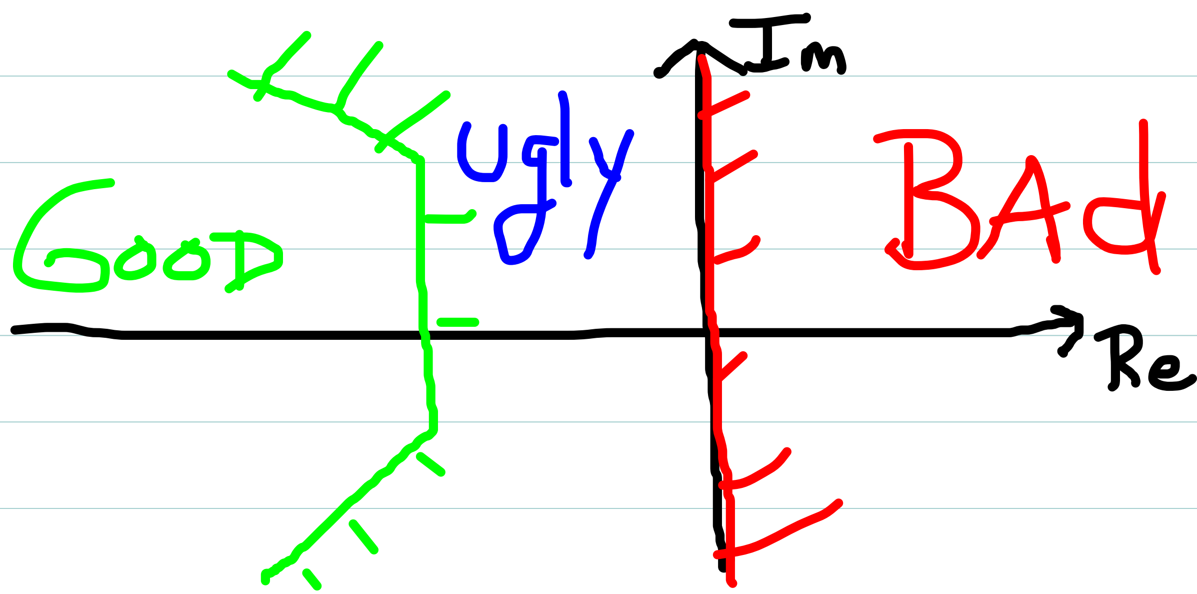
Will identify required properties for $L(s)$ so this is true, then work backwards to determine required properties of $H(s)$.

Fundamental Consideration: Closed-loop Stability

Most basic design consideration: !!

Closed-loop poles should be "good", and certainly <u>must be stable</u>.

Thus, sol'ns of $\boxed{\text{CE: } 1+L(s)=0}$ must be in left half of complex plane, preferably in "good region" (far from imag Axis, relatively close to or on the real Axis).



A Crucial Observation:

If $L(j\omega) = -1$ for some ω, then

$1 + L(s) = \emptyset$ has a sol'n $s = j\omega$ for some ω

$\Rightarrow$ closed-loop dynamics has poles at $\pm j\omega$, on imag Axis

$\Rightarrow$ Such poles are on the boundary between bad and ugly

$\Rightarrow$ This situation must be avoided!!!

Now if $\boxed{L(j\omega) = -1}$ for some $\omega > 0$, then:

$\Rightarrow$ Polar plot of $L(j\omega)$ passes through -1

$\Rightarrow$ $\omega_a = \omega_\gamma$ (both crossover freqs same)

$\Rightarrow$ $\boxed{a = 0 \text{ dB}, \ \gamma = 0^\circ}$ (both margins 0)

Any such feedback loop is bad! !!

Now, suppose $\exists \omega \geq 0 \ni: L(j\omega) \approx -1$ (i.e. close to, but not exactly -1)

By continuity of $L(s)$, $1 + L(s) = 0$ would have a sol'n very near (but not exactly on) the imag axis.

Some poles of $T(s)$ would be in bad or ugly region
$\Rightarrow$ Also undesireable!

Now, if $L(j\omega) \approx -1$ for some $\omega \geq \emptyset$

$\Rightarrow$ polar plot of $L(j\omega)$ comes very close to -1 but doesn't pass exactly through it

$\Rightarrow$ (typically) $|a_{dB}|$ and $|\gamma|$ very small (small margins)

$\Rightarrow$ $\boxed{\text{This should also be avoided.}}$

Thus, for $T(s)$ to have only good poles, we need conditions:

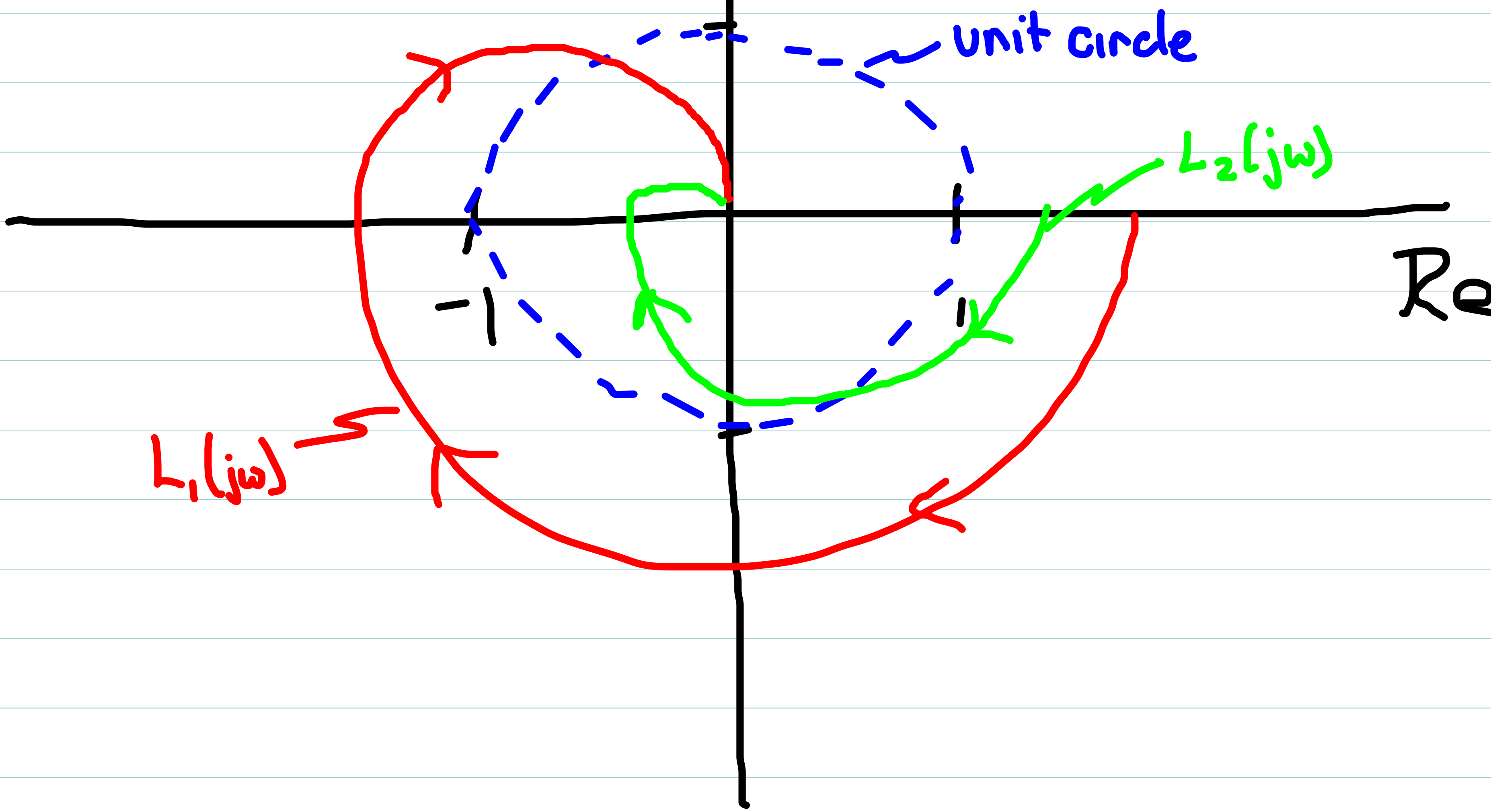
$\Rightarrow$ Gain and phase margins $\boxed{\text{of } L(s)}$ $\Leftarrow$!!! to be large

$\Rightarrow$ polar plot of $L(j\omega)$ avoids -1 by wide margins

Necessary, but <u>not sufficient</u>!

Both plots avoid -1 by large margins

Is one better?
Yes! But criterion is non-obvious!



Nyquist Stability Criterion

All roots of $1+L(s)=\emptyset$ are in LHP if

the Nyquist diagram (a modified polar plot) $\boxed{\text{of } L(j\omega)}$

circles the -1 point the correct number of times.

$\Rightarrow$ Major Theoretical result! Used extensively in control theory

$\Rightarrow$ Questions to answer

$\Rightarrow$ How to create diagram from polar?

$\Rightarrow$ How to count encirclements of -1?

$\Rightarrow$ How many encirclements needed?

Nyquist Diagram

When $L(s)$ is type $N \leq \emptyset$ (no poles at origin)

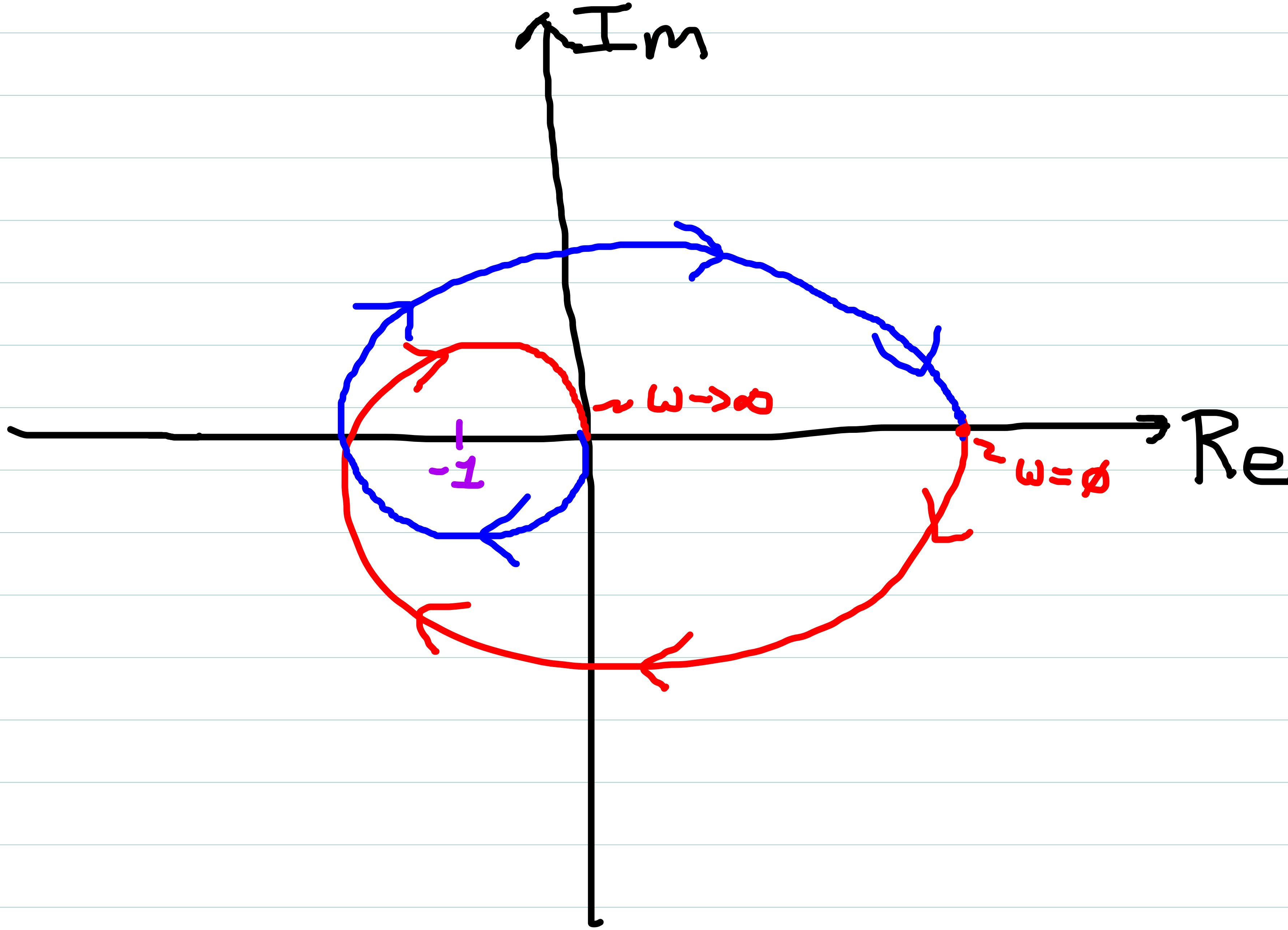
$\Rightarrow$ Draw polar of $L(j\omega)$

$\Rightarrow$ "Flip" polar of L about real axis
(this is the polar of $L(-j\omega)$, i.e. for negative frequencies)

$\Rightarrow$ Put arrows on flipped plot whose direction is consistent with direction of arrows on original polar plot
(i.e. arrows show direction of increasing frequency, from $\omega = -\infty$, through $\omega = \emptyset$, to $\omega = \infty$).

(We will modify for $N > \emptyset$ after we examine complete stability condition.

Example: $L(s) = \dfrac{K_B}{(\tau s+1)^3}$ $K_B, \tau > 0$



Counting Encirclements

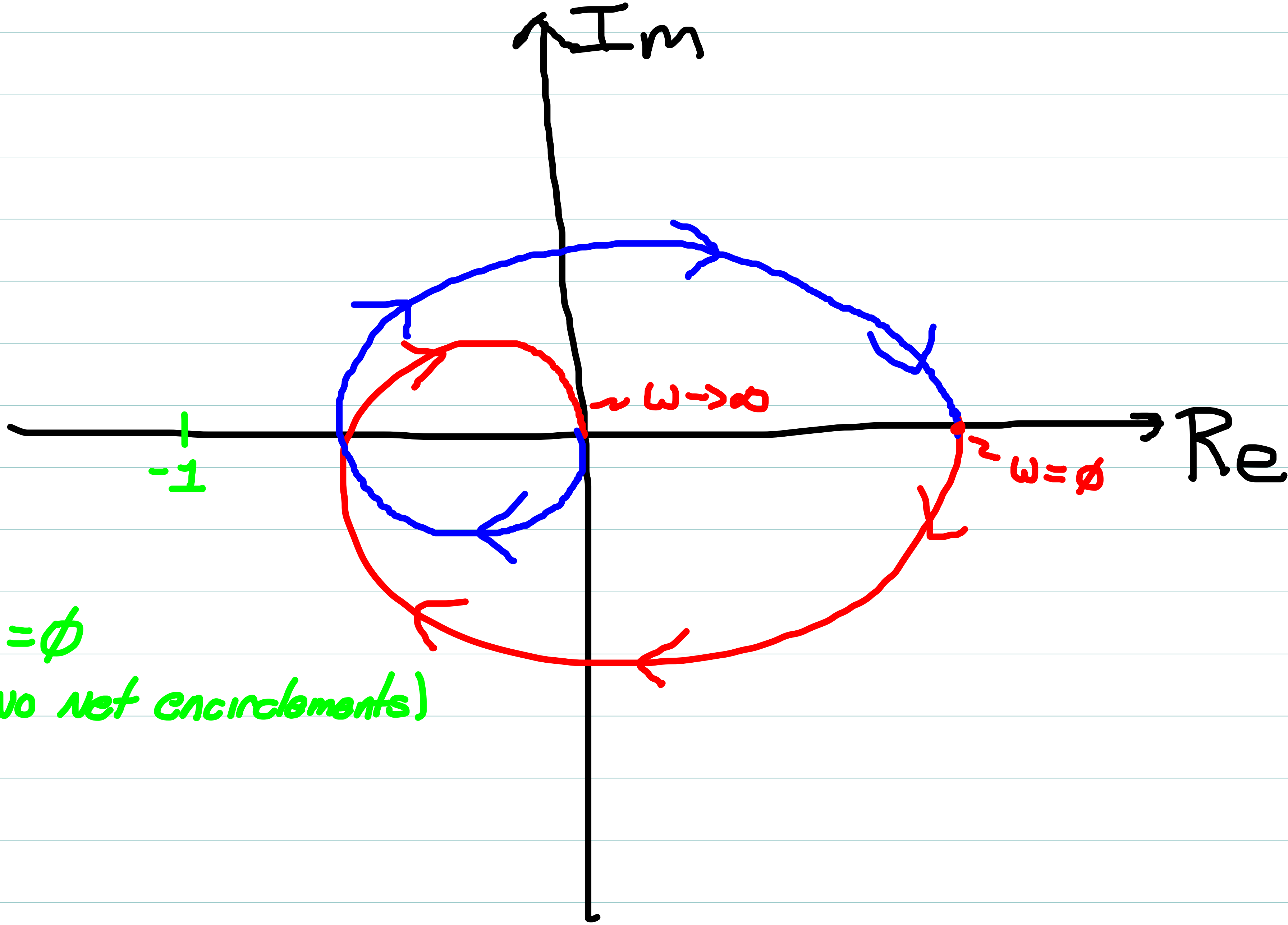
⇒ Count the number of complete loops the diagram makes around −1.

⇒ A clockwise loop counts as +1 encirclement
A counter-clockwise loop counts as −1 encirclement

⇒ Diagrams may have both CW or CCW loops around −1

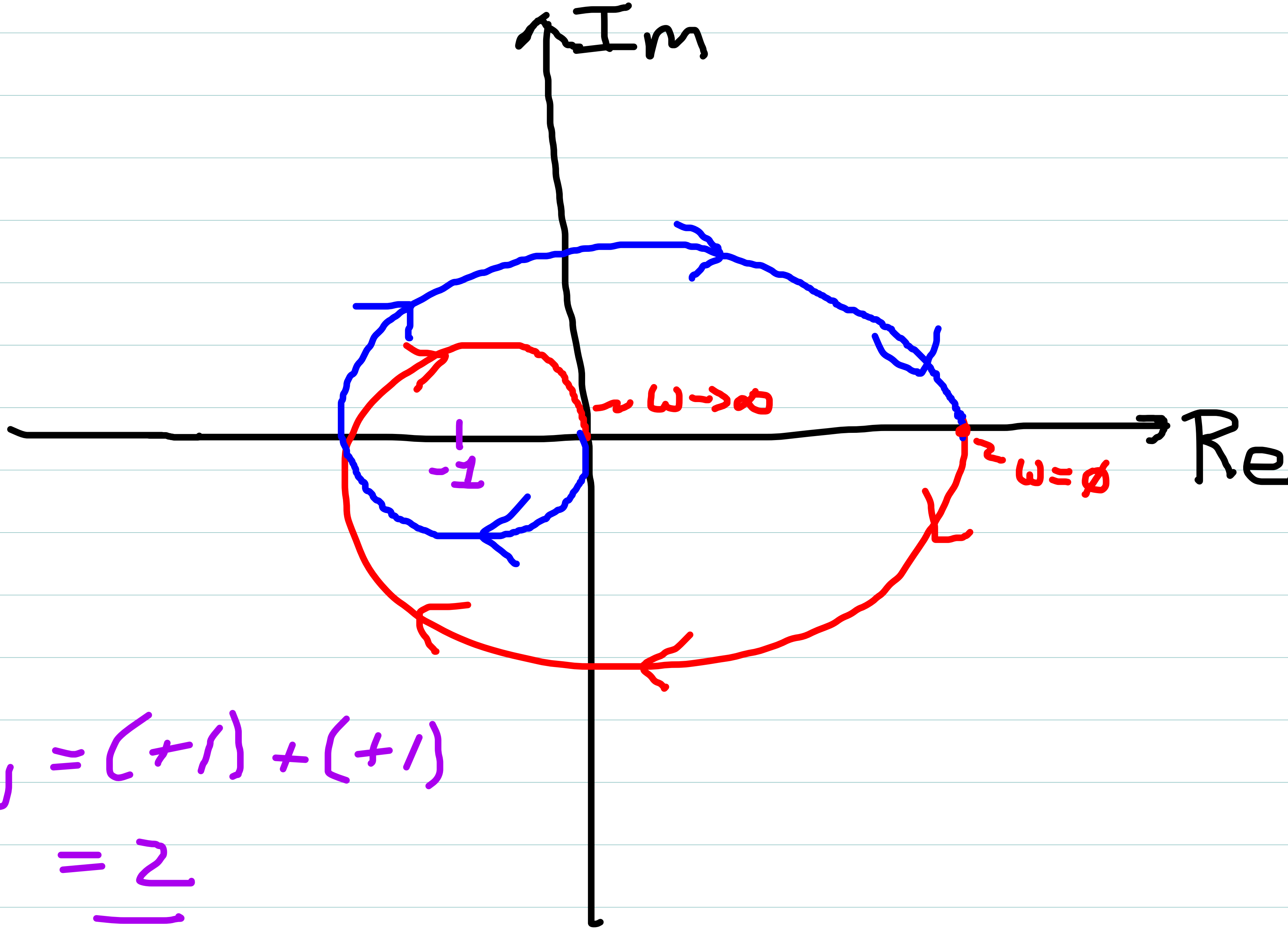
⇒ Let $N_{cw}(L)$ be the net number of CW encirclements for Nyquist diagram of L
(i.e. result of adding contribution of each loop using the ±1 convention above).

Example: $L(s) = \dfrac{K_B}{(\tau s+1)^3}$ $K_B, \tau > 0$



$N_{cw} = 0$
(no net encirclements)

Example: $L(s) = \frac{K_B}{(\tau s+1)^3}$ $K_B, \tau > 0$



$N_{CW} = (+1) + (+1)$

$= \underline{2}$

Counting Encirclements

⇒ Count the number of complete loops the diagram makes around -1.

⇒ A clockwise loop counts as $+1$ encirclement
A counter-clockwise loop counts as -1 encirclement

⇒ Diagrams may have both CW or CCW loops around -1

⇒ Let $N_{cw}(L)$ be the net number of CW encirclements for Nyquist diagram of L
(i.e. result of adding contribution of each loop using the ± 1 convention above).

Easy Way to Count Encirclements

"Ray trick"

=> Draw a ray radially outward from -1 in <u>any</u> direction

=> Looking along the ray, away from -1

=> Count $+1$ each time diagram crosses ray from <u>left</u> to right.

=> Count -1 each time ray is crossed right to left.

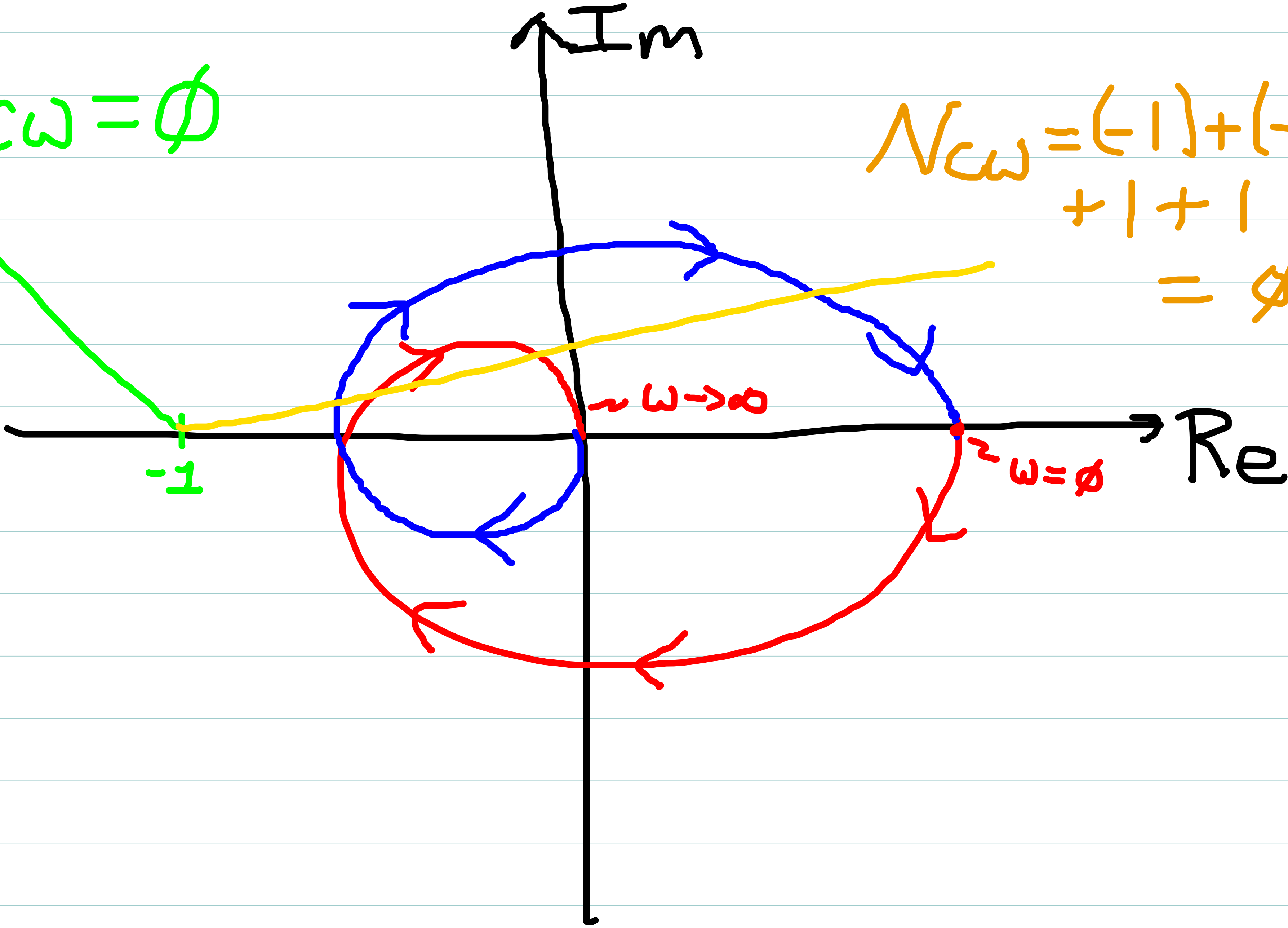
=> Same result regardless of ray direction

=> Choose direction with least number of intersections for easiest counting.

Example: $L(s) = \dfrac{K_B}{(\tau s+1)^3}$

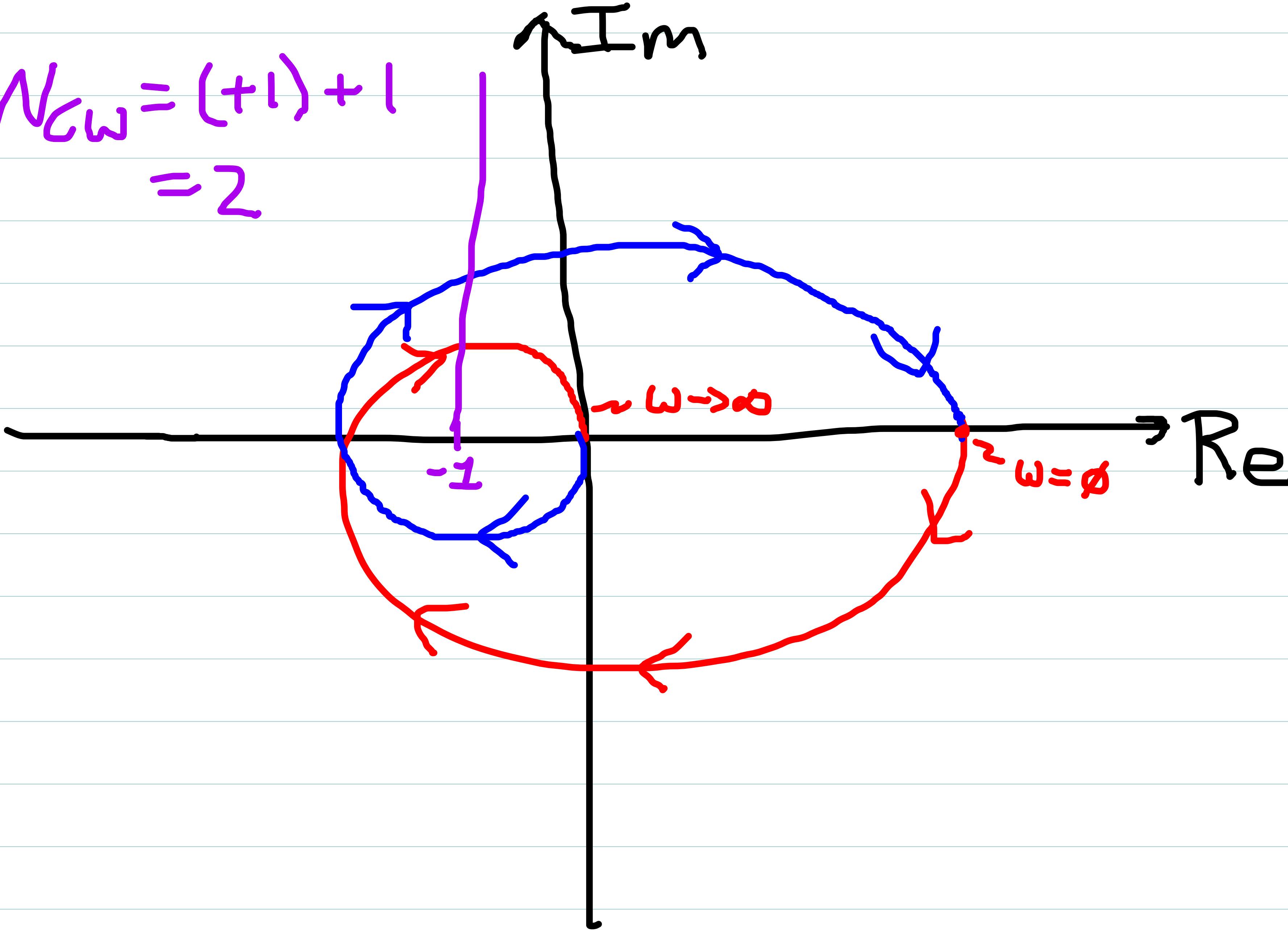
$N_{CW} = \emptyset$

$N_{CW} = (-1) + (-1) + 1 + 1 = \emptyset$



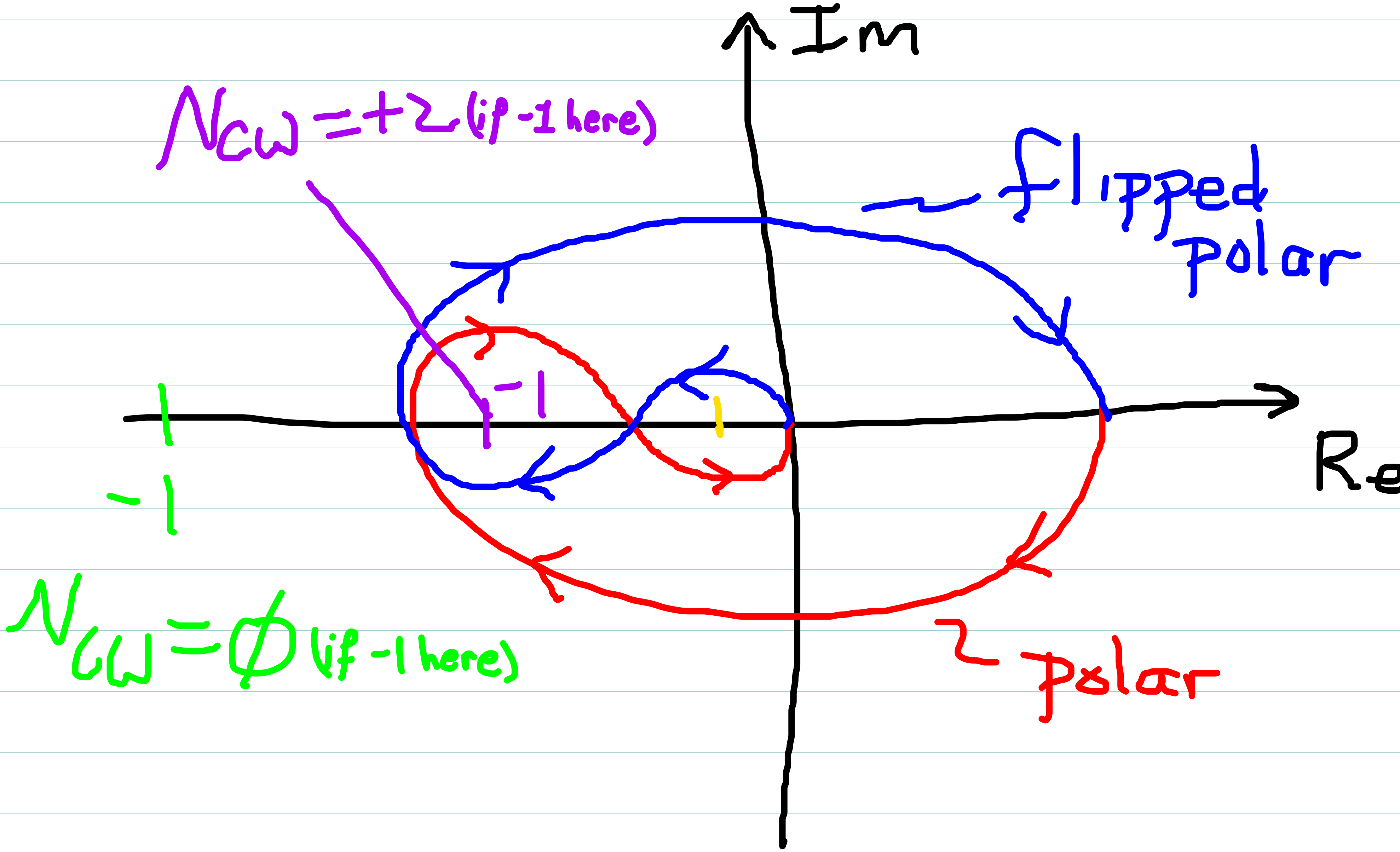
Example: $L(s) = \frac{K_B}{(\tau s+1)^3}$

$N_{CW} = (+1) + 1$
$= 2$



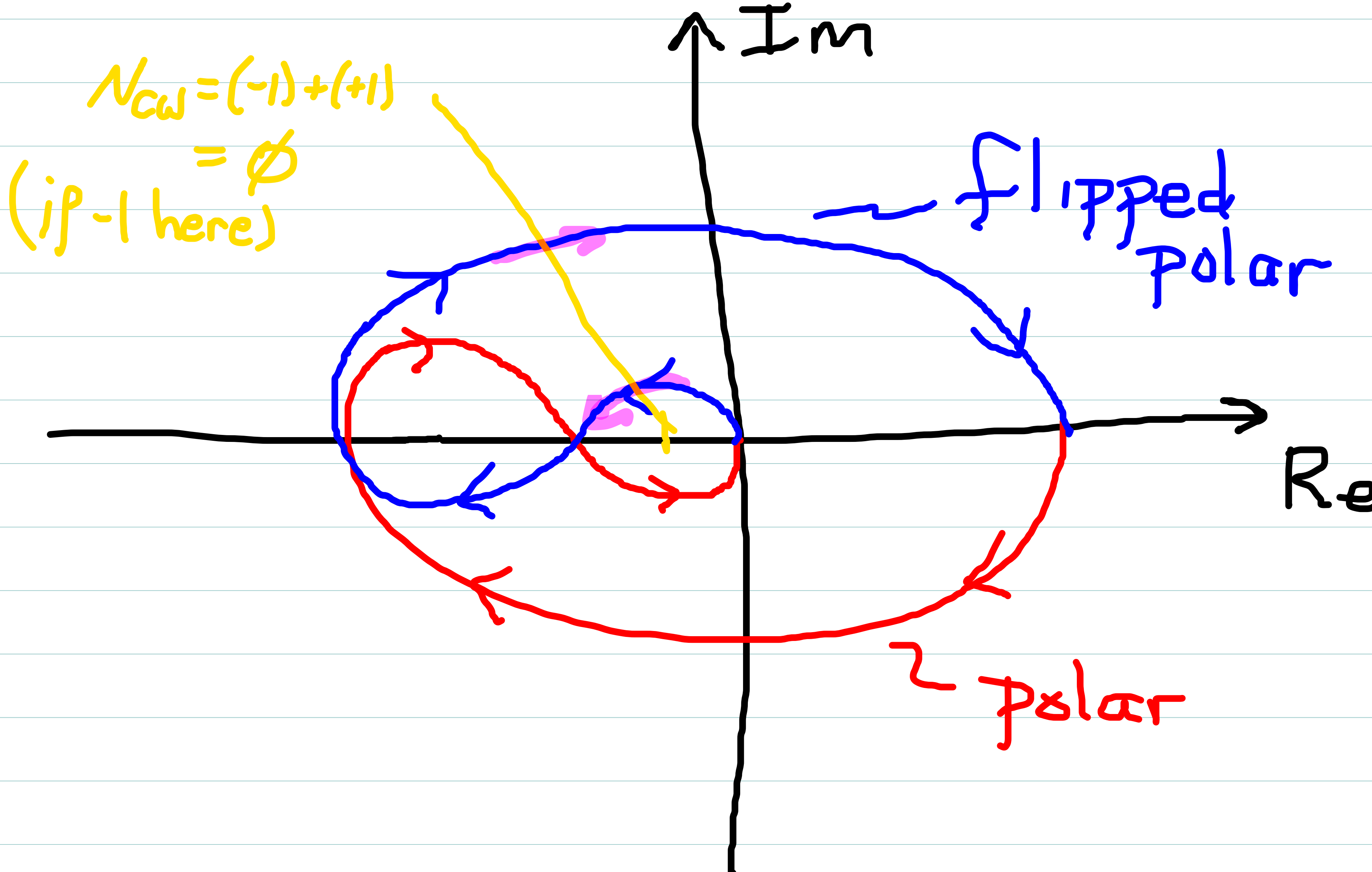
A more complicated Example:

$$L(s) = \frac{K_B(\tau_1 s+1)^2}{(\tau_2 s+1)^3} \qquad \tau_2 >> \tau_1 > \phi$$



A more complicated Example:

$$L(s) = \frac{K_B(\tau_1 s+1)^2}{(\tau_2 s+1)^3} \qquad \tau_2 \gg \tau_1 > \emptyset$$



Nyquist Stability Theorem

For an arbitrary transfer function $G(s)$, define

$$P_R(G) = \#\text{RHP (unstable) poles of } G(s)$$

Nyquist showed:

$$\boxed{N_{cw}(L) = P_R(T) - P_R(L)}$$

Re-arranging:

$$\underbrace{P_R(T)}_{\text{want to predict}} = \underbrace{P_R(L) + N_{cw}(L)}_{\text{Known}}$$

Note: $P_R(T) \geq 0$ <u>always</u>. If you compute $P_R(T) < 0$

$\Rightarrow$ you have drawn the diagram incorrectly, or

$\Rightarrow$ you have counted encirclements incorrectly.

Implication

$\Rightarrow$ We <u>must</u> have $P_R(T) = \emptyset$ (stable closed-loop system)

$\Rightarrow$ $\boxed{N_{cw}(L) = -P_R(L)}$ (Stability condition)

i.e. Nyquist diagram must show a net <u>negative</u> number of encirclements, equal to number of unstable poles of $L(s)$.

Recall negative CW encirclements are CCW encirclements

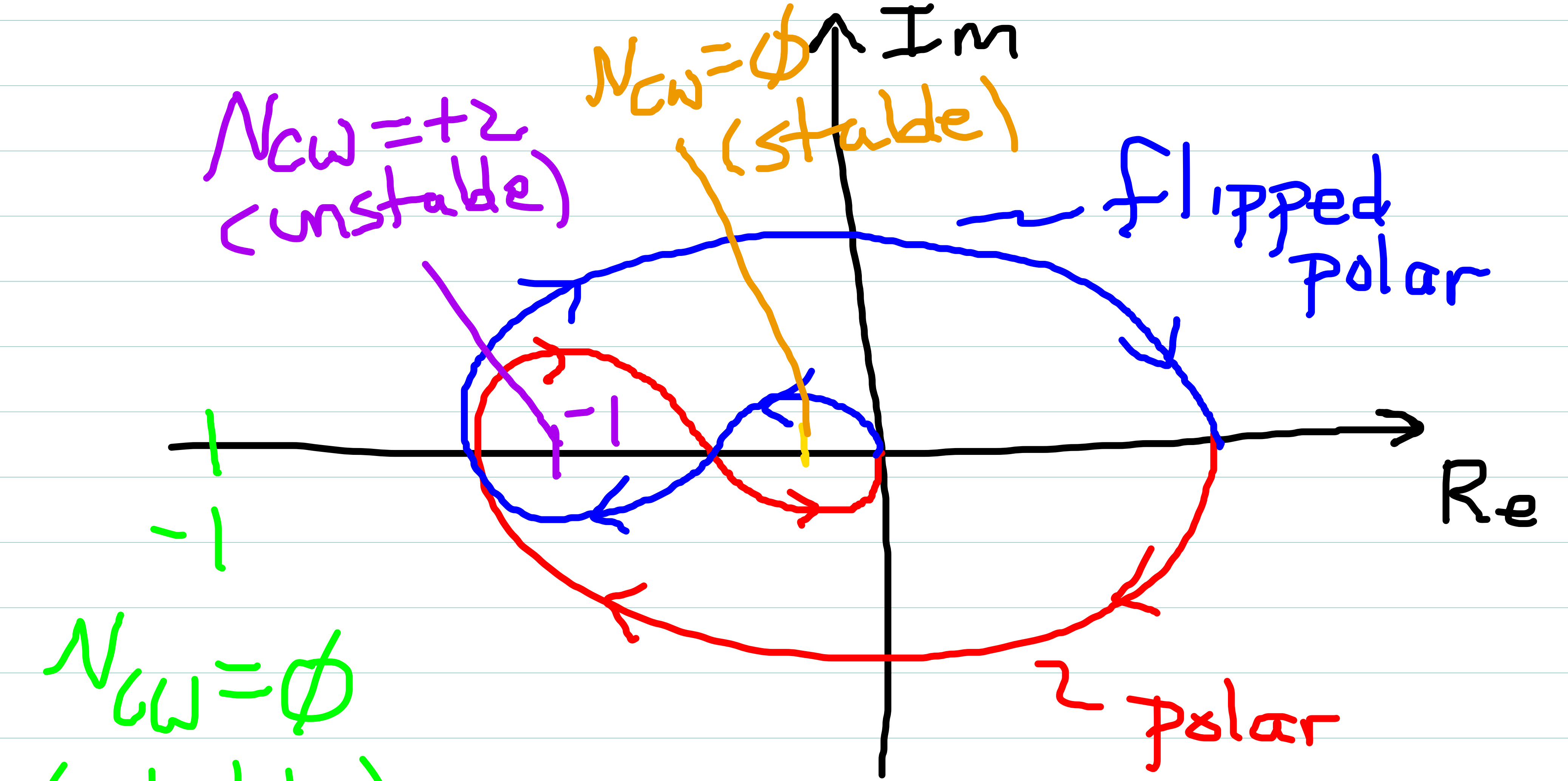
$\Rightarrow$ Nyquist diagram must show a net number of <u>CCW</u> encirclements equal to # unstable poles of $L(s)$

<u>Note</u>: if $P_R(L) = \emptyset$ ($L(s)$ is stable) then the diagram must show <u>NO</u> ($\emptyset$) net encirclements

A more complicated Example:

$$L(s) = \frac{K_B(\tau_1 s+1)^2}{(\tau_2 s+1)^3} \qquad \tau_2 \gg \tau_1 > 0$$

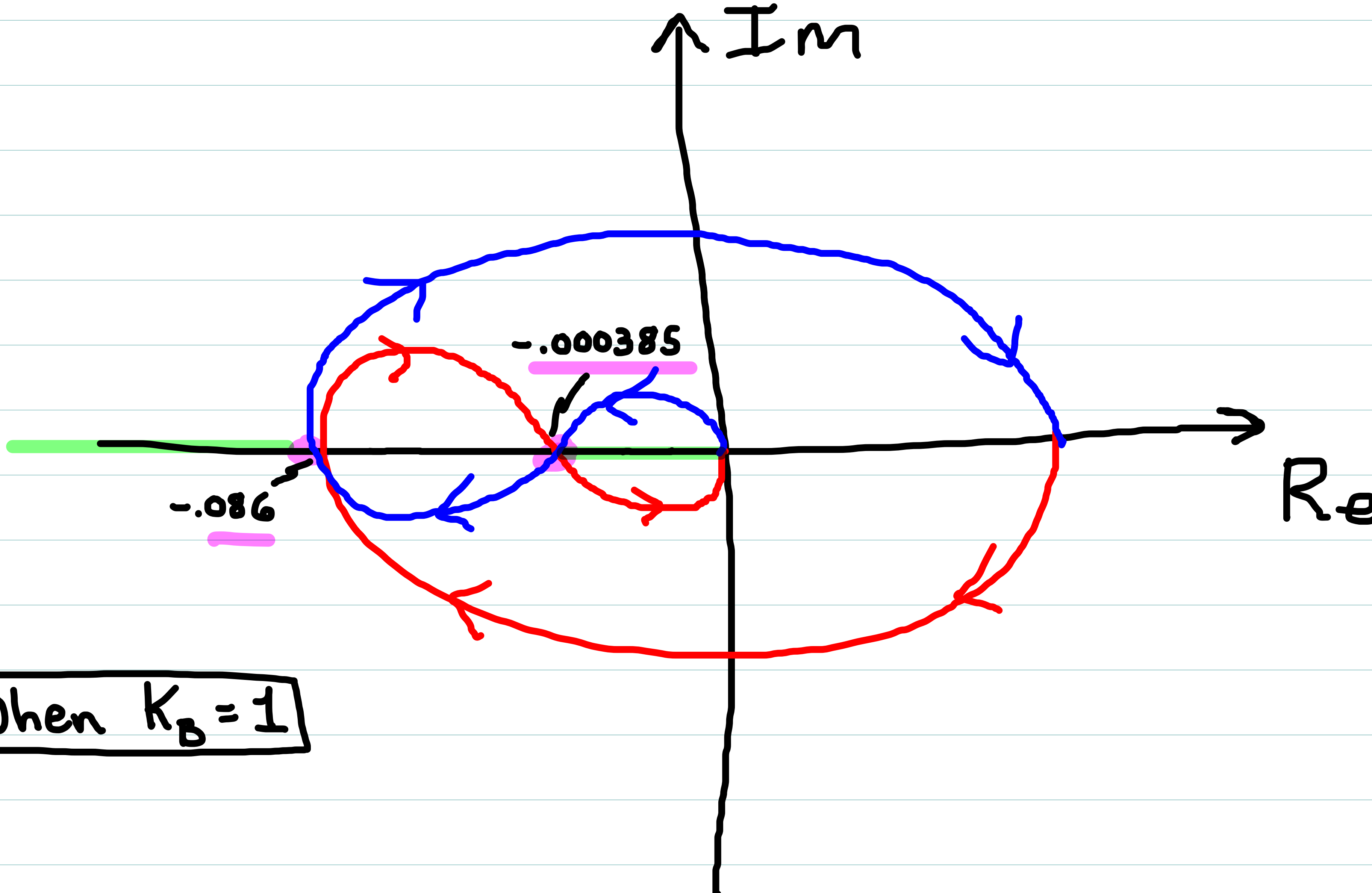
$P_R(L) = 0$



Stability depends on location of -1!

Effect of gain changes

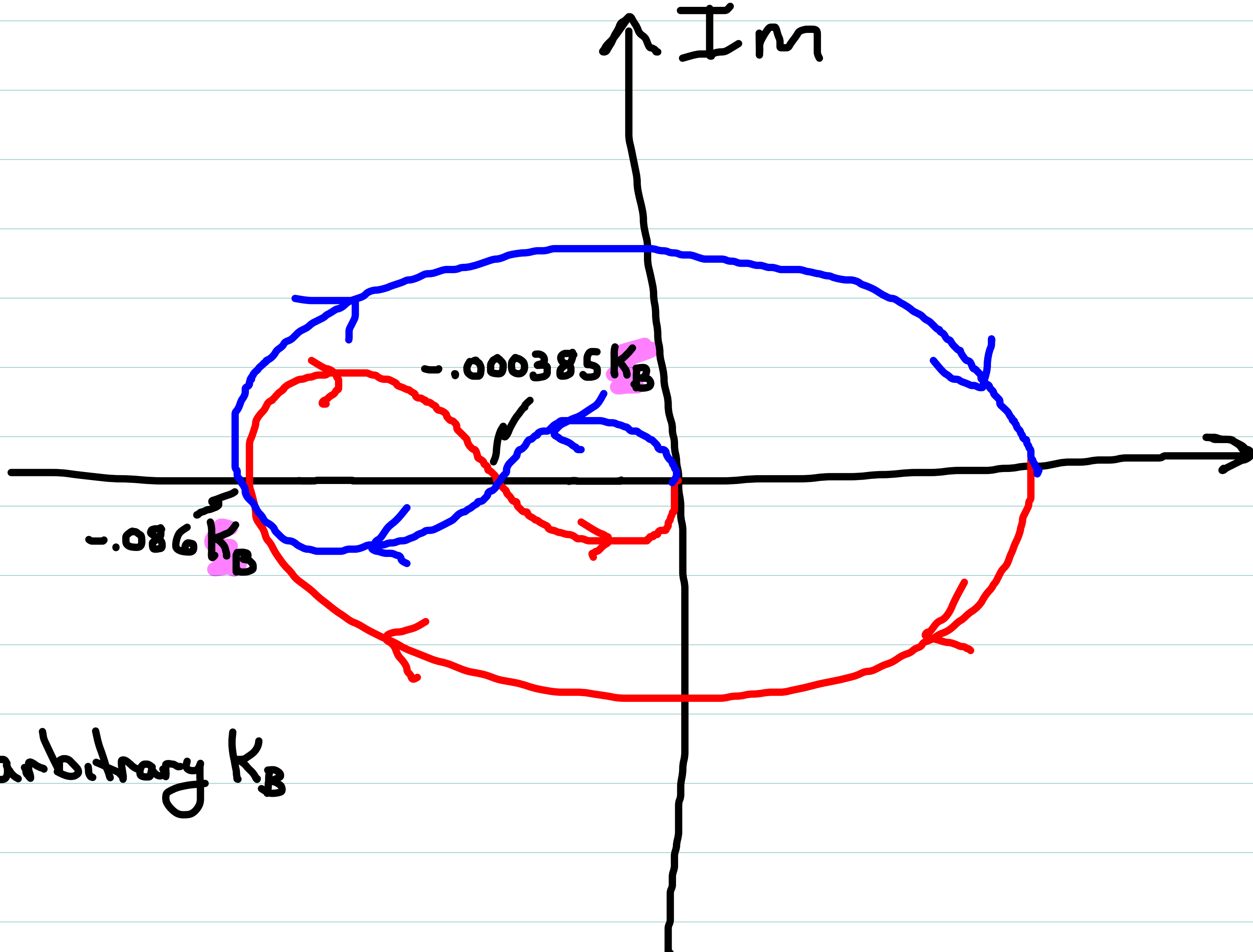
$$L(s) = \frac{K_B(\tau_1 s+1)^2}{(\tau_2 s+1)^3} \qquad \tau_1 = 10,\ \tau_2 = 1$$



$\boxed{\text{When } K_B = 1}$

Effect of gain changes

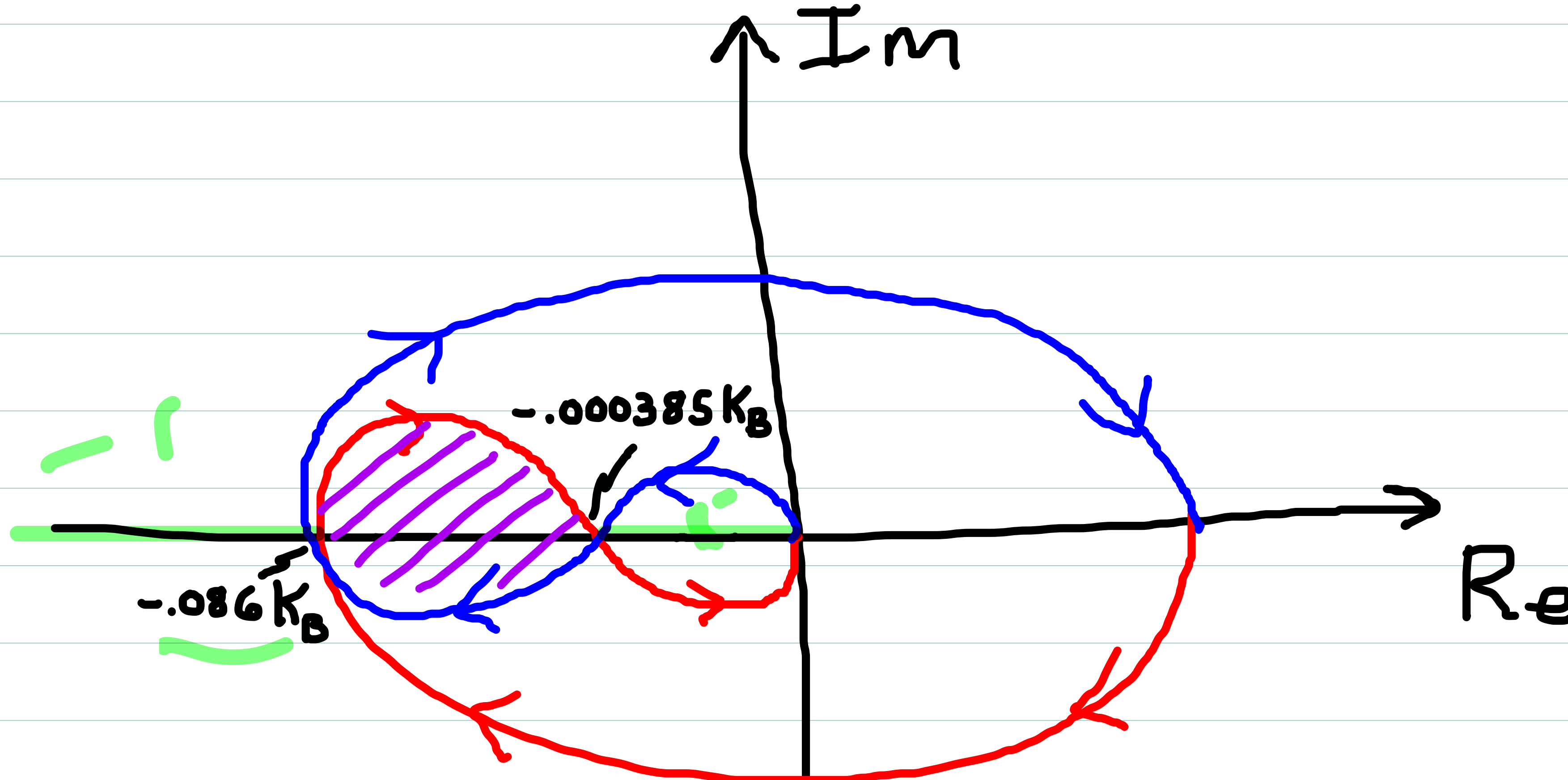
$$L(s) = \frac{K_B(\tau_1 s+1)^2}{(\tau_2 s+1)^3} \qquad \tau_1 = 10,\ \tau_2 = 1$$



for arbitrary K_B

Effect of gain changes

$$L(s) = \frac{K_B(\tau_1 s+1)^2}{(\tau_2 s+1)^3} \qquad \tau_1 = 10,\ \tau_2 = 1$$



for this system, T(s) stable unless -1 is in hashed area

Need: $-1 < -.086K_B$ or $-.0004K_B < -1$

Thus: Stable for $K_B < 1/.086 \cong 11.63$

or $K_B > 1/.000385 \cong 2597$

Note: Gain change is easy to accomplish with compensator:

$$H(s) = K \quad (\Rightarrow u(t) = Ke(t) \text{ "proportional" control})$$

$$L(s) = H(s)G(s) = KG(s) \text{ here}$$

$$\Rightarrow (K_B)_L = K(K_B)_G$$

However, gain change only affects "size" of polar (hence location of -1 relative to loops in Nyquist).

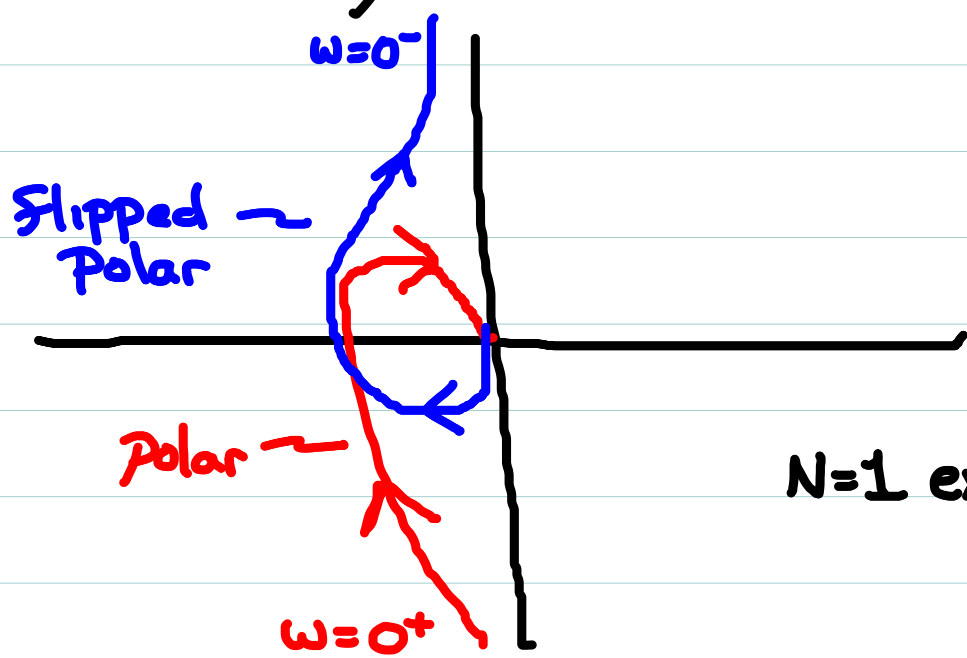
More substantial changes to polar/Nyquist diagram (changes to number and/or location of its loops) require also zeros/poles in $H(s)$.

Nyquist Diagram for $N > \emptyset$ systems

When $L(s)$ has type $N > \emptyset$ (one or more poles at origin) the first step to creating Nyquist diagram is same:

$\Rightarrow$ Draw polar of $L(j\omega)$
$\Rightarrow$ Flip polar about real Axis

However, the resulting diagram is Not connected; both halfs have "tails" parallel to coordinate AXES



N=1 example

Completing the diagram, $N > \emptyset$

$\Rightarrow$ Connect the $\omega = 0^-$ tail of flipped polar to $\omega = 0^+$ tail of original polar with a clockwise circular arc of total rotation $N\pi$

(i.e. ½ circle for every pole at origin in $L(s)$)

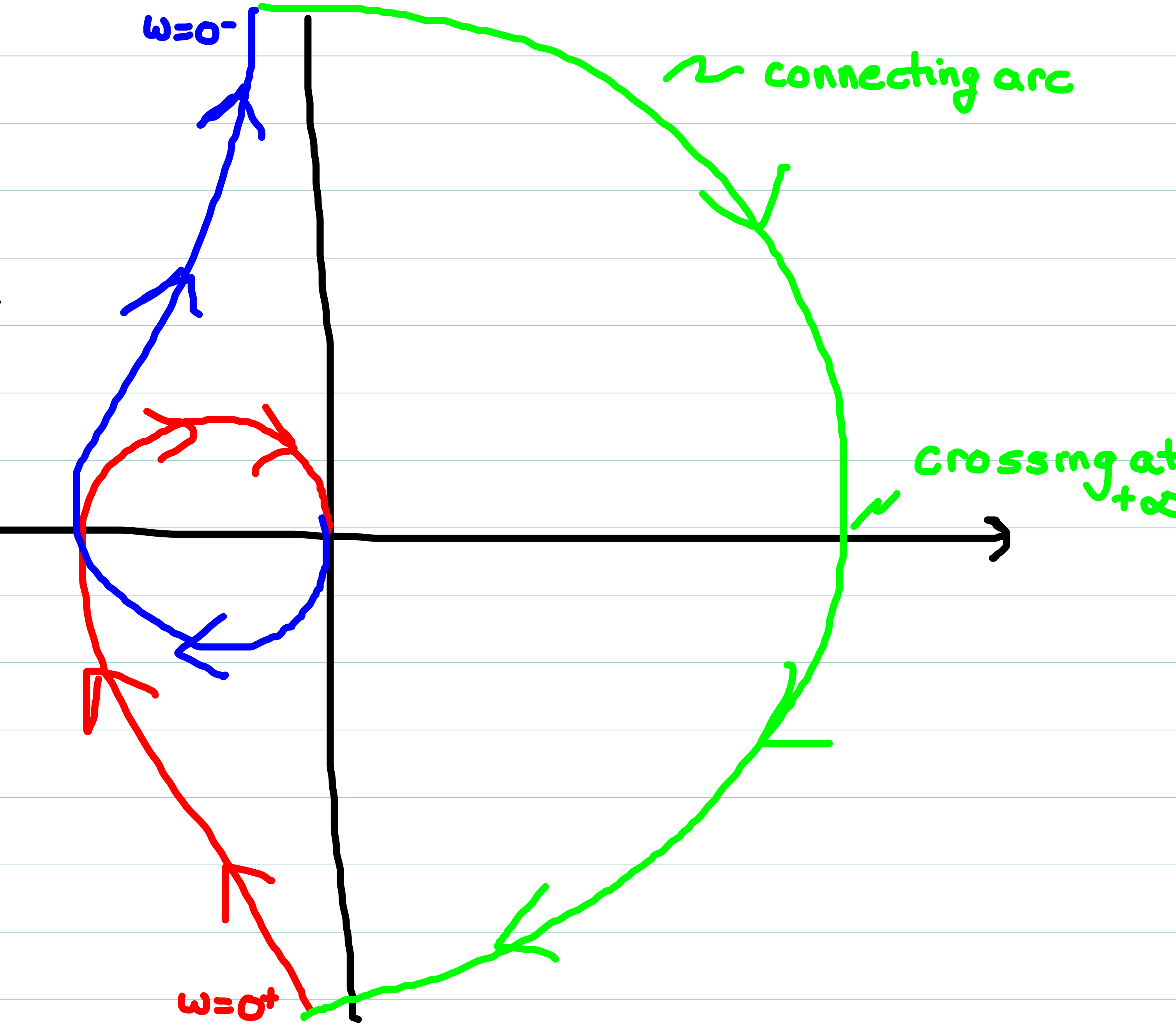
Note: Connecting arc has infinite radius, although we draw it as finite.

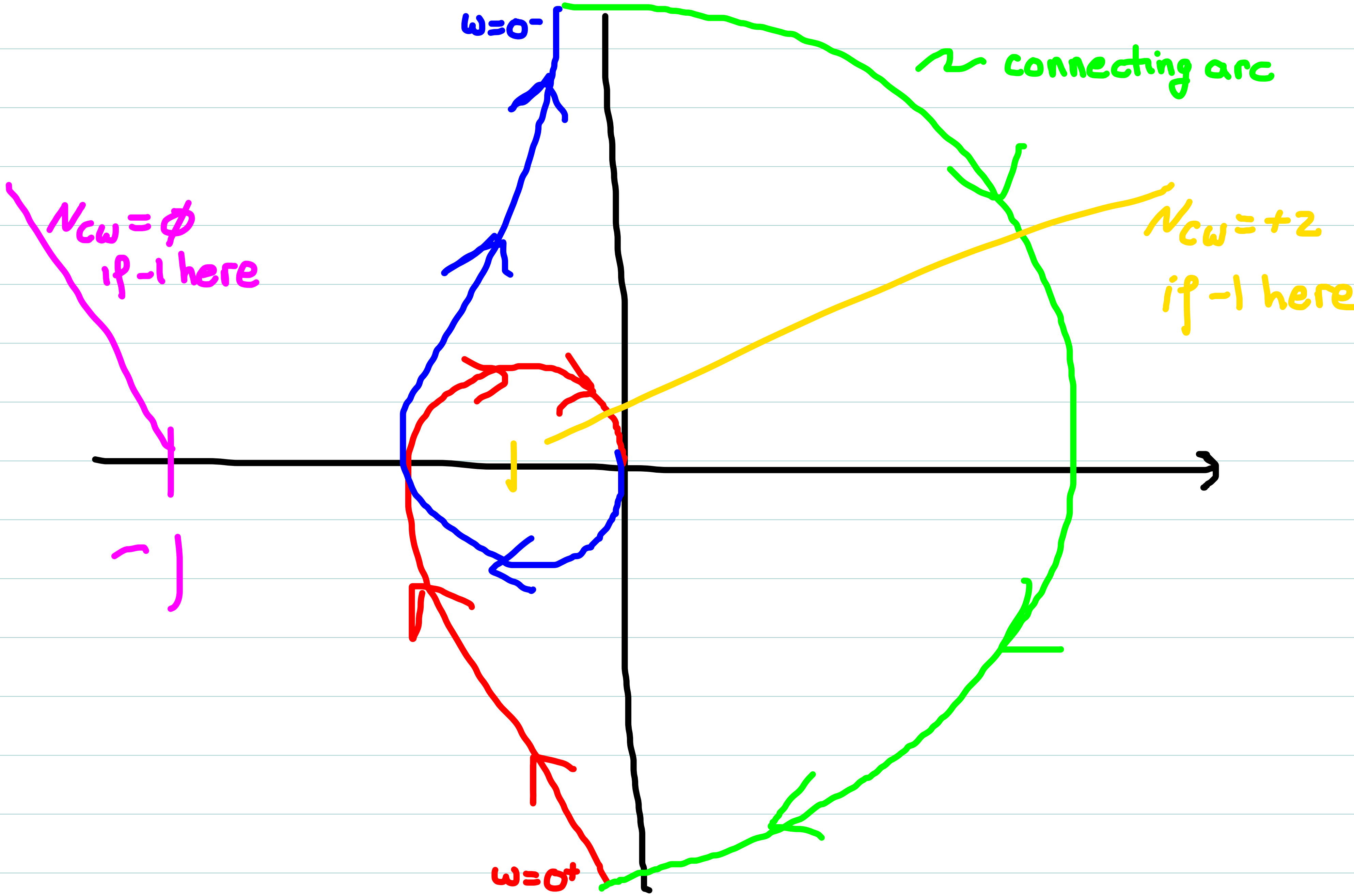
$\Rightarrow$ After connecting tails, compute $N_{cw}(L)$ as before.

N=1 system
for example

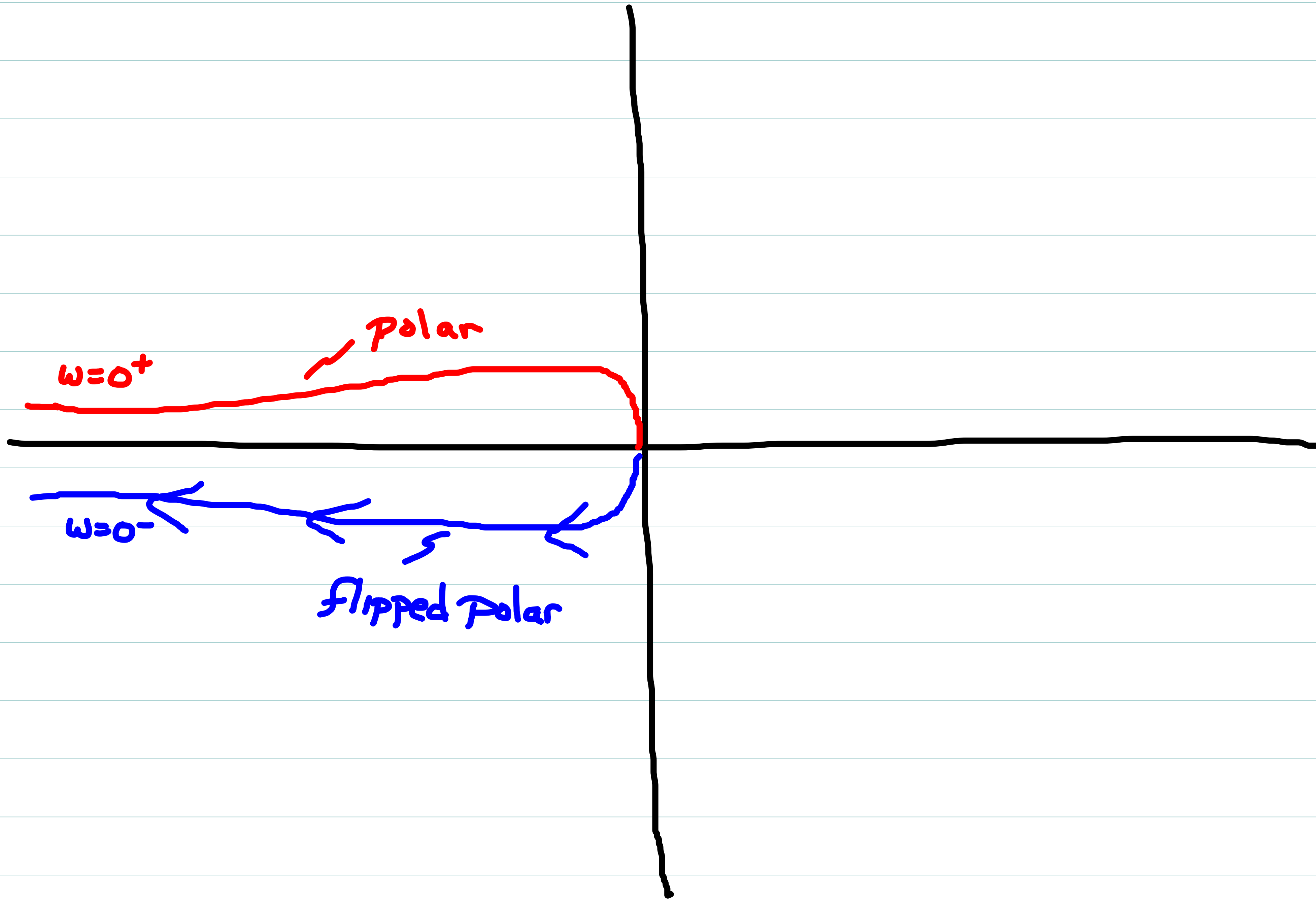
$$L(s) = \frac{K_B}{s(\tau s+1)^2}$$

$K_B, \tau > \emptyset$

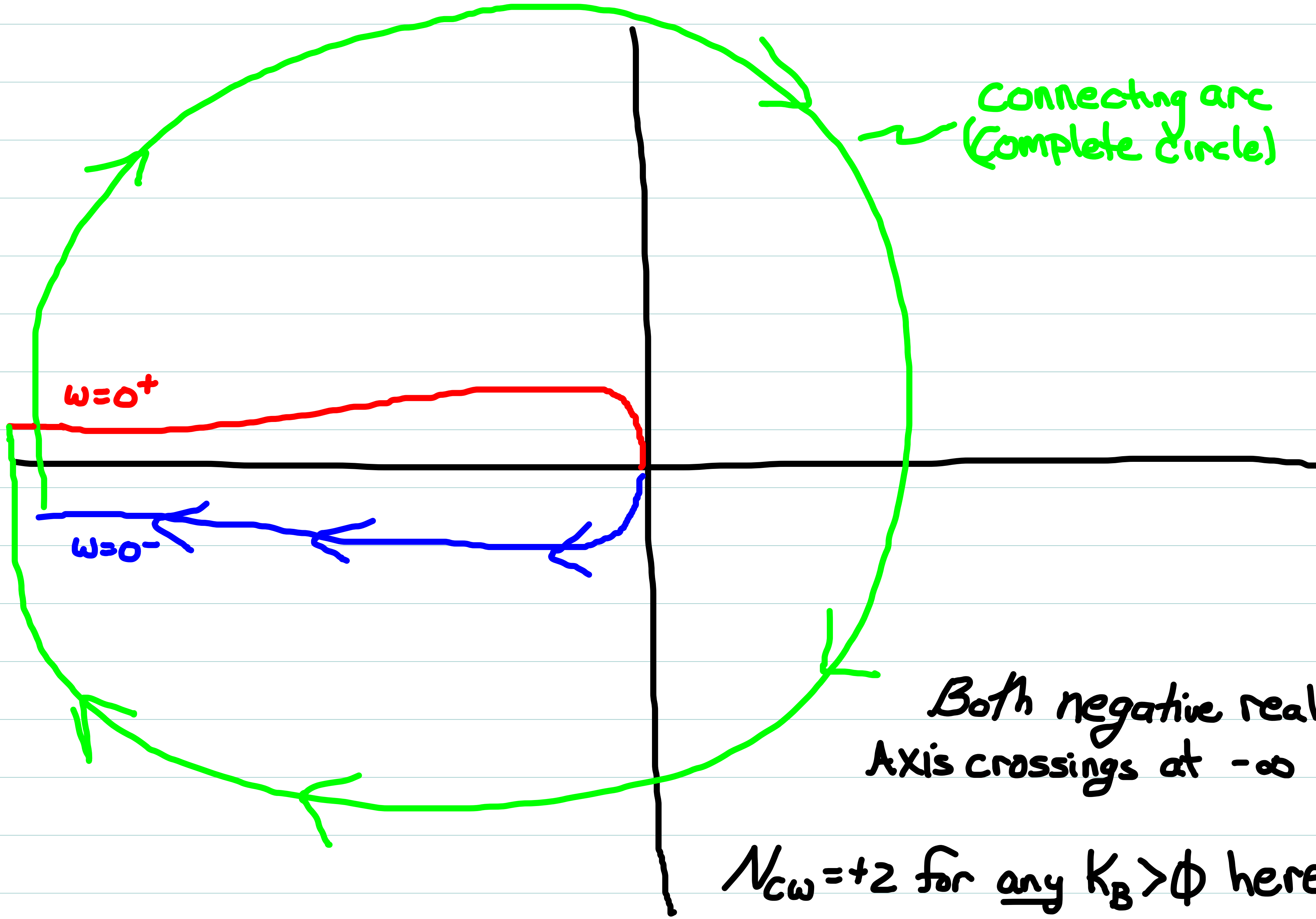



$\omega = 0^-$
connecting arc
$N_{cw} = \phi$
if -1 here
$N_{cw} = +2$
if -1 here
-1
1
$\omega = 0^+$

Be careful with N=2 systems!



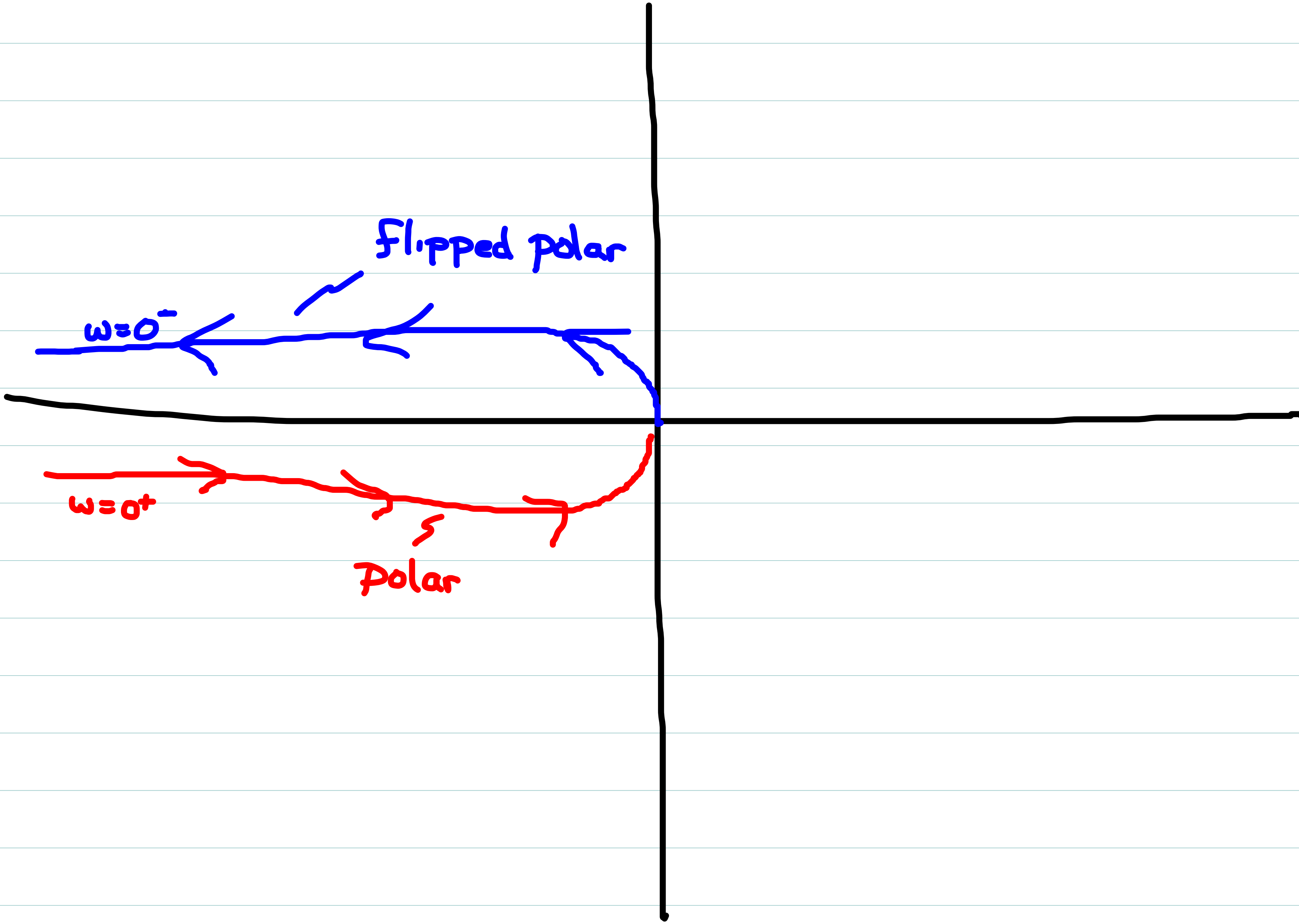
Be careful with N=2 systems!

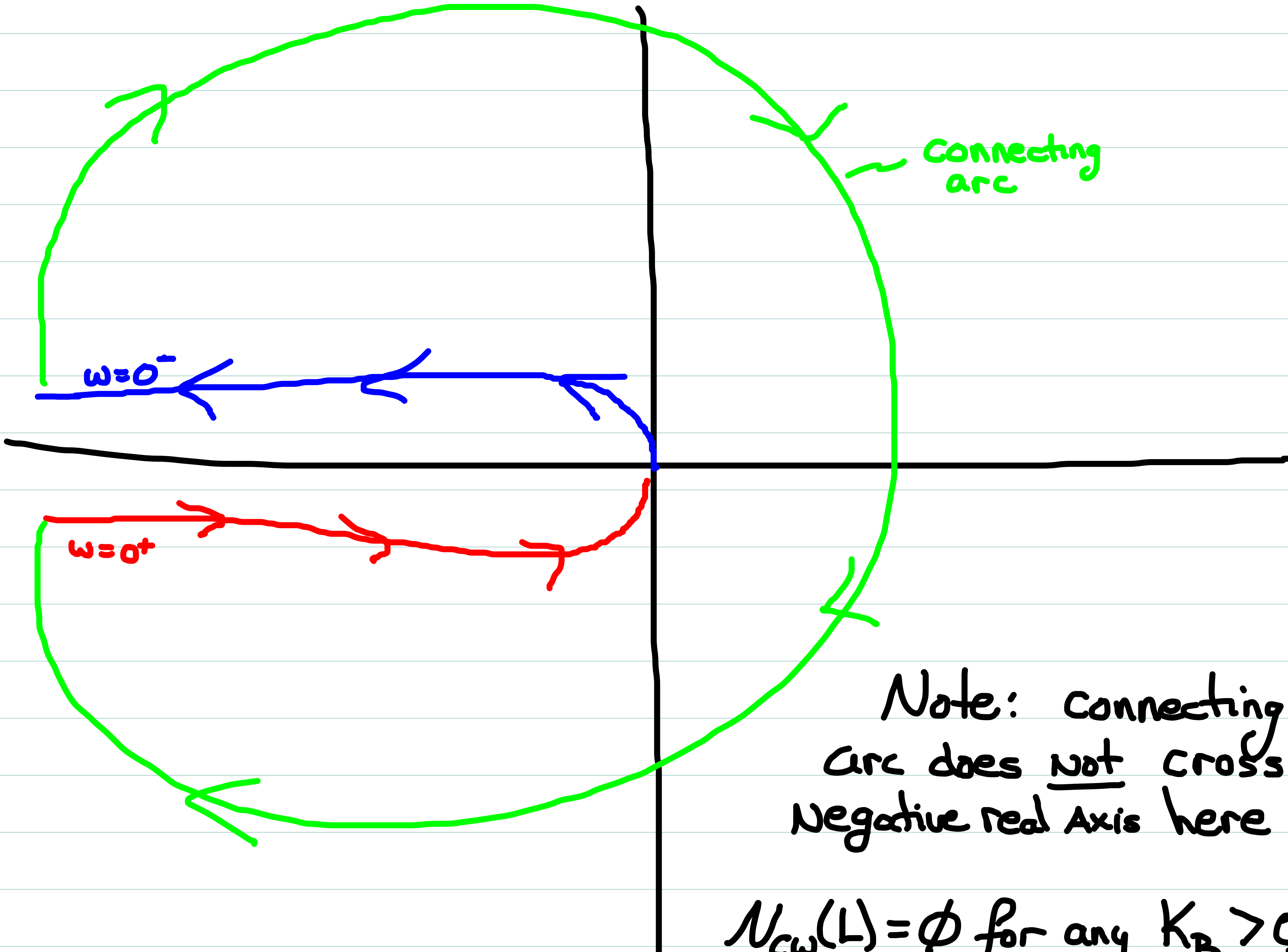


Both negative real Axis crossings at $-\infty$

$N_{cw} = +2$ for <u>any</u> $K_B > \phi$ here

An apparently similar system (N=2 still)





Note: connecting arc does <u>not</u> cross negative real axis here

$N_{cw}(L) = \emptyset$ for any $K_B > \emptyset$ here

Utility of gain/phase margin

⇒ a, γ measure how close polar comes to -1

⇒ If design is nominally stable (Nyquist shows required number of encirclements of -1), then

a, γ measure how much Nyquist plot can **Change** in a pure gain or phase fashion, before -1 would enter a different loop, changing the number of encirclements.

Thus: a, γ are measures of the "tolerance" of the system's stability to gain/phase changes in $L(s)$.

⇒ **Relative** stability measures.

Robustness (classical)

As measures of the tolerance of the control system stability to changes in shape of Nyquist, gain and phase margin are measures of the robustness of the design.

That is, the ability of the design to tolerate model errors which would create pure gain or pure phase errors in $L(s)$

Typically caused by errors in model of $G(s)$, since

$$L(s) = G(s)H(s)$$

and there is no uncertainty in $H(s)$.

Classical Robustness Requirements

A "robustly stable" design thus requires:

$\Rightarrow$ Correct number of Nyquist encirclements

AND $\Rightarrow$ Large $|a|$, $|\gamma|$

Typical professional requirements

$\Rightarrow |a_{dB}| \geq 6$ (i.e. $a > 6\,dB$ or $a < -6\,dB$)

$\Rightarrow |\gamma| \geq 30°$

Requirement on a is physically equivalent to no more than a factor of 2 uncertainty on gain of $G(s)$

Utility of gain/phase margin

⇒ a, γ measure how close polar comes to -1

⇒ If design is nominally stable (Nyquist shows required number of encirclements of -1), then

a, γ measure how much Nyquist plot can <u>change</u> in a pure gain or phase fashion, before -1 would enter a different loop, changing the number of encirclements.

Thus: a, γ are measures of the "tolerance" of the system's stability to gain/phase changes in $L(s)$.

⇒ <u>Relative</u> stability measures.

<u>Robustness</u> (classical)

As measures of the <u>tolerance</u> of the control system stability to changes in shape of Nyquist, gain and phase margin are measures of the <u>robustness</u> of the design.

That is, the ability of the design to tolerate model errors which would create pure gain or pure phase errors in $L(s)$

Typically caused by errors in model of $G(s)$, since

$$L(s) = G(s)H(s)$$

and there is no uncertainty in $H(s)$.

Classical Robustness Requirements

A "robustly stable" design thus requires:

$\Rightarrow$ Correct number of Nyquist encirclements

AND $\Rightarrow$ Large $|a|$, $|\gamma|$

Typical professional requirements

$\Rightarrow |a_{dB}| \geq 6$ (i.e. $a > 6\,dB$ or $a < -6\,dB$)

$\Rightarrow |\gamma| \geq 30°$

Requirement on a is physically equivalent to no more than a factor of 2 uncertainty on gain of $G(s)$

Recall: a, γ formally measure only how much Nyquist can change before encirclements change

(Assuming design is nominally stable, such changes would usually be bad!)

<u>By themselves</u> (separate from Nyquist) they are not reliable indicators of stability.

i.e. $a > \emptyset$ dB means Nyquist plot crosses neg. real axis to right of -1; $a < \emptyset$ dB means it crosses left of -1

Which is "better" (necessary for stability) depends on full Nyquist analysis.

However:

For a great many physical systems with:
a) $L(s)$ stable; b.) unique ω_γ; c.) $\gamma(L) > 0°$, the shape of Nyquist plot ensures $T(s)$ stable.

(True even for many $L(s)$ which satisfy c) but violate a) or b); however need to check actual Nyquist shape carefully here).

Common enough to be a major design guideline:

$\Rightarrow$ Design $H(s)$ to ensure that $L(s)$ has positive phase margin

$$\Rightarrow \measuredangle L(j\omega_\gamma) > -180°$$

Constraints for Stability

For most simple (and common) systems (and many not so simple systems) Nyquist will show stability if phase margin of $L(j\omega)$ is positive.

Design prescription: Add LHP zeros in H(s) to increase phase at magnitude xover.

Indeed, we will show using different techniques that it is rare that such a strategy would fail to stabilize.

$\Rightarrow$ Theoretically interesting counter-example: if G(s) has both a zero and a pole in RHP. Such a system may actually require a RHP pole in H(s) to stabilize!

Always check the Nyquist diagram when using simple Guidelines to design H(s)!

How much phase margin is "good"

Again, $\gamma > 30°$ is a typical minimum, and would ensure stability in common cases.

Why 30°? Is more better? Unfortunately, there is no simple correlation between freq. domain properties of $L(j\omega)$ and the <u>exact</u> location of poles of $T(s)$.

Nyquist tells us only $\mathrm{Re}\{p_k\} < \emptyset$ for each pole p_k of $T(s)$ when the stability condition is satisfied

However, we can develop some useful intuition correlating (γ, ω_γ) with transient properties of $T(s)$ by looking at some typical simple examples.

Simple Example

$$L(s) = \frac{K}{s(s+\alpha)} \Rightarrow T(s) = \frac{K}{s^2+\alpha s+K}$$

$(\alpha > \emptyset)$

If $K=\alpha^2\sqrt{2}$, then $\omega_\gamma = |\alpha|$ and $\gamma = 45°$
(prove this to yourself if not obvious!)

Closed-loop poles are complex since $\alpha^2 - 4K < 0$
$(\alpha^2 - 4\sqrt{2}\alpha^2) < 0$

and in fact closed-loop damping ratio is $\zeta = 0.42$ here.

$\Rightarrow$ Increasing K decreases γ here, and also decreases damping ratio of closed-loop poles.

$\Rightarrow$ Decreasing K increases γ here, and also increases DR of closed-loop poles

In fact, <u>for this system</u> we can show

$$\xi_{CL} \approx \frac{\gamma\,(\text{deg})}{100} \qquad (\text{for } 0 \lesssim \gamma \lesssim 70^\circ)$$

i.e. closed-loop damping ratio ξ_{CL} is directly proportional to the phase margin of L.

What about settling times for a step response of $T(s)$?
$\Rightarrow$ controlled by real parts of closed-loop poles.

Here the real parts are at $-\alpha/2 < 0$

$$t_s = \frac{4}{|\alpha/2|} = \frac{8}{|\alpha|} = \frac{8}{\omega_\gamma} \quad (\text{when } \gamma = 45^\circ \text{ as above})$$

i.e. t_s <u>inversely</u> prop. to ω_γ in this example.

Freq. Domain Constraints for Performance

When simple $\gamma(L) > \phi$ constraint works for stabilization, then typically:

$\Rightarrow$ larger γ gives higher damping for poles of $T(s)$

$\Rightarrow$ larger ω_γ gives faster settling time for $T(s)$ transients

Except for very simple systems, there are no direct mathematical connections between the freq. domain properties of $L(j\omega)$ (like γ and ω_γ) and the corresponding time domain properties of $T(s)$.

However, certain general trends have been found to hold:

For more complex systems, above observations do Not hold precisely, but general trends do:

Given 2 possible OL TF: $L_1(s)$, $L_2(s)$

a.) If L_1, L_2 have same ω_γ but suppose $\gamma(L_1) > \gamma(L_2)$ then the closed-loop TFs T_1 and T_2 will have comperable settling times, but T_1 will have a higher Damping ratio

b.) If L_1 and L_2 have same phase margin but different ω_γ, $\omega_{\gamma_1} > \omega_{\gamma_2}$, then T_1 and T_2 will have comperable damping, but T_1 will settle faster than T_2

$\Rightarrow$ Design guideline: make γ, ω_γ as large as possible.

Intro. to Controller Design

Stability and healthy margins are just the first of many different constraints for a good design

Often the constraints conflict, and we must use our judgement to achieve an acceptable trade-off

The general design process is typically:

1.) Look at Bode/Nyquist of $G(s)$.

2.) Determine how plots in 1.) must be changed to achieve desired design goals.

3.) From required changes in 2.), determine the ZPK structure $H(s)$ must have.

Controller Implementation, I

But can we actually have *any* H(s) we want?

Unfortunately no. There are *implementation* constraints:

i.e., can we actually calculate u(t) from e(t) *in real time*

Note that we *do not* calculate u(t) from

$$u(t) = \mathcal{L}^{-1}\{H(s)E(s)\}$$

Why not?

- $y_d(t)$ Not always Known ahead of time
  (may come from pilot inputs)

- y(t) cannot be predictedly exactly due to
  inaccurate model or external "disturbances"

Controller Implementation, II

So how do we implement the controller? By solving in real time the differential equation relating $u(t)$ to $e(t)$.

There are mathematical constraints under which this is possible, and these in turn constrain the "allowable" $H(s)$.

$$U(s) = H(s)\,E(s) = H(s)\left[Y_d(s) - Y(s)\right]$$

Suppose

$$H(s) = \frac{a(s)}{b(s)}$$

$a(s), b(s)$ polynomials in s. Then

$$b(s)\,U(s) = a(s)\,E(s)$$

Controller Implementation, II

$$U(s) = H(s)\,E(s)$$

Suppose
$$H(s) = \frac{a(s)}{b(s)}$$

$a(s), b(s)$ polynomials in s. Then

$$\mathcal{L}^{-1}\left\{ b(s)\,U(s) = a(s)\,E(s) \right\}$$

Gives a differential equation relating $u(t)$ ("output") to $e(t)$ ("input")

This diff'l equation must be <u>solvable in real time</u> using <u>only</u> measurements of

$$e(t) = y_d(t) - y(t)$$

Example

$$H(s) = \frac{6(s+1)^2}{(s+3)(s+5)} = \frac{\overbrace{(6s^2+12s+6)}^{a(s)}}{\underbrace{(s^2+8s+15)}_{b(s)}}$$

$$\Rightarrow (s^2+8s+15)U(s) = (6s^2+12s+6)E(s)$$

$$\Rightarrow \ddot{u}(t) + 8\dot{u}(t) + 15u(t) = 6\ddot{e}(t) + 12\dot{e}(t) + 6e(t)$$

DE which must be solved during operation of controller on vehicle.

Must be soluble using only measured $e(t)$.

$\dot{e}(t)$, $\ddot{e}(t)$ terms Not assumed to be available!

Note these terms come from zeros of $H(s)$...

Real-time implementation constraint

Computation of $u(t)$ must require only knowledge of $e(t)$,
(Not $\dot{e}(t)$, $\ddot{e}(t)$, etc.)

But note the DE from $\mathcal{L}^{-1}\{b(s)U(s) = a(s)E(s)\}$ will have derivatives of $e(t)$ on RHS.

If you think about it, this would seem to suggest $H(s)$ could never have any zeros [i.e. $a(s)$ must be a constant]

Fortunately this is not the case, if we think a little more deeply:

$$H(s)E(s) = \boxed{\left[\frac{a(s)}{b(s)}\right]}E(s) = \left[\underline{d(s) + \boxed{\frac{a'(s)}{b(s)}}}\right]E(s)$$

where $d(s)$ is the quotient polynomial of $\frac{a(s)}{b(s)}$ and $a'(s)$ is the remainder polynomial.

If $\deg\{a(s)\} \geq \deg\{b(s)\}$, then

$$H(s) = \frac{a(s)}{b(s)} = d(s) + \frac{a'(s)}{b(s)}$$

where $\deg\{d(s)\} = \deg\{a(s)\} - \deg\{b(s)\}$

and $\deg\{a'(s)\} = \deg\{b(s)\} - 1$

Since $\deg\{a'(s)\} < \deg\{b(s)\}$ we can expand

$$\boxed{\frac{a'(s)}{b(s)} = \sum_{k=1}^{M} \frac{c_k}{(s-\ell_k)}}$$

(M = # poles of $H(s)$ here!)

where ℓ_k are roots of $b(s)$ (poles of $H(s)$)

and:

$$c_k = \left\{ (s-\ell_k) \left[\frac{a'(s)}{b(s)} \right] \right\}_{s=\ell_k}$$

If instead $\deg\{a(s)\} < \deg\{b(s)\}$ then

$$H(s) = \frac{a(s)}{b(s)} = \sum_{k=1}^{M} \frac{c_k}{(s-\ell_k)} \quad \text{directly}$$

so that $d(s) = \emptyset$ and $a'(s) = a(s)$

in the above.

PFE of $H(s)$!

Thus generally:

$$H(s)E(s) = \boxed{\left[d(s) + \sum_{k=1}^{M} \frac{c_k}{s-\ell_k}\right]} E(s)$$

or $$H(s)E(s) = d(s)E(s) + \sum_{k=1}^{M} c_k \left[\frac{1}{s-\ell_k}\right] E(s)$$

Look at each of the terms individually

$$U(s) = H(s)E(s) = d(s)E(s) + \sum_{k=1}^{M} c_k \left[\frac{1}{s-\ell_k}\right]E(s)$$

Introduce:

$$X_k(s) = \left(\frac{1}{s-\ell_k}\right)E(s)$$

So that

$$U(s) = d(s)E(s) + \sum_{k=1}^{M} c_k X_k(s)$$

and

$$u(t) = \mathcal{L}^{-1}\{d(s)E(s)\} + \sum_{k=1}^{M} c_k x_k(t)$$

with $x_k(t) = \mathcal{L}^{-1}\{X_k(s)\}$

But note from above: $(s-\ell_k)X_k(s) = E(s)$

$$\Rightarrow \dot{x}_k(t) - \ell_k x_k(t) = e(t)$$

DE which only involves $e(t)$!

Thus, generally the control calculations required by $H(s)$ can be implemented using:

$$\boxed{u(t) = \mathcal{L}^{-1}\{d(s)E(s)\} + \sum_{k=1}^{M} c_k x_k(t)}$$

where

$$\boxed{\dot{x}_k(t) = \ell_k x_k(t) + e(t)} \leftarrow M \text{ different } 1^{st} \text{ order DEs for } x_k(t)$$

ℓ_k are poles of $H(s)$, and c_k are the residues:

$$c_k = \left\{(s-\ell_k)\left[\frac{a'(s)}{b(s)}\right]\right\}_{s=\ell_k}$$

What about $\mathcal{L}^{-1}\{d(s)E(s)\}$? Recall $d(s)$ is a polynomial with degree $\deg\{a(s)\} - \deg\{b(s)\}$

If $\deg\{d(s)\} > 1$ $\quad (\deg\{a(s)\} > \deg\{b(s)\})$

i.e.

$$d(s) = d_0 + d_1 s + d_2 s^2 + \cdots$$

then
$$\mathcal{L}^{-1}\{d(s)E(s)\} = d_0 e(t) + \underbrace{d_1 \dot{e}(t) + d_2 \ddot{e}(t) + \cdots}_{\text{Cannot be implemented with assumed measurements.}}$$

Thus, these add'l terms can only be implemented if

$$\deg\{d(s)\} = \emptyset$$ (i.e. $d(s)$ is just a constant)

Or equivalently $\deg\{\underbrace{a(s)}_{\text{numerator of } H(s)}\} \leq \deg\{\underbrace{b(s)}_{\text{Denom } H(s)}\}$

Relative Degree

The relative degree of a transfer function $G(s)$ is:

$$\rho(G) = \text{Degree of Denom poly} - \text{Degree of num poly}$$
$$= \#\text{poles of } G - \#\text{zeros of } G$$

From the above, the constraint for real-time implementation of compensator $H(s)$ is:

$$\rho(H) \geq \emptyset$$

$$\boxed{\text{i.e. } H(s) \text{ must have No more zeros than it has poles.}}$$

$\Rightarrow$ Will be a significant constraint on our designs!

Thus, generally the control calculations required by $H(s)$ can be implemented using:

$$\boxed{u(t) = \mathcal{L}^{-1}\{d(s)E(s)\} + \sum_{k=1}^{M} c_k x_k(t)}$$

where $\dot{x}_k(t) = \ell_k x_k(t) + e(t)$ ← M <u>different</u> 1^{st} order DEs. for $x_k(t)$

ℓ_k are poles of $H(s)$, and c_k are the residues:

$$c_k = \left\{ (s-\ell_k) \left[\frac{a'(s)}{b(s)} \right] \right\}_{s=\ell_k}$$

What about $\mathcal{L}^{-1}\{d(s)E(s)\}$? Recall $d(s)$ is a polynomial with degree $\deg\{a(s)\} - \deg\{b(s)\}$

If $\deg\{d(s)\} > 1$ $\left(\deg\{a(s)\} > \deg\{b(s)\}\right)$

i.e.

$$d(s) = d_0 + d_1 s + d_2 s^2 + \cdots$$

This term is OK!

$$\text{then } \mathcal{L}^{-1}\{d(s)E(s)\} = d_0 e(t) + \underbrace{d_1 \dot{e}(t) + d_2 \ddot{e}(t) + \cdots}$$

Cannot be implemented with assumed measurements.

Thus, these add'l terms can only be implemented if

$$\deg\{d(s)\} = \emptyset$$ (i.e. $d(s)$ is just a <u>constant</u>)

Or equivalently $\deg\{\underline{a(s)}\} \leq \deg\{\underline{b(s)}\}$

numerator of $H(s)$

Denom $H(s)$

Eq'ns are naturally in state-space format:

$$\Longrightarrow \begin{cases} \dot{x}_K(t) = \ell_K x_K(t) + e(t) \\ u(t) = \sum c_K x_K(t) + d_0 e(t) \end{cases}$$

$K = 1, \dots, M = \#$ Poles $H(s)$

ℓ_K = pole of $H(s)$

Note this is a state-space model for $H(s)$!

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ \vdots \\ x_m(t) \end{bmatrix} = \underbrace{\begin{bmatrix} \ell_1 & 0 & \cdots & 0 \\ 0 & \ell_2 & 0 \cdots & 0 \\ 0 & \cdots & 0 & \ell_m \end{bmatrix}}_{A_c}\begin{bmatrix} x_1(t) \\ \vdots \\ x_m(t) \end{bmatrix} + \underbrace{\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}_{B_c} e(t)$$

$$u(t) = \underbrace{\begin{bmatrix} c_1 & c_2 & \cdots & c_m \end{bmatrix}}_{C_c}\begin{bmatrix} x_1(t) \\ \vdots \\ x_m(t) \end{bmatrix} + \underbrace{d_0}_{D_c} e(t)$$

D_c "0" if $\rho(H) > 0$

$[A_c, B_c, C_c, D_c]$ state-space "realization" of Compensator TF $H(s)$

(One possible form => others are possible)

Relative Degree

The relative degree of a transfer function $G(s)$ is:

$$\rho(G) = \text{Degree of Denom poly} - \text{Degree of num poly}$$
$$= \#\text{poles of } G - \#\text{zeros of } G$$

From the above, the constraint for real-time implementation of compensator $H(s)$ is:

$$\rho(H) \geq 0$$

$\boxed{\text{i.e. } H(s) \text{ must have no more zeros than it has poles.}}$

$\Rightarrow$ Will be a significant constraint on our designs!

Examples, cont

1.) $H(s) = 6(s+1)^2 = 6s^2 + 12s + 6$ $\quad (p=-2)$ ✗

$\Rightarrow u(t) = \underline{6\ddot{e}(t) + 12\dot{e}(t)} + 6e(t)$ **Not implementable**

2.) $H(s) = \dfrac{6(s+1)^2}{(s+3)} = 6s - 6 + \dfrac{24}{s+3}$ $\quad (p=-1)$ ✗

$$\Rightarrow \begin{cases} u(t) = \underline{6\dot{e}(t)} - 6e(t) + 24x_1(t) \\ \dot{x}_1(t) = -3x_1(t) + e(t) \end{cases}$$

Not implementable

3.) $H(s) = \dfrac{6(s+1)^2}{(s+3)(s+5)} = 6 + \dfrac{12}{s+3} - \dfrac{48}{s+5}$ $\quad (p=0)$ ✓

$$\Rightarrow \begin{cases} u(t) = 6e(t) + 12x_1(t) - 48x_2(t) \\ \dot{x}_1(t) = -3x_1(t) + e(t) \\ \dot{x}_2(t) = -5x_2(t) + e(t) \end{cases}$$

Implementable!

- Now that we have a practical constraint on $H(s)$ we can start to consider designing it to meet these and other constraints. Let's consider some simple examples.

- Typical freq domain constraints are to have a mag xover (ω_γ) sufficiently large to ensure a fast settling time, and phase marg. (γ) sufficiently large to ensure small overshoot.

- Most physical systems have phase that approaches or falls below $-180°$ at higher freqs, which means typically $H(s)$ will need to increase the phase in at least a band of higher freqs.

Design Study I:

Suppose $G(s) = \frac{3}{s(s+2)}$, and we <u>want</u> a stable CL system with $\omega_\gamma = 6$, $\gamma = 45°$.

With $H(s) = K$, these constraints are not achievable since $\angle G(6j) < -135°$.

Using above example, we know specs are met if:

$$L(s) = \frac{6^2\sqrt{2}}{s(s+6)}$$ ($\alpha = 6$ in prev. Example)

$\Rightarrow$ Choose

$$H(s) = \left(\frac{6^2\sqrt{2}}{3}\right)\frac{(s+2)}{(s+6)}$$

So $L(s) = G(s)H(s)$ has desired properties

Note: Design here uses <u>stable</u> pole-zero cancellation.

Design Study, II

Suppose instead want $\omega_\gamma = 6$, $\gamma = 60°$. Specs can't be met so easily as above.

<u>Need</u>: $\measuredangle L(j\omega_{Des}) = -120° = \gamma_{Des} - 180°$ ($\gamma_{des} = 60°$ here, and $\omega_{Des} = 6$ here)

But $\measuredangle L(j\omega_{Des}) = \measuredangle G(j\omega_{Des}) + \measuredangle H(j\omega_{Des})$

Hence: $\gamma_{Des} - 180° = \measuredangle G(j\omega_{Des}) + \measuredangle H(j\omega_{Des})$

Or: $\measuredangle H(j\omega_{Des}) = \gamma_{Des} - 180° - \measuredangle G(j\omega_{Des})$

$$= \phi_{req} \quad \text{"phase Deficit"}$$

ϕ_{req} is required phase (typically positive) that compensator must provide at ω_{Des} to meet specs.

For $G(s) = \frac{3}{s(s+2)}$, $\gamma_{Des} = 60°$, $\omega_{Des} = 6$

$\measuredangle G(6j) = -161.6°$

$$\phi_{req} = 60° - 180° - 161.56°$$

$$\Rightarrow \phi_{req} = 41.56°$$

Now suppose we could ideally implement only a LHP zero in $H(s)$

$$\Rightarrow H(s) = K(s - z_c) \quad \text{in ZPK form}$$

(Note we can't do this generally, but it is a convenient hypothetical starting point to illustrate the thought process).

$\Phi_{req} = 41.56^\circ$, $H(s) = K(s - z_c)$ $\quad K, z_c > \emptyset$

Choose K, z_c so that

$$\measuredangle(j\omega_{Des} - z_c) = \Phi_{req}$$

and

$$|L(j\omega_{Des})| = 1$$

<u>Decoupled!</u>

$$\measuredangle(j\omega_{Des} - z_c) = \tan^{-1}\left(\frac{\omega_{Des}}{-z_c}\right) = \tan^{-1}\left(\frac{\omega_{Des}}{|z_c|}\right) \text{ since } z_c < \emptyset$$

$$\Rightarrow \frac{\omega_{Des}}{|z_c|} = \tan\Phi_{req} \quad \text{or} \quad z_c = -\left[\frac{\omega_{Des}}{\tan\Phi_{req}}\right]$$

Here $z_c = -\left[\dfrac{6}{\tan 41.56^\circ}\right] = -6.77$

So now $H(s) = K(s + 6.77)$. Find K

Finding K

Let $L_o(s) = L(s)\Big]_{K=1} \Rightarrow L(s) = K L_o(s)$

Then choose

$$K = \frac{1}{|L_o(j\omega_{des})|}$$

Since then

$$|L(j\omega_{Des})| = |K L_o(j\omega_{Des})| = |K|\,|L_o(j\omega_{Des})|$$

$$= 1$$

i.e. ω_{Des} is mag xover freq. for $L(j\omega)$, as desired

Here: $$L_o(s) = \frac{3(s+6.77)}{s(s+2)}$$

$$|L_o(6j)| = 0.715 \Rightarrow K \cong 1.4$$

$$H(s) = 1.4(s+6.77)$$

$$U(s) = H(s)E(s)$$

$$= [1.4s + 1.4 \times 6.77]\,E(s)$$

$$\Rightarrow \quad u(t) = \underbrace{1.4}_{K_D}\,\dot{e}(t) + \underbrace{1.4 \times 6.77}_{K_P}\,e(t)$$

"PD" controller

Above is not generally implementable $\quad p(H) < \phi$

$\Rightarrow$ $H(s)$ must contain at least 1 (LHP) pole to balance the zero (make $p(H) \geq \phi$)

$\Rightarrow$ LHP poles contribute negative phase, hence work against our objectives

One strategy (not necessarily the best, but easy to do): put pole of $H(s)$ so that its impact on phase of $L(j\omega)$ is negligible at least near desired crossover ω_{Des}

i.e. $H(s) = K\left[\frac{(s-z_c)}{(s-p_c)}\right]$

with $|p_c| \geq 10\,\omega_{Des}$

(Recall p_c will change phase starting for $\omega \gtrsim \frac{|p_c|}{10}$; want this above ω_{Des}).

If $P_c = -10\,\omega_{Des}$, then

$$\measuredangle(j\omega_{Des} - P_c) = \measuredangle(j\omega_{Des} + 10\,\omega_{Des})$$

$$= \tan^{-1}\left(\frac{1}{10}\right) \approx 5.7^\circ$$

and $\measuredangle H(j\omega_{Des}) = \measuredangle(j\omega_{Des} - Z_c) - \measuredangle(j\omega_{Des} - P_c)$

$$= \measuredangle(j\omega_{Des} - Z_c) - 5.7^\circ$$

Still need: $\measuredangle H(j\omega_{Des}) = \Phi_{req}$

So choose:

$$\measuredangle(j\omega_{Des} - Z_c) = \boxed{\Phi_{req} + 5.7^\circ}$$

Then choose K as before.

For our example we need

$\measuredangle(6j - z_c) = 41.56° + 5.7° = 47.26°$

$\Rightarrow \quad z_c = -5.54$

$$\Rightarrow L_o(s) = \frac{3(s+5.54)}{s(s+2)(s+60)}$$

$\Rightarrow$ $K = 93.37$

$$\boxed{H(s) = \frac{93.37(s+5.54)}{(s+60)}}$$

$\dot{x}_1 = -60x_1 + e$

$u = \underline{K_1} e + \underline{K_2} x_1$

Note big increase in K! Generally associated with bigger u(t). Must check for saturation!

Ideal (pure zero) result obtained As $p_c \to -\infty$ (pole very far into LHP), but is associated with very large control inputs. Can do some simple z_c, p_c optimization to moderate control magnitude

Simple optimization of required location for p_c

We have seen a simple strategy for choosing required pole in $H(s)$ is to make $p_c < -10\,\omega_{Des}$

$\Rightarrow$ Ensures p_c subtracts no more than $5.7°$ from $\measuredangle H(j\omega_{Des})$, easy to adjust location of z_c to "make up" this phase loss to maintain $\measuredangle H(j\omega_{Des}) = \phi_{req}$.

However, such a strategy often results in undesireably large $u(t)$.

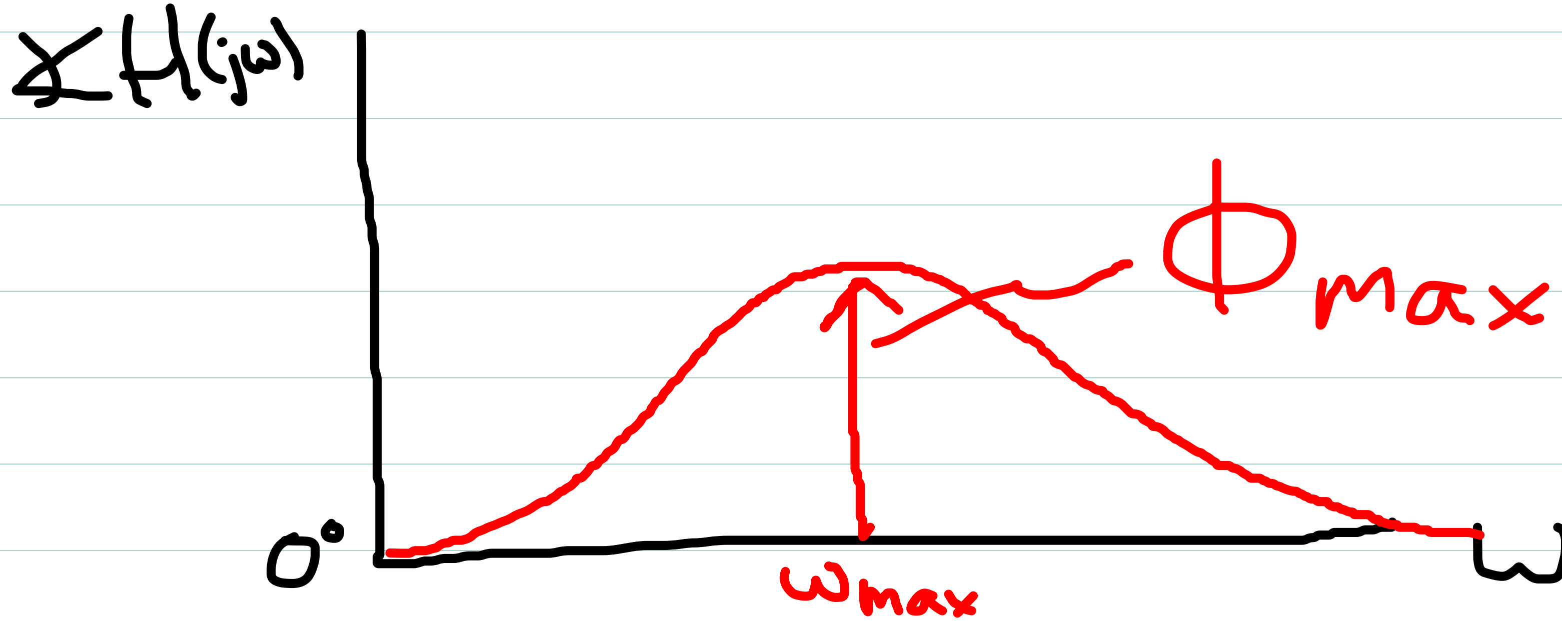
Try to balance the competing requirements by finding minimum possible ratio p_c/z_c which still provides $\measuredangle H(j\omega_{Des}) = \phi_{req}$.

Compensators of general form:

$$H(s) = K\left[\frac{\beta\tau s+1}{\tau s+1}\right] \quad \text{with} \quad \begin{cases} \tau > 0 \\ \beta > 1 \end{cases}$$

$\Rightarrow$ $z_c = -1/\beta\tau$, $p_c = -1/\tau$ so $|z_c| < |p_c|$ (zero closer to imag axis) and $\beta = p_c/z_c$ ("lead ratio")

Are called "lead compensators", since $\measuredangle H(j\omega) > 0$ for all $\omega > 0$ (positive phase is called "lead").



Note here that:

$$U(s) = H(s)E(s) = K\left[\frac{\beta T s+1}{Ts+1}\right]E(s)$$
$$= K\left[\beta - \frac{\beta-1}{Ts+1}\right]E(s)$$

$$\boxed{\begin{aligned} &\Rightarrow u(t) = K\beta e(t) + K(1-\beta)x_1(t) \\ &\quad \tau \dot{x}_1(t) = -x_1(t) + e(t) \end{aligned}}$$

Corresponding implementation equations

So that $|u(t)| \propto \beta$ generally.

Want to find smallest value of $\beta = {}^{p_c}/_{z_c}$ which ensures $\phi_{max} \geq \phi_{req}$.

We can compute:

$$\phi_{max} = \sin^{-1}\left[\frac{\beta-1}{\beta+1}\right]$$

Which is an <u>increasing</u> function of $\beta > 1$

Thus, the minimum required β value is obtained when

$$\phi_{max} = \phi_{req}.$$

Selection of τ can then be achieved using analytical result

$$\omega_{max} = \frac{1}{\tau\sqrt{\beta}}$$

and setting $\omega_{max} = \omega_{Des}$.

Revisit previous example:

$$G(s) = \frac{3}{s(s+2)} \quad \text{want } \gamma_{Des} = 60°, \ \omega_{Des} = 6$$

for which $\phi_{req} = 41.56°$

Previous design used $H(s) = 93.4\left[\dfrac{s+5.54}{s+60}\right]$ $(\beta = 10.8)$

New design:

$$41.56^\circ = \Phi_{req} = \Phi_{max} = \sin^{-1}\left[\frac{\beta-1}{\beta+1}\right]$$

$$\Longrightarrow \beta = 4.94$$

Then $\quad 6 = \omega_{Des} = \omega_{max} = \dfrac{1}{\tau\sqrt{\beta}} \quad \Rightarrow \tau = \dfrac{1}{\omega_{Des}\sqrt{\beta}}$

So $\tau = 0.075$ here.

and thus $\quad H(s) = K\left[\dfrac{0.375s+1}{0.075s+1}\right] \quad \begin{pmatrix} z_c = -2.67 \\ p_c = -13.3 \end{pmatrix}$

Choose K as before: $\quad$ Let $L_0(s) = [L(s)]_{K=1}$

Then take $\quad K = \dfrac{1}{|L_0(j\omega_{Des})|} = 5.63$ here

Comparison of Designs:

For a unit step $y_d(t)$:

Old design: $t_s \approx 0.96$ sec, $M_p \approx 15\%$, $u_{max} \approx 92$ "far pole"

New design: $t_s \approx 0.86$ sec, $M_p \approx 10\%$, $u_{max} \approx 27$ "lead"

New design is essentially the same (a little better) in transient performance, and requires a factor of 3 less maximum control effort.

$\Rightarrow$ minimizing β is very beneficial!

Another Example

Suppose $G(s) = \dfrac{3}{s(s-2)}$

and we again want $\gamma_{Des} = 45°$ with $\omega_{Des} = 6$, which we know is assured if:

$$L(s) = \frac{6^2\sqrt{2}}{s(s+6)}$$

You might be tempted to try $H(s) = \left(\dfrac{6^2\sqrt{2}}{3}\right)\left[\dfrac{s-2}{s+6}\right]$

Don't do it! Pole-zero cancellation cannot be guaranteed to be exact here, and CL system <u>will</u> have an unstable pole (try it! Use $(s-2.1)$ in numerator and see what happens).

Instead, meet targets using only LHP poles/zeros.

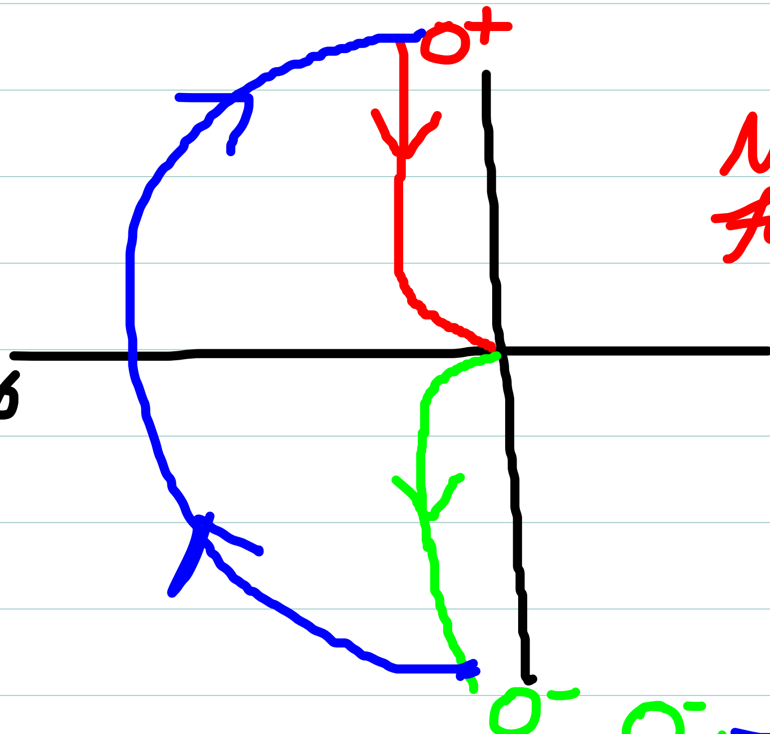
Check Nyquist: Ensure $\gamma > 0$ still stabilizing

Using $H(s) = K > 0$

$P_R(L) = 1$

$N_{cw}(L) = 1$ for any $K > 0$

$\Rightarrow T(s)$ unstable

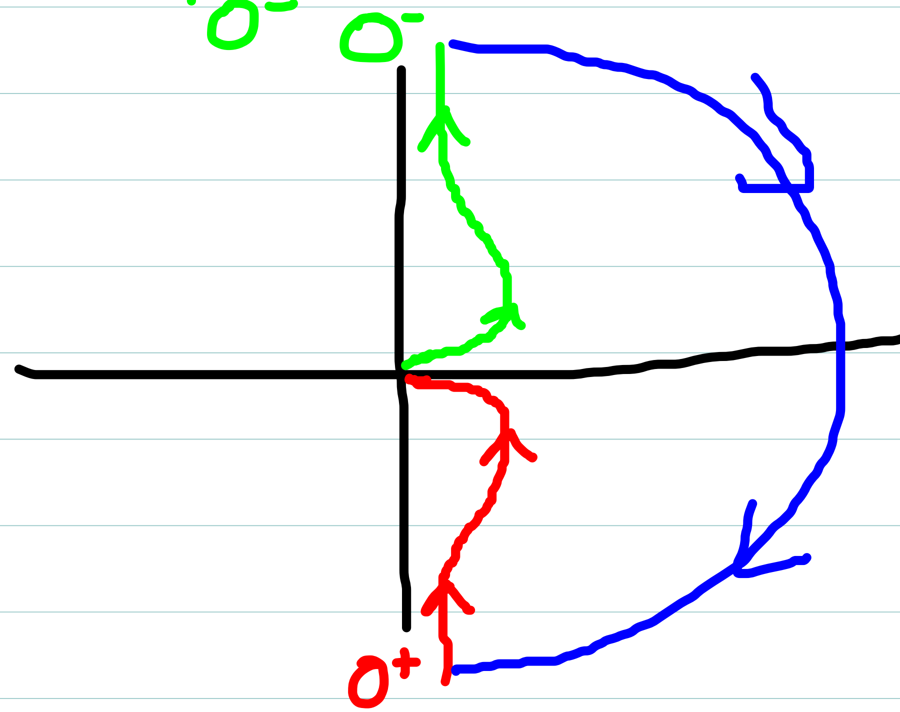


Need $N_{cw}(L) = -1$ for stability!

Using $H(s) = K < 0$

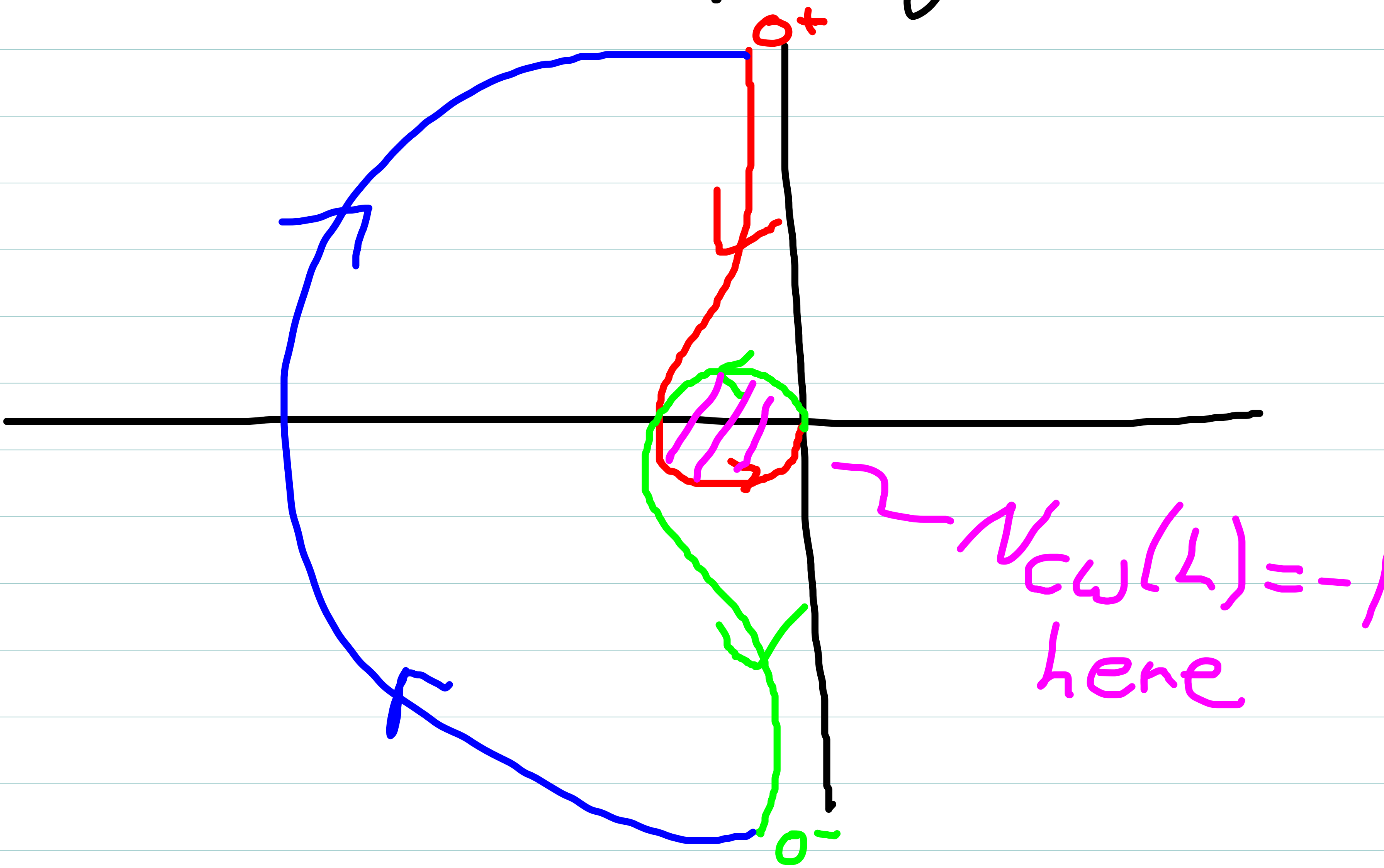
$N_{cw}(L) = 0$ for any $K < 0$

$T(s)$ unstable

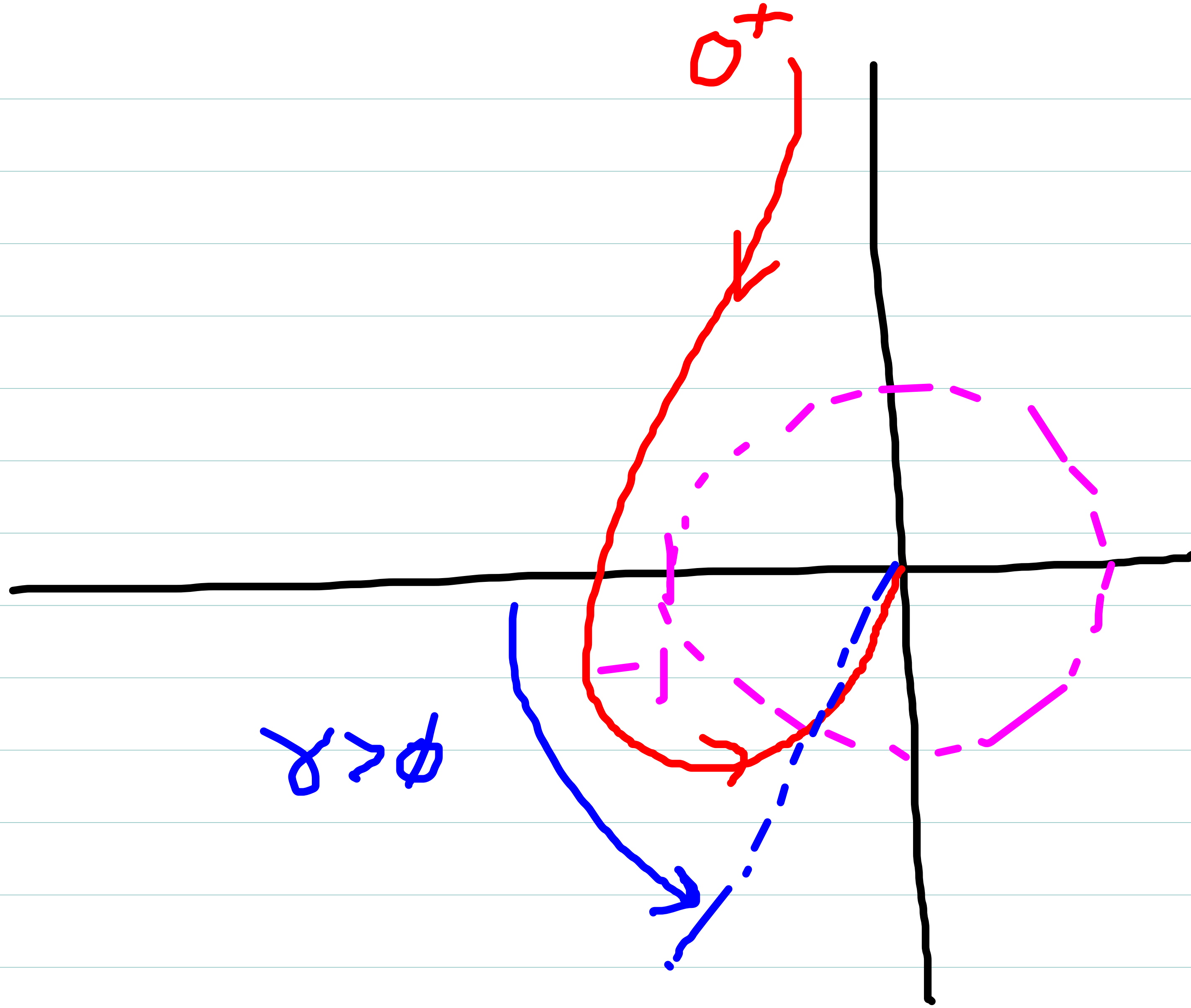


Conclusion: Cannot stabilize with $H(s) = K$

If $H(s) = K(\tau s + 1)$, $K > \phi$, $\tau > \phi$
(LHP zero, positive gain)



Will stabilize if -1 in shaded area!



With -1 in desired region, phase margin will be positive $\Rightarrow$ agrees with our basic design guideline positive γ <u>is</u> stabilizing here!

For $G(s) = \dfrac{3}{s(s-2)}$, $\gamma_{des} = 45^\circ$, $\omega_{des} = 6$:

$$\Phi_{req} = 45 - 180 - \measuredangle G(6j)$$

$$= 63.43^\circ$$

$$\Rightarrow z_c = -\left[\frac{\omega_{des}}{\tan \Phi_{req}}\right] = -3$$

$$\Rightarrow K = \frac{1}{|L_0(6j)|} = 1.89. \text{ So } H(s) = 1.89(s+3) \text{ works}$$

For an implementable design, simple approach is again to put pole at $-10\omega_{des}$, increase Φ_{req} by 5.7°, giving

$$H(s) = 118.79\left[\frac{(s+2.29)}{(s+60)}\right]$$

(Or, we could do a lead comp design, if the above requires excessively large $u(t)$.)

$$G(s) = \frac{3}{s(s-2)}$$

Suppose we want instead $\gamma_{Des} = 70°$ at $\omega_{Des} = 6$

$\phi_{req} = 70 - 180 - \measuredangle G(j\ ) = 88.43°$, Add $+5.7°$ for pole, need
$\measuredangle(j\omega_{Des} - z_c) = 94.13°$

This condition cannot be satisfied with a single LHP zero
Need $\underline{\underline{2}}$ zeros here.

With 2 zeros, can add up to $+180°$ to $\measuredangle L(j\omega)$. Lots of choices for 2 zeros adding up to $94.13°$ at $\omega = 6$. Can simplify design if we assume zeros/poles repeated

$$H(s) = K\left[\frac{(s-z_c)^2}{(s-p_c)^2}\right]$$

So $\measuredangle H(j\omega) = 2\measuredangle(j\omega - z_c) - 2\measuredangle(j\omega - p_c)$

$$\Rightarrow \measuredangle(j\omega_{Des} - z_c) = \frac{\phi_{req}}{2} + \measuredangle(j\omega_{Des} - p_c)$$

for example, using $p_c = -10\omega_{Des} = -60$ again

We get $z_c = -5.05$, $K = 747.89$

and $H(s) = 747.89 \left[\dfrac{(s+5.05)^2}{(s+60)^2} \right]$

(Again, could instead do a lead comp design - See next page).

Note: Above considerations apply any time we would need $\measuredangle H(j\omega_{Des}) > 90°$. Not specific to this example.

We can apply similar thinking for more complicated situations: i.e. if we would need

$$\measuredangle H(j\omega_{Des}) > 180°$$

Then we need $\underline{3}$ LHP zeros in $H(s)$, etc.

Alternate lead comp design for above

Using again $\gamma_{des} = 70°$, $\omega_{des} = 6$

Set $\Phi_{max} = \frac{\Phi_{req}}{2} \Longrightarrow \beta = 5.61$

$\omega_{max} = 6 \Longrightarrow \tau = .07$

Then:

$$H(s) = K\left[\frac{(.395s+1)^2}{(.07s+1)^2}\right]$$

and

$$K = \frac{1}{|L_o(6j)|} = 2.255$$

Comparison of different designs

$\boxed{\gamma = 45°}$:

zero/far pole: $t_s = 1.95$ sec, $M_p = 51\%$, $u_{max} = 119$

reduced

Lead: $t_s = 1.4$ sec, $M_p = 53\%$, $u_{max} = 52$

$\boxed{\gamma = 70°}$

(less overshoot)

zero/far pole: $t_s = 3.0$ sec, $M_p = 40\%$, $u_{max} = 700$

much less

Lead: $t_s = 3.9$ sec, $M_p = 47\%$, $u_{max} = 71$

Which is better...? That becomes a judgement call
Also need to consider tracking, bandwidth, robustness, and how "tight" your requirements actually are...

$$\left[\; G(s) = \frac{3}{s(s-2)}, \quad \omega_{Des} = 6 \;\right]$$

(Intro:) Effect of Uncertainty

Suppose again we had wanted $\gamma = 45°$, $\omega_\gamma = 6$ and we (naively) ignored my caveat and chose

$$H(s) = \left(\frac{6^2\sqrt{2}}{3}\right)\left(\frac{s-2}{s+6}\right)$$

If the G(s) really is:

$$G(s) = \frac{3}{s(s-1.9)}$$

[5% uncertainty in unstable pole location]

we'd have CL poles at: $-3.04 \pm 6.5j$, 1.97 unstable!

But...

With the two designs above, we'd instead have CL poles:

Zero/far pole: $-2.1 \pm 3.3j$, -53.9
Lead: $-2.8 \pm 2.3j$, -18

Both still stable!

Now we have an idea of the constraints on $L(s)$ for closed-loop stability and transient performance

$\Rightarrow$ Make $L(j\omega)$ have large positive phase margin γ and large crossover freq ω_γ;
(but check Nyquist in unusual or unfamiliar cases)

Let's examine constraints on $L(s)$ which ensure good <u>tracking</u>, i.e. which ensure $|e_{ss}(t)|$ is small for a variety of $y_d(t)$.

Recall that $e(t)$ for a given $y_d(t)$ is governed by sensitivity transfer function $S(s)$ where

$$E(s) = S(s)Y_d(s) \qquad \text{with} \qquad S(s) = \frac{1}{1+L(s)}$$

Intuitively, we make $e(t)$ small by making $L(s)$ "big"

Simple Relationships

Already seen:

$\Rightarrow$ $e_{ss}(t) = 0$ when $y_d(t) = A$ (constant)

if $S(0) = 0$

$\Rightarrow$ $|e_{ss}(t)| \leq 0.7A$ ($\leq 70\%$ error)

if $y_d(t) = A\cos(\omega t + \psi)$

for any ω such that $|S(j\omega)| \leq -3\text{dB}$

And we call the range of such ω the "tracking bandwidth" ω_B of the system.

More general observations

$$|S(j\omega)| = \left|\frac{1}{1+L(j\omega)}\right| = \frac{1}{|1+L(j\omega)|}$$

All physical systems with implementable controllers satisfy:

$$\boxed{|L(j\omega)| \to \emptyset \text{ as } \omega \to \infty}$$

i.e. $L(s)$ has relative degree of 1 or more (at least one more pole than zeros).

Since $H(s)$ is constrained to have relative degree zero or greater, and all physical systems have $G(s)$ with relative degree 1 or greater. Thus $L(s) = G(s)\,H(s)$ has relative degree at least 1

Implication: $\boxed{|S(j\omega)| \to 1 \ (\emptyset \text{ dB}) \text{ as } \omega \to \infty}$

$|S(j\omega)|_{dB} \to \emptyset$ as $\omega \to \infty$

Thus there is always an <u>upper</u> bound on bandwidth.

Let's see if we can more precisely characterize this bound in terms of properties of $L(s)$.

Looking at lower freqs: $\omega \to \emptyset$

$$S(\emptyset) = \frac{1}{1+L(\emptyset)}$$

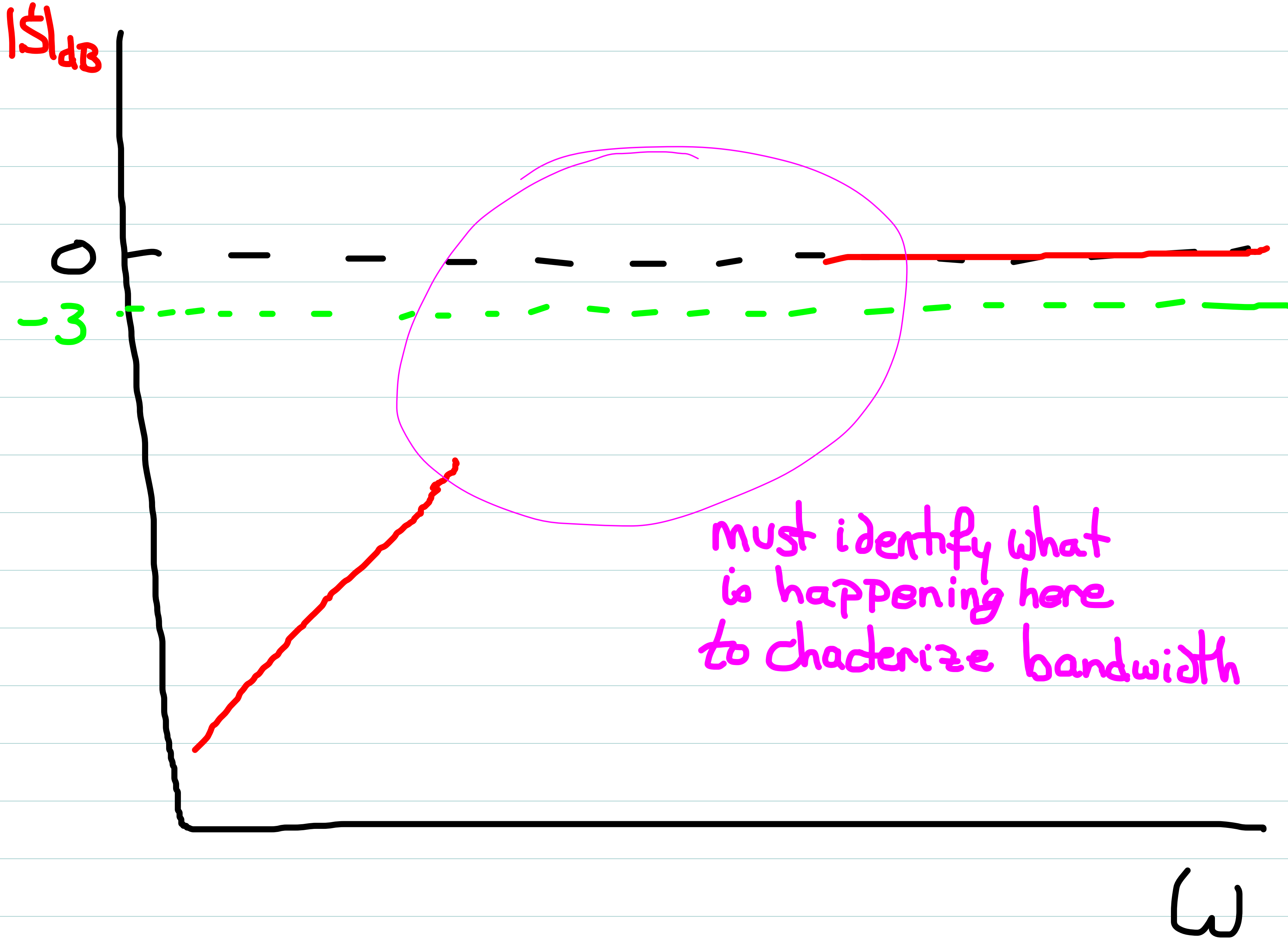
$S(\emptyset)=\emptyset \Rightarrow L(\emptyset)=\infty \Rightarrow L(s)\Big]_{s=\emptyset} = \infty \Rightarrow$ $\boxed{L(s) \text{ has pole at origin}}$

$\Downarrow$

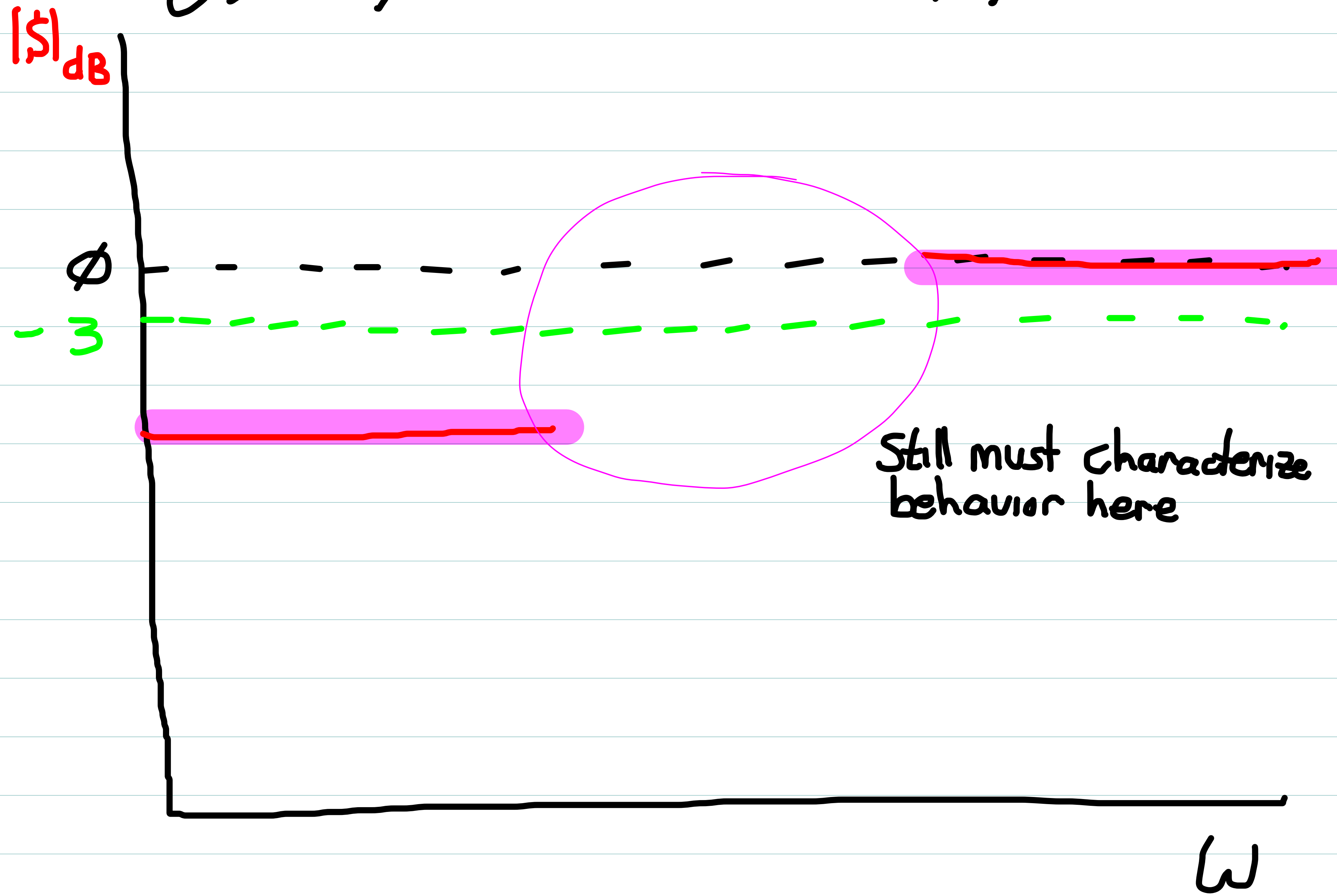
$\boxed{S(s) \text{ has } \underline{\text{zero}} \text{ at origin}}$

Remember this correlation! We will see it again!

$\Rightarrow$ low freq slope of $|S|$ is <u>positive</u>


|S|dB
0
-3
must identify what
is happening here
to chacterize bandwidth
ω

If $L(0) \neq \infty$, then $|S(0)|$ is constant
and mag plot of $|S(j\omega)|$ has zero low freq slope



Bandwidth is region for which $|S(j\omega)| \leq -3\text{dB}$

in actual units $|S(j\omega)| \leq \frac{1}{\sqrt{2}}$

And hence is the region for which $|1+L(j\omega)| \geq \sqrt{2}$

Want to identify constraints on $L(j\omega)$ which guarantee this.

If $|L(j\omega)| > 1$ (ϕdB), then it is true that

$$|1+L(j\omega)| \geq |L(j\omega)| - 1$$

Hence, if $|L(j\omega)| \geq 1+\sqrt{2}$ (~ 7.7 dB), then

$$|1+L(j\omega)| \geq \sqrt{2} \quad \text{and} \quad |S(j\omega)| \leq -3\text{dB}$$

So, tracking bandwidth is guaranteed to be <u>at least</u> the range of ω for which $|L(j\omega)| \gtrsim 7.7$ dB

This is pretty close to ω_γ ($|L(j\omega_\gamma)| = \emptyset$ dB)
Let's see if we can more precisely relate ω_B to ω_γ:

Assume that $|L(j\omega)|$ is decreasing with slope at least -20dB/dec from +7.7 dB through $\emptyset$ dB (typical, but not always).

Then $|L(j\omega)| \geq 7.7$ dB starting at frequencies (7.7/20) of a decade <u>below</u> ω_γ

i.e. for $\omega \leq \left(10^{-7.7/20}\right)\omega_\gamma \approx \omega_\gamma/2.5$

$\omega_B \cong \omega_\gamma/2.5$

Now, let's look more precisely at what is happening at ω_γ

$$|S(j\omega_\gamma)| = \frac{1}{|1+L(j\omega_\gamma)|}$$

$|1+L(j\omega_\gamma)|$ depends on phase $\measuredangle L(j\omega_\gamma)$ and hence on phase margin γ:

Since $|L(j\omega_\gamma)| = 1$ by definition:

$$L(j\omega_\gamma) = e^{j\phi} \quad \text{where } \phi = \measuredangle L(j\omega_\gamma)$$

By definition of $\gamma = 180 + \measuredangle L(j\omega_\gamma)$, $\phi = \gamma - 180°$

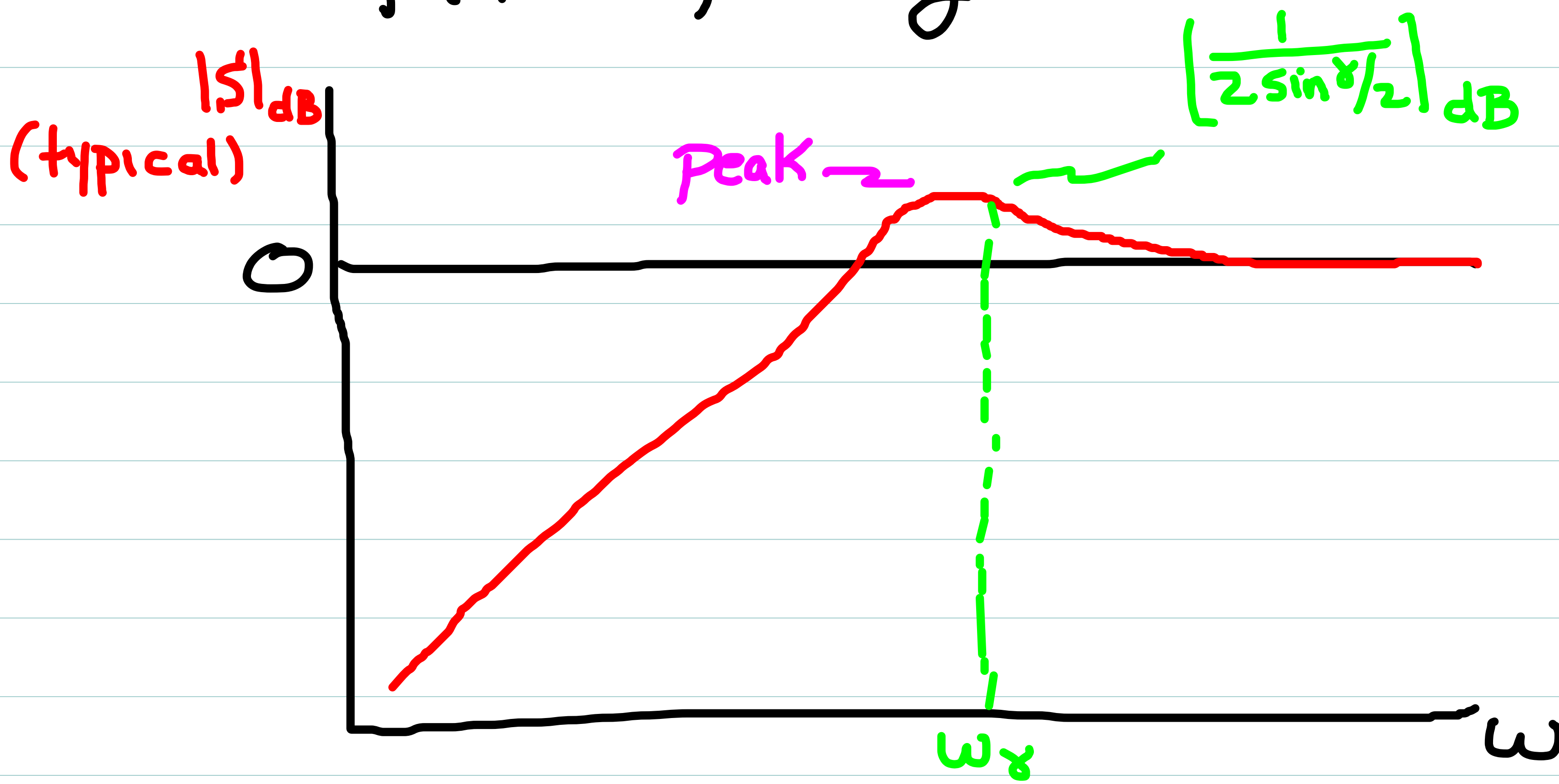
So $1 + L(j\omega_\gamma) = 1 + e^{j(\gamma-\pi)} = (1+\cos(\gamma-\pi)) + j\sin(\gamma-\pi)$

and $|1+L(j\omega_\gamma)| = 2\sin(\gamma/2)$

Hence:

$$\boxed{|S(j\omega_\gamma)| = \frac{1}{2\sin\gamma/2}}$$

Note $|S(j\omega_\gamma)| > 1$ when $\gamma < 60°$, thus generally $|S(j\omega)|$ will exhibit a peak of height at least as tall as $|S(j\omega_\gamma)|$ (may be higher)



$$|S(j\omega_\gamma)| = \frac{1}{2\sin\gamma/2}$$

Note if $\gamma = 90°$ $|S(j\omega_\gamma)| = \frac{1}{\sqrt{2}}$ $(= -3\,dB)$

Together with previous observations we can conclude that typically for a feedback system with $0 \le \gamma \le 90°$

$$\boxed{\frac{\omega_\gamma}{2.5} \le \omega_B \le \omega_\gamma}$$

And in particular increasing ω_γ increases tracking bandwidth ω_B

Thus in this sense, our design guidelines for performance are aligned with the design guidelines for good tracking

$\Rightarrow$ Larger ω_B means a greater range of sinusoidal $y_d(t)$ which can be tracked with minimal error.

But, this isn't the whole Story!

Many times we require our designs to have $|e_{ss}(t)| = \phi$ ("perfect tracking") for specified classes of $y_d(t)$ (even sinusoidal)

When can this be guaranteed?

Let $L(s) = \frac{N(s)}{D(s)}$ $N(s), D(s)$ Polynomials

Then $S(s) = \frac{1}{1+L(s)} = \frac{D(s)}{N(s)+D(s)}$

$\Rightarrow$ Zeros of $S(s)$ are poles of $L(s)$

In particular, perfect tracking of step $y_d(t)$ requires $S(\phi) = \phi \Rightarrow D(\phi) = \phi \Rightarrow L(s)$ has at least 1 pole at origin, as we have seen.

More generally, suppose

$$Y_d(s) = \frac{a(s)}{b(s)} \qquad a(s), b(s) \text{ polynomials}$$

Then $E(s) = S(s) Y_d(s)$

$$= \left[\frac{D(s)}{N(s)+D(s)}\right]\left[\frac{a(s)}{b(s)}\right]$$

Now, assuming our controller at least stabilizes the feedback loop, the poles of $S(s)$ [same as poles of $T(s)$] are stable

If all poles of $Y_d(s)$ (roots of $b(s)$) are stable, then partial fraction expansion and inverse transform of $E(s)$ will give $e(t)$ as a sum of decaying exponential functions.

$\Rightarrow e_{ss}(t) = \emptyset$ here

Above result makes sense:

For a stable system, $y(t)$ naturally "wants" to converge to ϕ. If $Y_d(s)$ has all stable poles, then $y_d(t)$ is a sum of decaying exponentials and $y_d(t) \to \phi$

So asymptotically, we are requiring the system to do what it already wants to do, and thus we get perfect steady-state tracking.

More interesting is when $y_d(t) \not\to \phi$ As $t \to \infty$. So suppose that poles of $Y_d(s)$ are not stable.

$$E(s) = \left[\frac{D(s)}{N(s)+D(s)}\right]\left[\frac{a(s)}{b(s)}\right]$$

and $e(t)$ will contain same non-stable poles as $Y_d(s)$, unless...

$$E(s) = \left[\frac{D(s)}{N(s)+D(s)}\right]\left[\frac{a(s)}{b(s)}\right]$$

Unless, the non-stable poles of $Y_d(s)$ are cancelled by zeros of $S(s)$

i.e. if $D(s) = D'(s)\,b(s)$, $D'(s)$ polynomial

For $y_d(t) = A \Rightarrow Y_d(s) = \frac{A}{s}$ ($b(s) = s$, root at origin)

Need $D(s) = s D'(s)$ i.e. $D(s)$ also has root at origin

$\Rightarrow L(s)$ has pole at origin (as we have already seen)

But the above result is much more general!

Suppose $y_d(t) = At \Rightarrow Y_d(s) = \frac{A}{s^2} \Rightarrow b(s) = s^2$

If $L(s)$ has 2 poles at origin $D(s) = s^2 D'(s)$, Non-stable terms will cancel.

General Result (tracking)

If $L(s)$ has the same non-stable poles as $Y_d(s)$ then $e_{ss}(t) = \emptyset$

If true, we say that $L(s)$ has an "internal model" of $y_d(t)$, and the above fact is known as the "internal model principle" (IMP)

Note: while theoretically this applies even if $Y_d(s)$ has unstable poles, practically we use this only for marginally stable poles of Y_d, i.e. poles on imag axis.

positive integer

One common special case: "type p" (polynomial) $y_d(t)$, i.e.

(p integer $\geq \emptyset$) $y_d(t) = \left(\frac{A_p}{p!}\right) t^p$ A_p constant

$\Rightarrow Y_d(s) = \frac{A_p}{s^{p+1}}$

p = power of t in $y_d(t)$ $\Leftrightarrow$ $p+1$ poles at $\emptyset$

Imp Examples:

1.) $y_d(t) = \text{const} \Rightarrow Y_d(s)$ has pole at $\emptyset$

$\Rightarrow e_{ss}(t) = \emptyset$ if $L(s)$ has pole at origin

2.) $\boxed{y_d(t) = A\cos(\omega t + \varphi) \Rightarrow Y_d(s) \text{ has poles at } \pm j\omega}$

$\Rightarrow e_{ss}(t) = \emptyset$ if $L(s)$ has poles at $\pm j\omega$

(i.e. denom of $L(s)$ has factor of $\underline{(s^2+\omega^2)}$

(Note that this result is true regardless of A or φ!)

Type p: $Y_d(s) = \frac{A_p}{s^{p+1}}$

Via IMP: perfect tracking ($e_{ss} = \emptyset$) requires $L(s)$ to have $p+1$ poles at origin

$p = \emptyset$, $y_d(t) = A_0$ (constant) $\Longrightarrow$ $L(s)$ needs 1 pole at origin

$p = 1$, $y_d(t) = A_1 t$ (linear) $\Longrightarrow$ $L(s)$ needs 2 poles at origin

and so on.

Now, suppose $L(s)$ does not have enough poles at origin what happens? Look more closely at

$$E(s) = S(s)Y_d(s) = \left[\frac{D(s)}{D(s)+N(s)}\right]\left(\frac{A_p}{s^{p+1}}\right)$$

when $Y_d(t)$ is type p.

$$E(s) = \left[\frac{D(s)}{D(s)+N(s)}\right]\frac{A_p}{s^{p+1}}$$

Pull out any poles $L(s)$ has at origin: Let

$$D(s) = s^N D'(s) \quad \left(N = \text{\# poles of } L(s) \text{ at origin} \right.$$
$$\left. \text{"type" of } L(s)\right)$$

So $E(s) = A_p\left[\frac{s^N}{s^{p+1}}\right]\left[\frac{D'(s)}{N(s)+D(s)}\right]$

If $N \geq p+1$, $E(s)$ will have only stable poles remaining and $e(t) \to \emptyset \Rightarrow e_{ss}(t) = \emptyset$

If $N = p$, however, ($L(s)$ has one less pole at origin than $Y_d(s)$) then

$$E(s) = \left(\frac{A_p}{s}\right)\left[\frac{D'(s)}{D(s)+N(s)}\right] = \frac{C_0}{s} + \underbrace{\frac{C_1}{s-d_1} + \cdots}_{\text{from stable poles of } S(s) \text{ (and } T(s))}$$

So $e_{ss}(t) = C_0$ constant here

We can compute c_0 in this case using residue formula:

$$c_0 = A_p \left[\frac{D'(s)}{D(s)+N(s)} \right]_{s=\phi}$$

But recall $D'(s) = {}^{D(s)}/s^p$ (since $p=N$ here)

So $c_0 = A_p \left[\frac{D(s)}{s^p D(s) + s^p N(s)} \right]_{s=\phi} = \left[\frac{A_p}{s^p + s^p L(s)} \right]_{s=\phi}$

But note (again since $N=p$ here)

$$\left[s^p L(s) \right]_{s=\phi} = K_{B,L} \quad \text{Bode gain of } L(s)$$

So:

$$c_0 = \left[\frac{A_p}{s^p + K_{B,L}} \right]_{s=\phi} = \begin{cases} \dfrac{A_p}{1+K_{B,L}} & p=\phi \\ \dfrac{A_p}{K_{B,L}} & p>\phi \end{cases}$$

$\Rightarrow$ $e_{ss}(t)$ inversely prop. to Bode gain of L

Now suppose $N = p-1$ (1 less poles at origin in $L(s)$)

Then $E(s) = A_p \left(\frac{s^N}{s^{p+1}}\right)\left[\frac{D'(s)}{D(s)+N(s)}\right] = \frac{A_p}{s^2}\left[\frac{D'(s)}{D(s)+N(s)}\right]$

$$= \frac{c_0}{s} + \frac{c_1}{s^2} + \underbrace{\frac{c_2}{(s-d_1)} + \cdots}_{\text{from stable poles of } S(s)}$$

So $e_{ss}(t) = c_0 + c_1 t \longrightarrow \infty$ as $t \to \infty$

Diverges

Easy to show similar phenomenon for any $N < p$.

i.e. if $N = p-2$ then

$$e_{ss}(t) = c_0 + c_1 t + c_2 t^2 \longrightarrow \infty$$

etc.

Summary: Tracking for "type-p" $y_d(t)$

$y_d(t) = \left(\frac{A_p}{p!}\right)t^p$, N = # poles at origin in $L(s)$.

$$\boxed{e_{ss}(t) = \begin{cases} \emptyset & N > p \\ c_0 \neq \emptyset & N = p \\ \infty & N < p \end{cases}}$$

Where

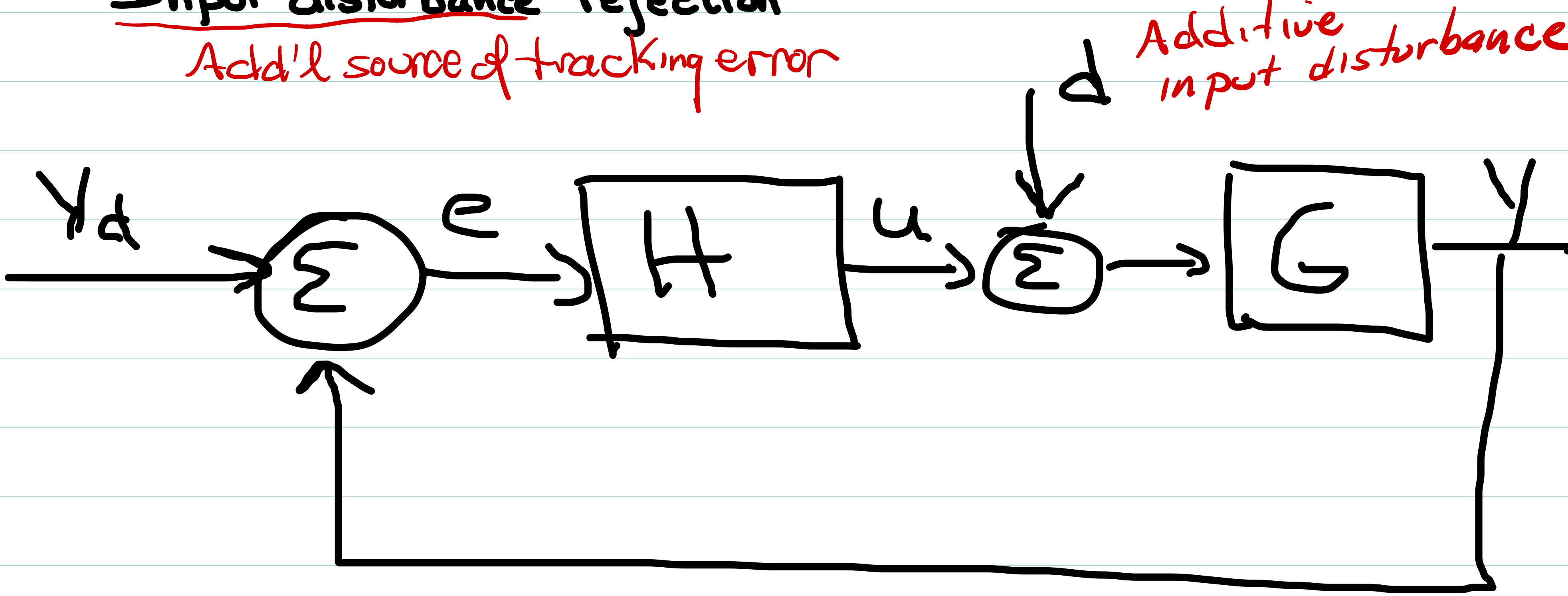
$$\boxed{c_0 = \begin{cases} \dfrac{A_p}{1 + K_{B,L}} & p = \emptyset \\ \dfrac{A_p}{K_{B,L}} & p > \emptyset \end{cases}}$$

and $K_{B,L}$ is the Bode gain of $L(s)$.

Very important design constraint!

Input disturbance rejection

Add'l source of tracking error



d(t) external "disturbance" input to the system: Not under our direct control, and cannot be predicted or measured during operation of the system.

What affect will this have on stability or accuracy?

"rejection": ability to maintain $e_{ss}(t)$ small even when $d(t) \neq \emptyset$.

Re-derive feedback loop equations:

$$Y(s) = G(s)\left[U(s) + D(s)\right] = G(s)\left[H(s)E(s) + D(s)\right]$$

$$= G(s)H(s)E(s) + G(s)D(s)$$

$$= G(s)H(s)\left[Y_d(s) - Y(s)\right] + G(s)D(s)$$

So $\left(1+L(s)\right)Y(s) = L(s)\,Y_d(s) + G(s)D(s)$

or $$Y(s) = \underbrace{\left[\frac{L(s)}{1+L(s)}\right]}_{T(s)} Y_d(s) + \underbrace{\underbrace{\left[\frac{G(s)}{1+L(s)}\right]}_{S_i(s)} D(s)}_{\text{Added term due to disturbance}}$$

$S_i(s)$ "input sensitivity" T.F.

Note: poles of $S_i(s)$ same as $T(s) \Rightarrow S_i(s)$ is stable if $T(s)$ is.

$$\boxed{\Rightarrow \text{Disturbance cannot destabilize system!}}$$

Disturbance can, however, worsen tracking:

$$Y(s) = T(s)Y_d(s) + S_i(s)D(s)$$

$$E(s) = Y_d(s) - Y(s) = \underbrace{(1-T(s))}_{S(s)}Y_d(s) - S_i(s)D(s)$$

So:

$$E(s) = \boxed{S(s)Y_d(s)} - S_i(s)D(s)$$

ideal tracking error

extra error due to disturbance

Want to quantify the added errors due to disturbance

Can analyze similarly to above, but need a bit more care:

$$S_i(s) = \frac{G(s)}{1+L(s)}$$

Let $G(s) = \dfrac{N_G(s)}{D_G(s)}$, $H(s) = \dfrac{N_H(s)}{D_H(s)}$ so $L(s) = \dfrac{N_G(s)N_H(s)}{D_G(s)D_H(s)}$

Dist. rejection

Want $|S_i(j\omega)| \ll 1$ for ω in freq range of $d(t)$

i.e. if $d(t)$ has sig. freq content in $[\omega_1, \omega_2]$

want $|S_i(j\omega)| \ll 1$ for $\omega \in [\omega_1, \omega_2]$

Note $S_i(s) = \dfrac{G(s)}{1+L(s)} = \dfrac{1}{G^{-1}(s) + H(s)}$

So $|S_i(j\omega)| \ll 1 \Rightarrow$ either

- $|G(j\omega)| \ll 1$, or
- $|H(j\omega)| \gg 1$ $\Leftarrow$ Can design for this!

Note: $|G(j\omega)| \to 0$ as $\omega \to \infty$ for physical systems

$\Rightarrow |S_i(j\omega)| \to 0$ as $\omega \to \infty$

But usually $|G(j\omega)| \gtrsim 1$ for mid/low freq.

Dist rejection, cont

$|S_c(j\omega)| \to 0$ As $\omega \to \infty$

<u>But</u> freq. band where dist. is signficant $[\omega_1, \omega_2]$ is typically at mid-low freqs where $|G(j\omega)| \approx 1$

$\Rightarrow$ Need $|H(j\omega)| >> 1$ at these <u>freqs</u>!

Thus, good dist. rejection typically Requires

$|H(j\omega)| >> 1$ for $\omega \in \underbrace{[\omega_1, \omega_2]}_{\text{freqs. where dist is significant}}$

Note: reqt on $H(s)$ <u>only</u>!

As with tracking error and $S(s)$, the IMP provides Add'l insights.

Then:

$$S_i(s) = \frac{G(s)}{1+L(s)} = \frac{N_G(s)D_H(s)}{D_G(s)D_H(s) + N_G(s)N_H(s)}$$

Again suppose $D(s) = \frac{a(s)}{b(s)}$; a, b polynomials

Then additional error:

$$S_i(s)D(s) = \left[\frac{N_G(s)D_H(s)}{D_G(s)D_H(s) + N_G(s)N_H(s)}\right]\left[\frac{a(s)}{b(s)}\right]$$

Internal model principle again!

If $N_G(s)D_H(s)$ cancels non-stable roots of $b(s)$ then in steady-state $\mathcal{L}^{-1}\{S_i D\} = \emptyset$

i.e. disturbance creates <u>NO</u> additional error!

Implications:

If $N_G(s)D_H(s)$ cancels non-stable roots of $b(s)$, then cancellation is either due to:

$\Rightarrow$ $N_G(s)$ cancelling (extremely rare)

$\Rightarrow$ $D_H(s)$ cancelling (can design for this)

So generally, external disturbances create NO Add'l error if compensator contains an internal model of disturbance.

That is, if compensator $H(s)$ has same non-stable poles as the disturbance.

"perfect rejection" of dist.

i.e. if $d(t) = d_0$ (constant), no add'l tracking error if $H(s)$ has pole at origin.

Summary of error analysis

For perfect tracking of "type p" desired behaviors

$$y_d(t) = \left(\frac{A_p}{p!}\right) t^p$$

L(s) must have p+1 poles at origin

For perfect rejection of type p disturbances d(t),
H(s) must have p+1 poles at origin

in Both cases, p poles at origin (one less) will ensure finite, but nonzero errors

Note: tracking objectives can be satisfied if required poles come from plant, compensator, or a combination of both

But dist. rejection reqt's can be satisfied only by poles in the compensator.

$\Rightarrow$ Above are special cases of IMP.

Good accuracy thus often requires $H(s)$ to have at least one pole at origin.

$\Rightarrow$ This pole adds -90° of phase at all frequencies!

$\Rightarrow$ Works against our stability/performance guidelines of increasing phase margin.

$\Rightarrow$ Even adding a LHP zero doesn't help here:

$$H(s) = K\left[\frac{s - z_c}{s}\right] \qquad z_c < \emptyset$$

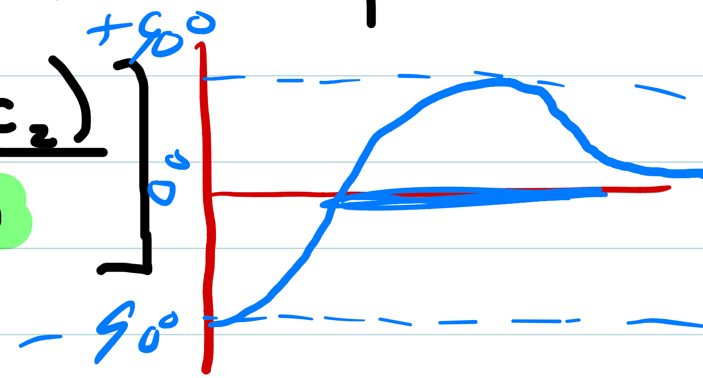
has $\measuredangle H(j\omega) < \emptyset^\circ$ for all $\omega \geq 0$.

$\Rightarrow$ May be acceptable if $\measuredangle G(j\omega)$ already has "adequate" positive phase, so $\measuredangle L = \measuredangle G + \measuredangle H$ can tolerate a small reduction.

More generally, we'd require extra LHP zero(s) to still provide positive phase changes to $L(s)$ despite required pole at origin

Implementability then requires an additional LHP pole:

$$H(s) = K\left[\frac{(s-z_{c_1})(s-z_{c_2})}{s\,(s-p_c)}\right]$$



4 degrees of freedom total!

Things get even more complicated if $H(s)$ needs ≥ 2 poles at origin to achieve tracking objectives!

Remember: Tracking of $y_d(t)$ depends on IMP properties of $L(s)$

Disturbance rejection depends on IMP properties of $H(s)$

More generally, we'd require extra LHP zero(s) to still provide positive phase changes to $\angle L(s)$ despite required pole at origin

Implementability then requires an additional LHP pole:

$$H(s) = K\left[\frac{(s-z_{c_1})(s-z_{c_2})}{s(s-p_c)}\right]$$

4 degrees of freedom total!

Things get even more complicated if $H(s)$ needs ≥ 2 poles at origin to achieve tracking objectives!

<u>Remember</u>: Tracking of $y_d(t)$ depends on IMP properties of $L(s)$

Disturbance rejection depends on IMP properties of $H(s)$

Robustness

Robustness is range of inaccuracy in our Nominal model $G(s)$ that we can tolerate before feedback loop *might* become unstable.

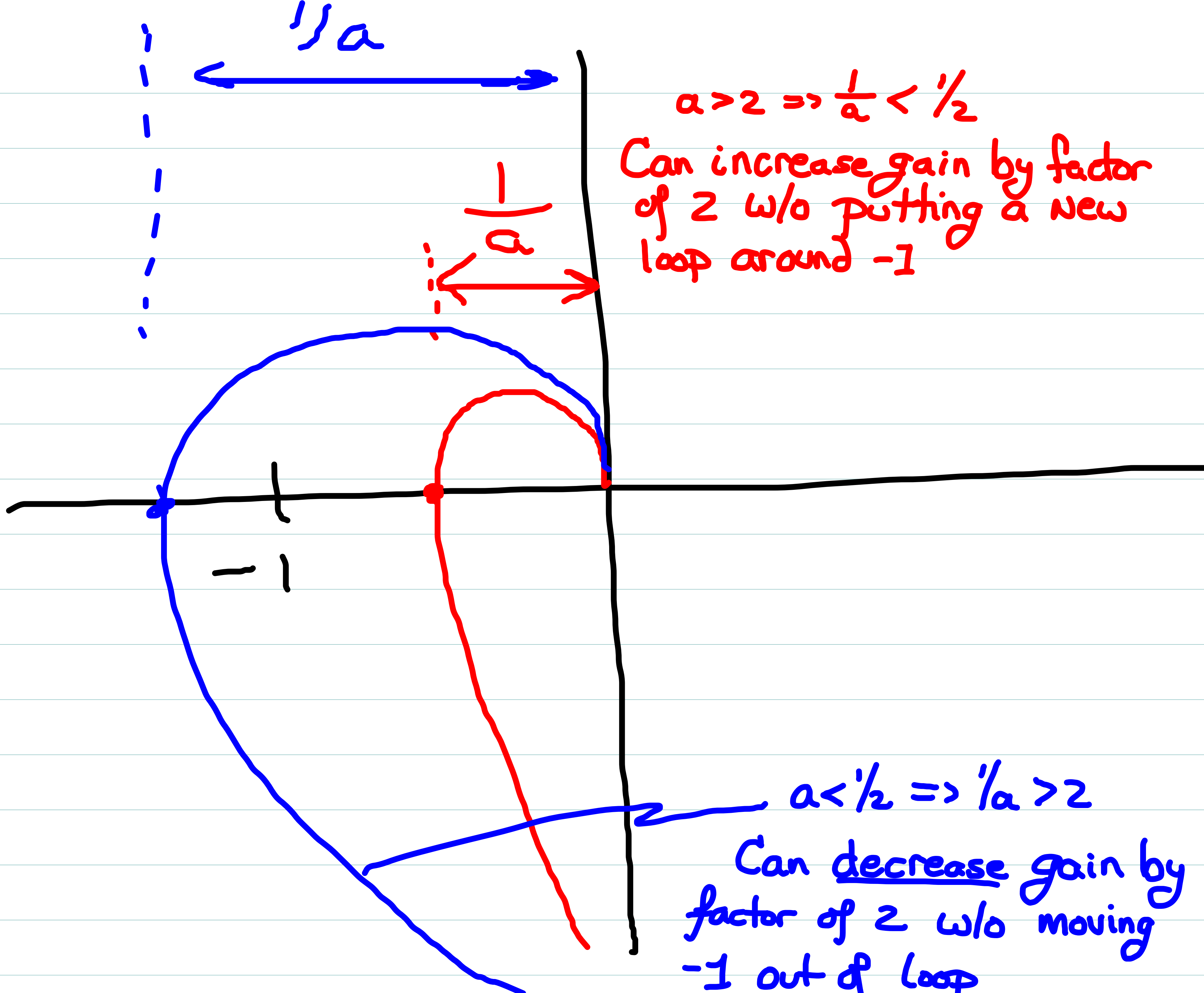
"Perturbations" to Nyquist analysis: how much can polar plot of $L(j\omega)$ be changed *without* changing the number of -1 encirclements.

Simple measures:

① gain margin: Measures tolerance to pure gain uncertainty

Common requirement: $|a|_{dB} \geq 6 \Rightarrow a \geq 2$ or $a \leq 1/2$

$\Rightarrow$ Plant gain could be a factor of 2 larger or smaller and -1 encirclements will not change.



$a > 2 \Rightarrow \frac{1}{a} < 1/2$

Can increase gain by factor of 2 w/o putting a new loop around -1

$a < 1/2 \Rightarrow 1/a > 2$

Can <u>decrease</u> gain by factor of 2 w/o moving -1 out of loop

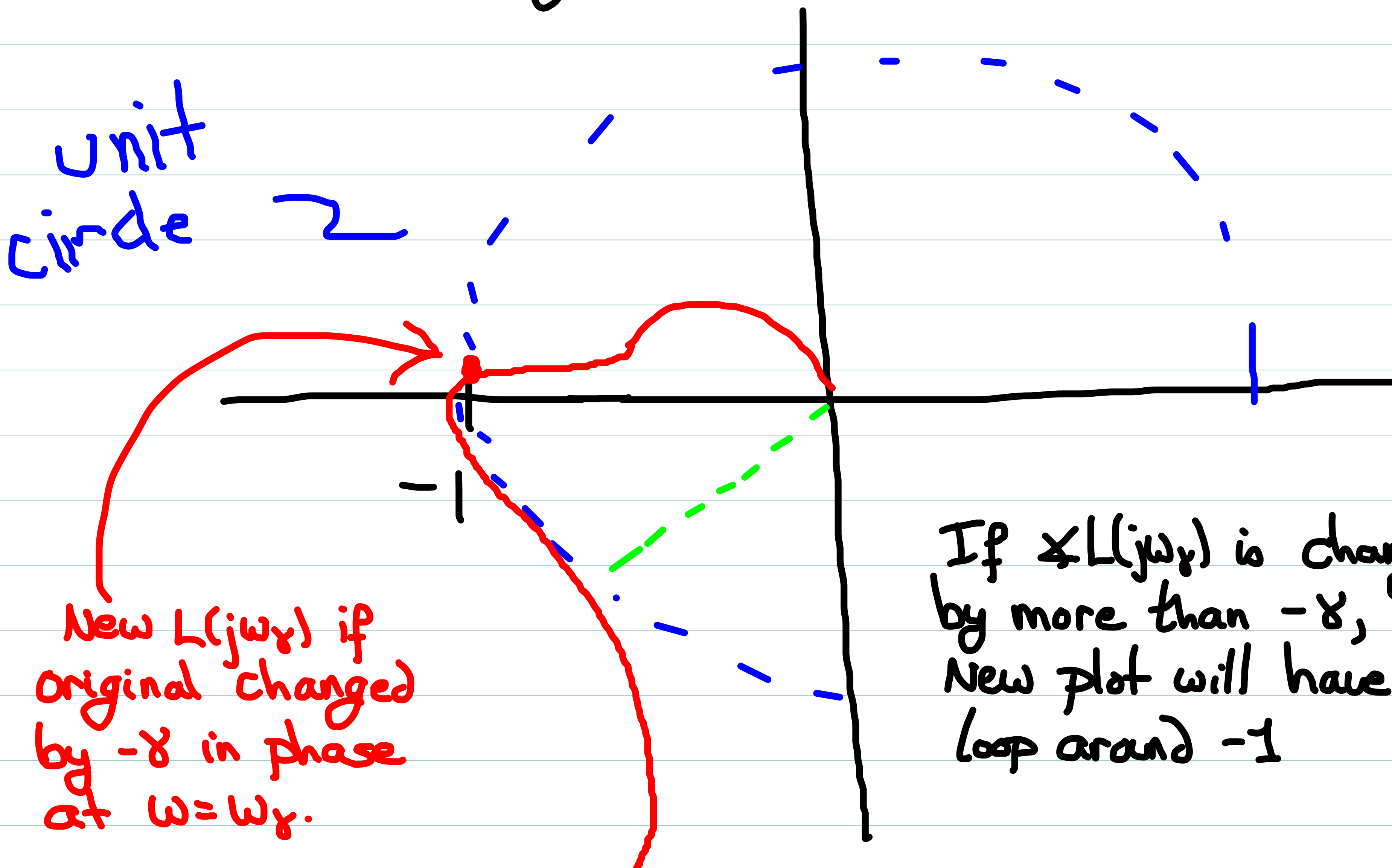
Simple Robustness measure #2: phase margin, γ

Measures pure <u>phase</u> uncertainty tolerable before -1 encirclements change

If $\measuredangle L(j\omega_\gamma)$ is changed by more than $-\gamma$, the new plot will have a loop around -1

Simple Robustness measure #2: phase margin, γ

Measures pure <u>phase</u> uncertainty tolerable before -1 encirclements change



If $\measuredangle L(j\omega_\gamma)$ is changed by more than $-\gamma$, the New plot will have a loop around -1

Physical Sources of <u>Pure</u> Phase Change

Phase margin is an important metric, so there must be an important, common physical mechanism which can introduce pure phase changes. What is it?

$$\boxed{\text{Time Delay}}$$

We've been modeling our controller as <u>continuously</u> evolving, just like the physical system being controlled.

But the controller is <u>different</u> than a physical system with dynamics governed by continuous diff'l equations.

Models of these differences will create pure phase changes to $L(j\omega)$.

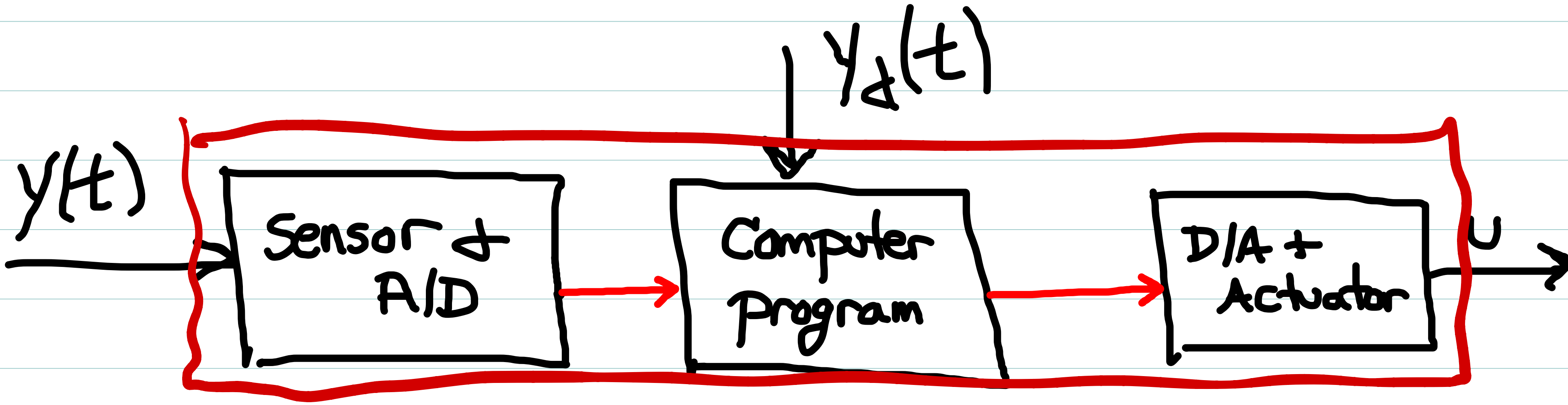
Time Delay

Three typical steps in controller implementation

1. Measure output $y(t)$, and input to computer
2. Compute $u(t)$ via computer program
3. Output $u(t)$ from computer to physical actuator

Each of these steps requires nonzero amount of time!

1. A/D conversion and transmission/read time
2. Time to execute program
3. D/A conversion and transmission time



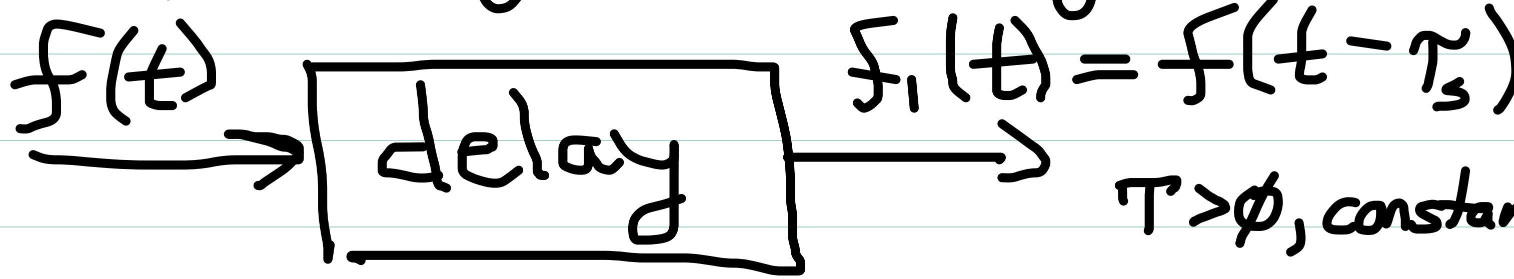
Each block, and each red arrow, requires nonzero time to operate. Call total required time T_s

T_s may be small (msec), but is always > 0!

The implication is that the $u(t)$ which actually gets applied to the plant depends on the measurement taken T_s seconds ago, i.e. $y(t - T_s)$

We haven't modeled this!

Laplace analysis of ideal delay



$\tau > \emptyset$, constant

By def'n:

$$F_1(s) = \int_{0^-}^{\infty} f_1(t)\, e^{-st}\, dt = \int_{0^-}^{\infty} f(t-\tau_s)\, e^{-st}\, dt$$

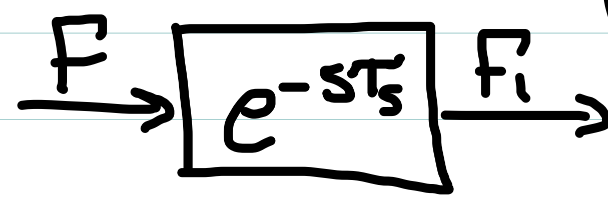
Let $\sigma = t - \tau_s$, $d\sigma = dt$

$$F_1(s) = \int_{-\tau_s}^{\infty} f(\sigma)\, e^{-s(\sigma+\tau_s)} \;=\; e^{-s\tau_s}\int_{\emptyset^-}^{\infty} f(\sigma)\, e^{-s\sigma}\, d\sigma$$

Since τ_s is constant, and Laplace assumes $f(t) = \emptyset$ for $t < \emptyset$

Thus:

$$F_1(s) = e^{-s\tau_s} F(s)$$



transfer function!

So really:

$$L(s) = \underbrace{e^{-sT_s}}_{\text{transfer function of delay}} \underbrace{\left[G(s)H(s)\right]}_{L_0(s)}$$

$L_0(s)$: "ideal" (no delay) open-loop TF.

Now, e^{-sT_s} is difficult to deal with in standard TF manipulations, because it is Not rational. Cannot be described with a finite number of poles and zeros.

It's impact on freq. Domain properties of $L(j\omega)$ are easy to determine, however.

$$|L(j\omega)| = \boxed{|e^{-j\omega T_s}|}\, |L_0(j\omega)|$$

$$\angle L(j\omega) = \boxed{\angle e^{-j\omega T_s}} + \angle L_0(j\omega)$$

↳ what are these?

Recall for complex number in polar form:

$$z = re^{j\theta} \Longleftrightarrow |z| = r, \ \measuredangle z = \theta$$

$$e^{-j\omega T_s} = 1 \cdot e^{j(-\omega T_s)} \Longrightarrow r = 1, \ \theta = -\omega T_s$$

So $|e^{-j\omega T_s}| = 1$ for all ω, and

$$\measuredangle e^{-j\omega T_s} = -\omega T_s \quad \text{for all } \omega$$

Hence:

$$|L(j\omega)| = |L_0(j\omega)|$$ (unaffected by delay)

$$\measuredangle L(j\omega) = \measuredangle L_0(j\omega) - \omega T_s$$

$\boxed{\text{Effect of delay is pure phase change in } L(j\omega)!}$

Delay thus acts to reduce phase margin:

$$\gamma = 180° + \measuredangle L(j\omega) = \underbrace{180° + \measuredangle L_0(j\omega_\gamma)}_{\gamma_0 \text{: expected Phase margin w/o delay}} - \underbrace{\omega_\gamma T_s}_{\text{reduction in actual phase margin due to delay}}$$

i.e. $\boxed{\gamma = \gamma_0 - \omega_\gamma T_s}$ ⟸ Key equation!

Note: ω_γ in rad/sec, T_s in sec ⟹ $\omega_\gamma T_s$ in rad

γ_0, γ expressed in deg, so must convert $\omega_\gamma T_s$ to deg here

Recall we typically need $\gamma > 0°$ for Nyquist to show stability

$$\Rightarrow \gamma_0 - \omega_\gamma T_s > 0 \quad \text{or} \quad T_s < \frac{\gamma_0(\text{rad})}{\omega_\gamma}$$

$\boxed{T_{max} = \dfrac{\gamma_0(\text{rad})}{\omega_\gamma}}$ is the maximum tolerable delay, or the "delay margin"

Now, typically T_s is fixed by available hardware.

Then $\gamma_0 - \omega_\gamma T_s > 0$ Becomes a design constraint

$\Rightarrow$ Cannot have $\omega_\gamma T_s$ "too big" or it will be impossible to design $H(s)$ to provide necessary positive phase for γ_0.

Typical guideline: Keep $\boxed{\omega_\gamma T_s < 0.1}$ $\left(\omega_\gamma < \frac{1}{10 T_s}\right)$

Then $\gamma = \gamma_0 - \omega_\gamma T_s \geq \gamma_0 - 5.7°$ $(0.1\,\text{rad} = 5.7°)$

Can design $H(s)$ to provide additional $+5.7°$ of phase margin in γ_0 to offset (or, just tolerate the small reduction)

$\Rightarrow$ Note this constrains ω_γ in a manner which works against guideline for good performance (big ω_γ)

$\Rightarrow$ **Sample rate fundamentally restricts performance!**

Different uses of delay eq'n

① Delay margin => $T_{max} = \gamma/\omega_\gamma$. Max tolerable delay w/o creating instability. ('common figure of merit')

② If T_s fixed, rule of thumb $\omega_\gamma T_s < .1$ restricts ω_γ, i.e. $\omega_\gamma < 1/(10T_s)$ (common)

③ If T_s can be changed (hardware upgrade) then $T_s < \frac{1}{10\omega_\gamma}$ needed to keep delay effect "small" (uncommon, except in early design phase)

④ $\gamma = \gamma_0 - \omega_\gamma T_s$.

Given fixed ω_γ, T_s, target $\gamma_0 = \gamma_{des} + \omega_\gamma T_s$

so $\gamma = \gamma_{des}$

i.e. design $H(s)$ so $L_0(s) = \tilde{G}(s)H(s)$ (OL TF w/o delay) has PM $\gamma_{des} + \omega_\gamma T_s$ at desired ω_γ (rare, but possible sometimes)

Gain and Phase margins are measures of robustness

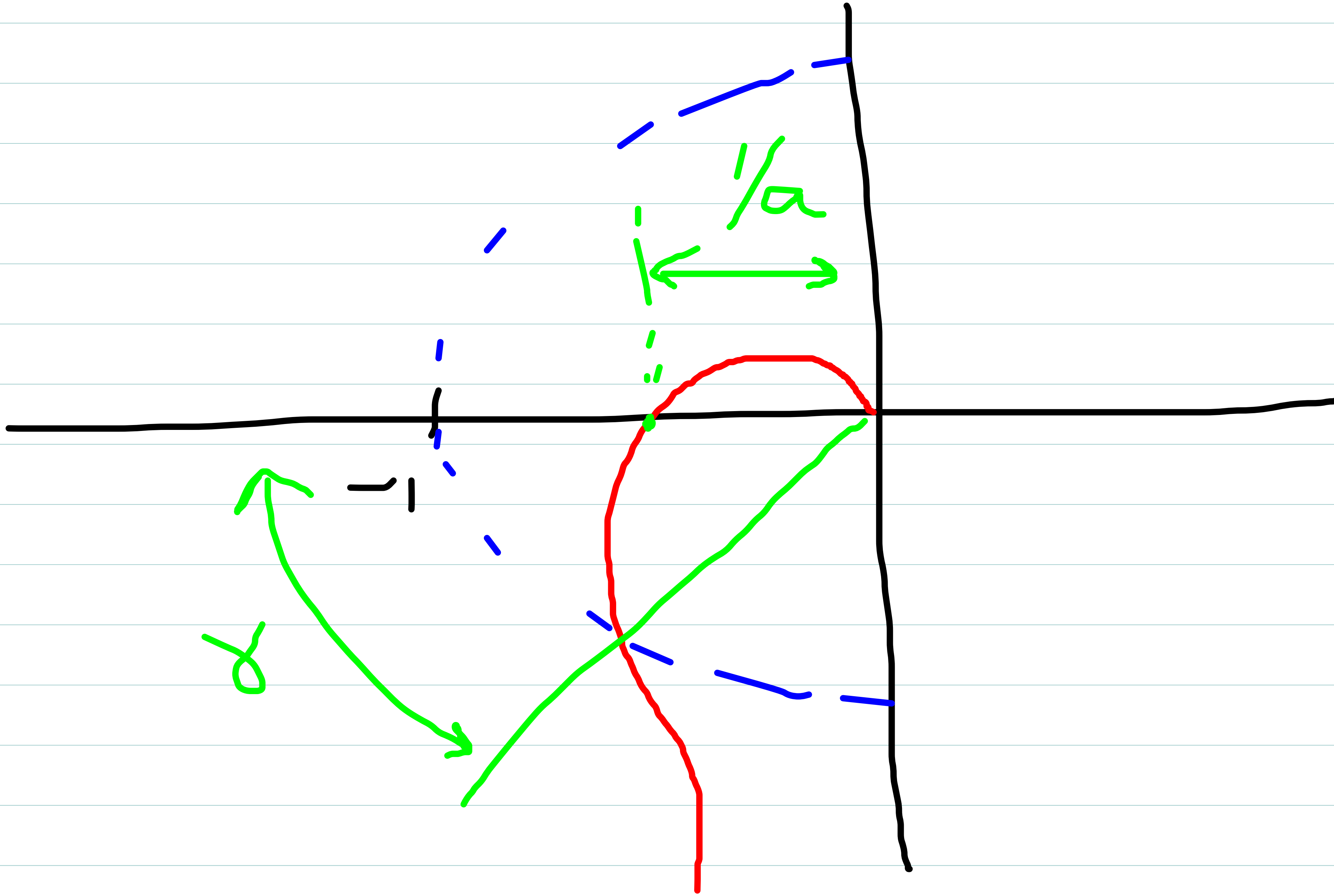
They quantify how close Nyquist diagram comes to -1 in two simple senses

Very common and popular since each corresponds to a physical source of possible model inaccuracy

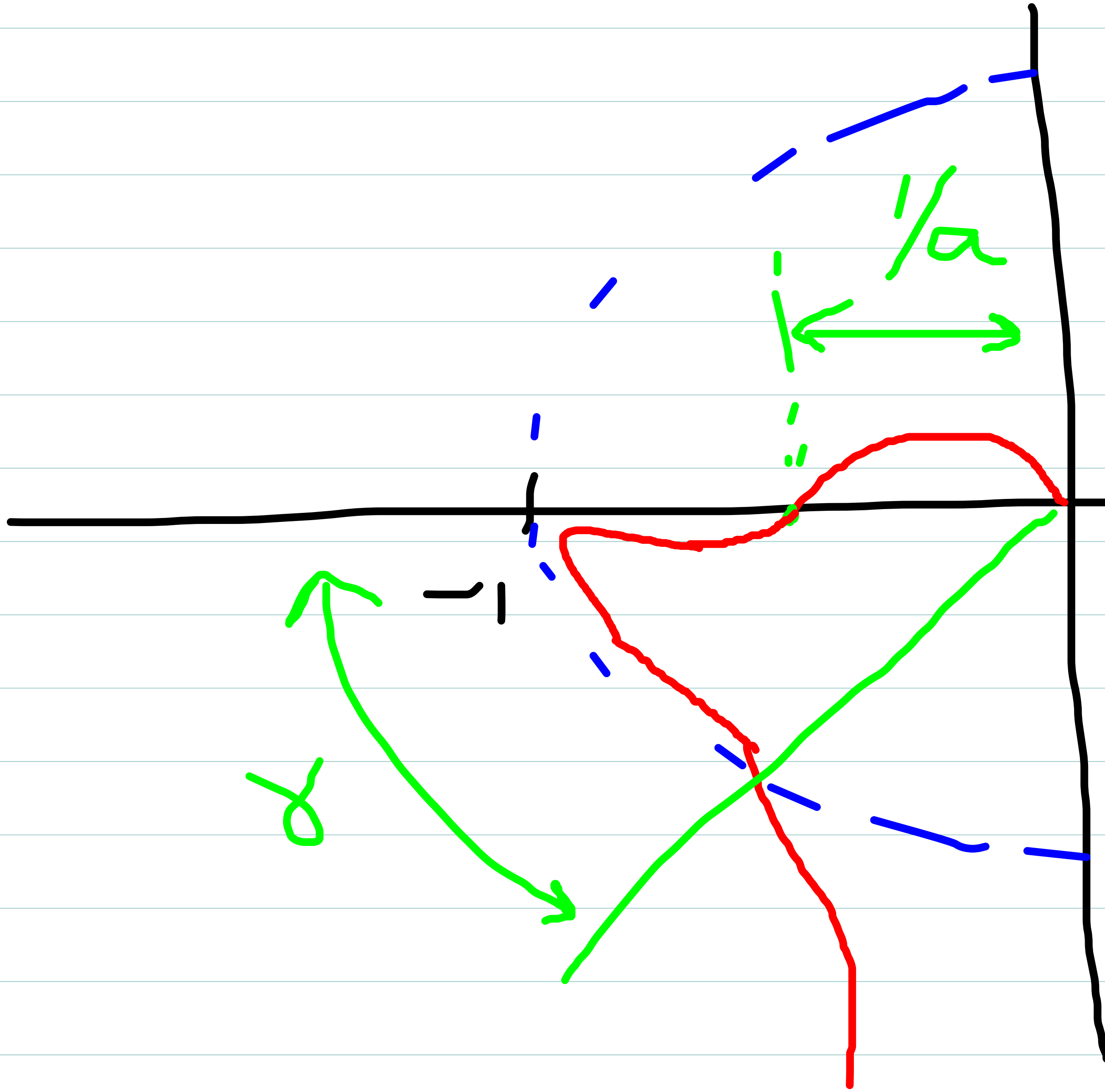
⇒ gain margin: tolerance to variations in overall gain of plant (typically overall mass or inertia)

⇒ phase margin: tolerance to time delays associated with computer implementation of controller

However: mathematically they are poor indicators of the tolerance of the Nyquist diagram to small perturbations



A typical case



Gain+phase margin same, but Nyquist (polar) comes arbitrarily close to -1.

⇒ arbitrarily small change could destabilize!

But this is possible too!

Gain and phase margin are useful, intuitive measures but cannot capture the effects of simultaneous gain and phase changes to $L(j\omega)$

Such changes would occur due to:

⇒ mismodeling of pole/zero locations in $G(s)$

⇒ Incompleteness of $G(s)$ model, i.e. physics has additional dynamics which are too uncertain, or too difficult, to model accurately

⇒ "real" $G(s)$ has additional poles/zeros which aren't present in model we use for design!

⇒ Want a robustness test which can also handle these!

"Phasor" Notation

Observation: Complex number add'n/sub'n follows same rules as 2D (planar) vectors

$z_1 = a_1 + b_1 j$, $z_2 = a_2 + b_2 j$ $\qquad$ $\underline{V}_1 = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}$ $\qquad$ $\underline{V}_2 = \begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$

$z_3 = z_1 + z_2$ $\qquad$ $\underline{V}_3 = \underline{V}_1 + \underline{V}_2$

$= (a_1 + a_2) + (b_1 + b_2) j$ $\qquad$ $= \begin{bmatrix} a_1 + a_2 \\ b_1 + b_2 \end{bmatrix}$

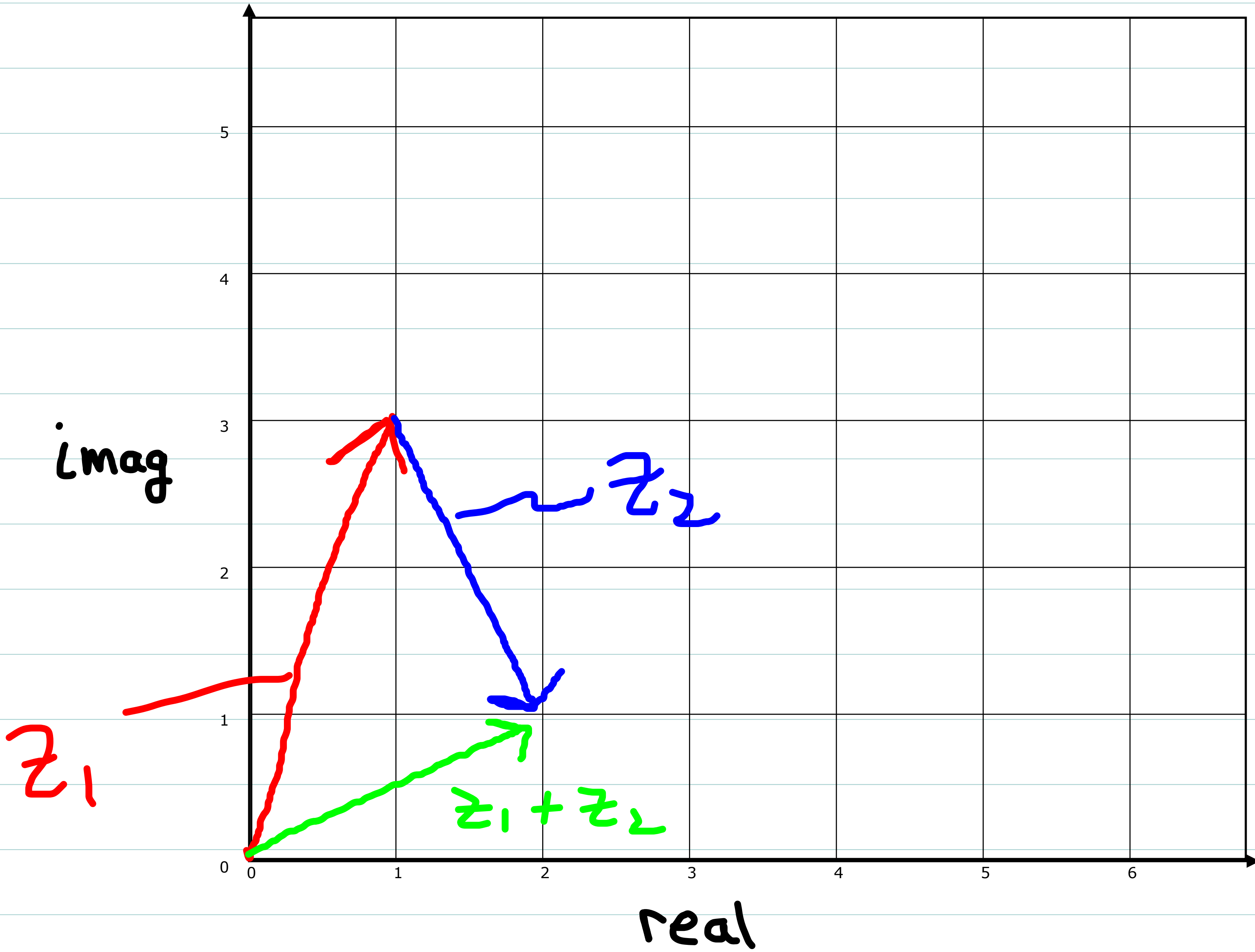
i.e. identify real part with 1st component of 2D vector, imag part with 2nd component.

$\Rightarrow$ Can interpret complex numbers as planar vectors

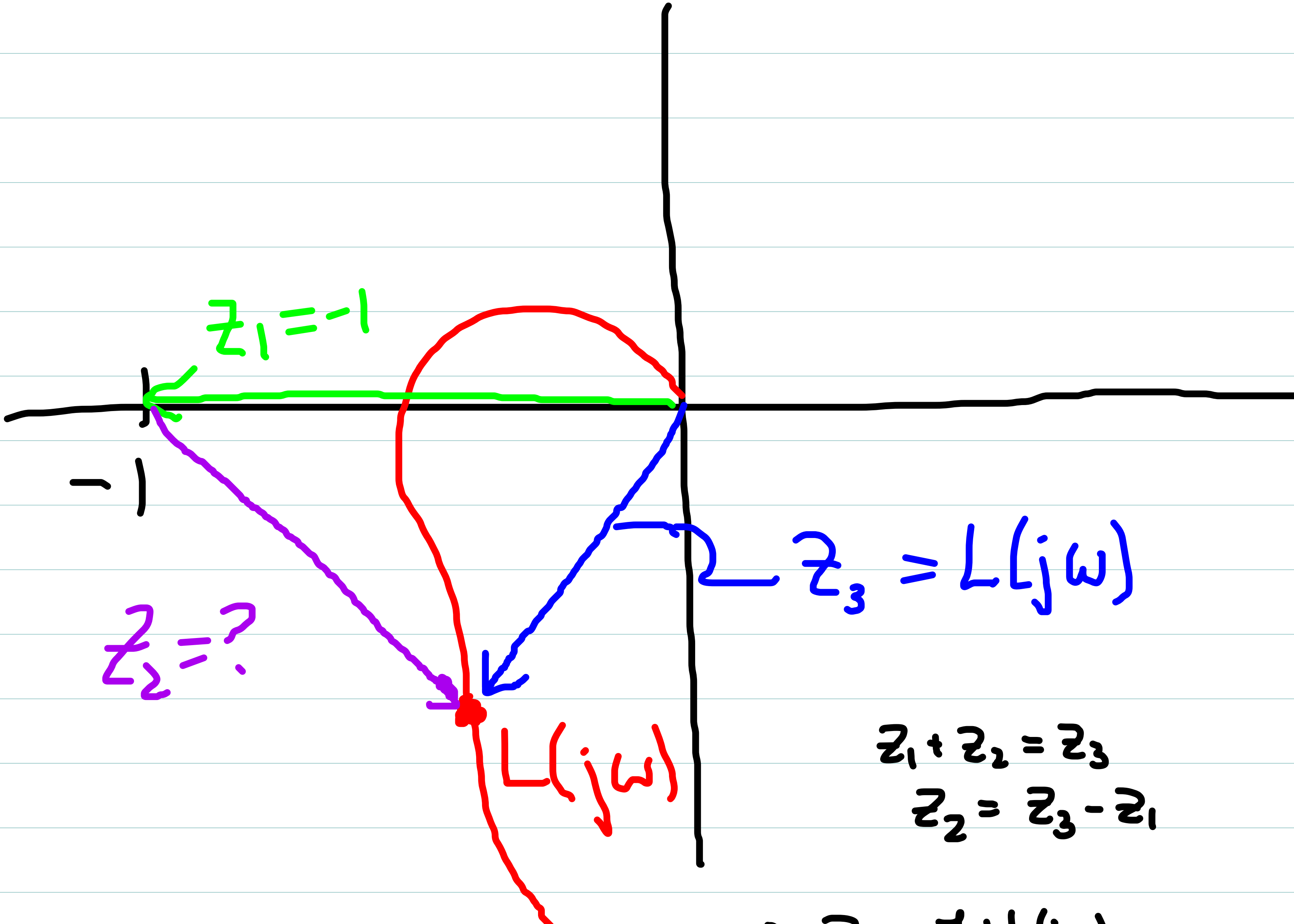
$\Rightarrow$ Can use vector graphical add'n tricks for complex numbers

Example

$$z_1 = 1+3j, \quad z_2 = 1-2j \Rightarrow z_3 = z_1+z_2 = 2+j$$



Important Application



$$z_1 + z_2 = z_3$$

$$z_2 = z_3 - z_1$$

$$\Rightarrow z_2 = 1 + L(j\omega)$$

Thus:

Complex number $1+L(j\omega)$ can be graphically visualized as the phasor from -1 to $L(j\omega)$ on polar plot.

$\Rightarrow$ $|1+L(j\omega)|$ is the distance from -1 to polar plot at freq ω.

$\Rightarrow$ Good robustness requires this doesn't get too small!

$\Rightarrow$ But note: $|1+L(j\omega)| = |S(j\omega)|^{-1}$

$\Rightarrow$ Thus, good robustness requires $|S(j\omega)|$ not get too big.

$\Rightarrow$ Good designs have $|S(j\omega)|$ which do not exhibit a large <u>peak</u>!

$$\left[\max_{\omega} |S(j\omega)|\right]^{-1} = \min_{\omega} |1 + L(j\omega)|$$

$= $ smallest distance from -1 to the polar/Nyquist plot

We do not want this to be too small, hence

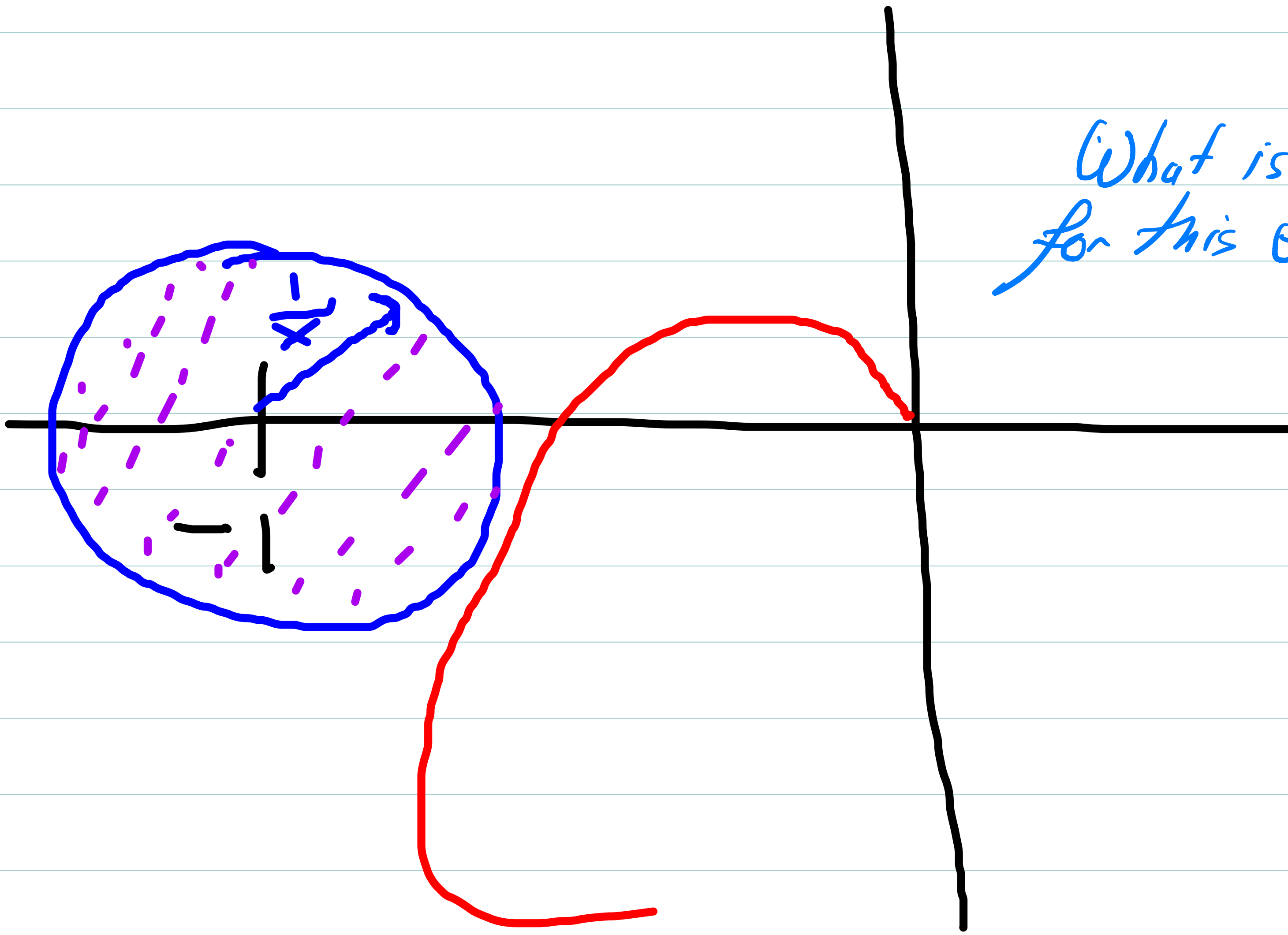
we need to ensure $\max_{\omega} |S(j\omega)|$ is not too big.

What is an appropriate target for $\max_{\omega} |S(j\omega)|$?

Now:

$$|S(j\omega)|_{max} < X \Rightarrow |1 + L(j\omega)| > \frac{1}{X} \text{ for all } \omega \geq 0$$

$\Rightarrow$ Polar (Nyquist) diagram of $L(j\omega)$ cannot enter a disk of radius $\frac{1}{X}$ centered at -1



What is a "good" size for this exclusion disk?

This property guarantees certain minimum phase + gain margins

for example, can show: $|S(j\omega)|_{max} < 2$ (+6 dB) $\Rightarrow |1+L(j\omega)| > 1/2$

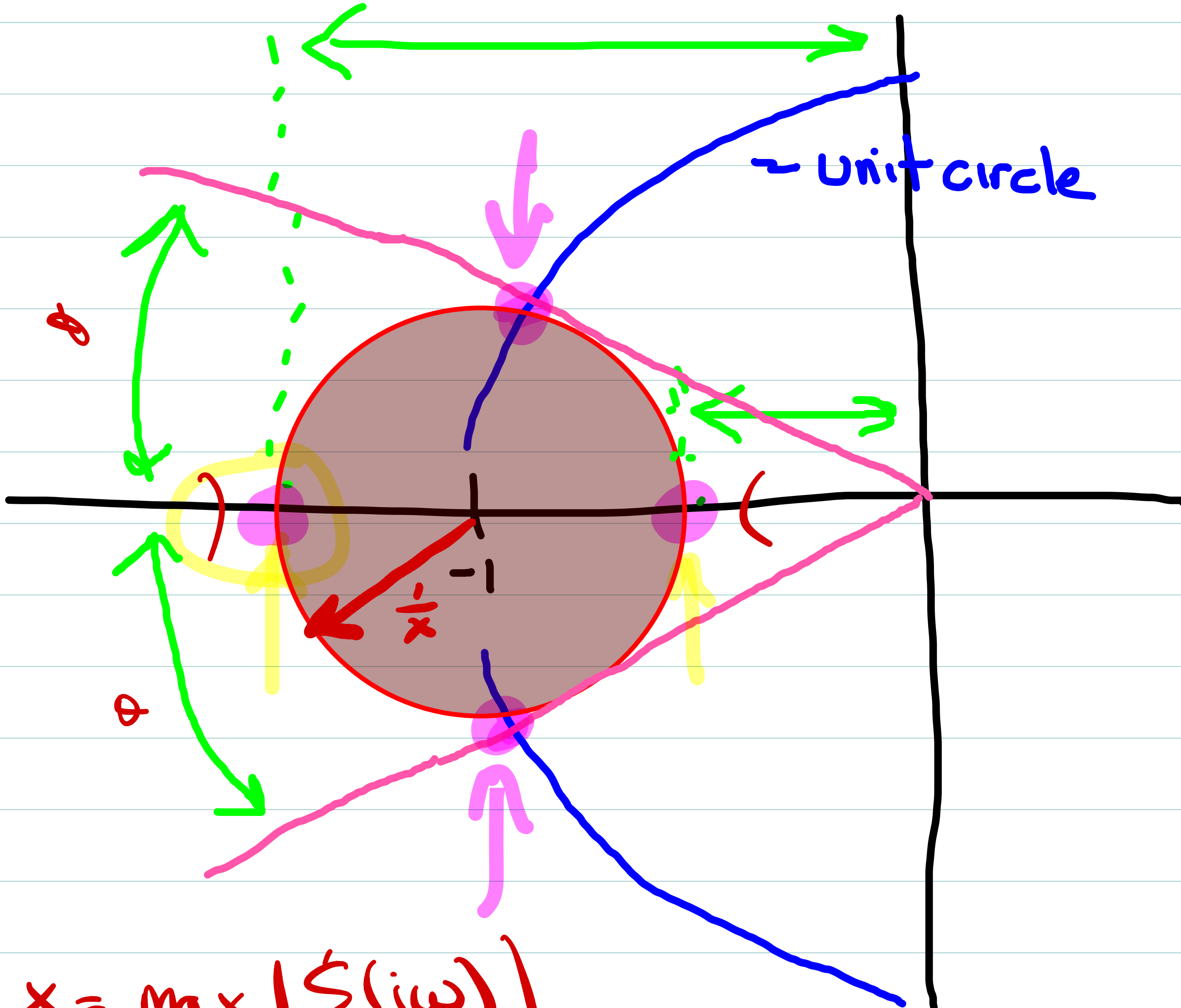
$\Rightarrow a < 2/3$ (−3.5 dB), $a > 2$ (+6 dB)

$\Rightarrow |\gamma| > 29°$

(Note that these are pretty close to the common industry standard req'ts: $|a|_{dB} \geq 6$, $|\gamma| \geq 30°$)

<u>However</u>, a specific set of gain, phase margins <u>does</u> <u>not</u> conversely guarantee a bound on $|S(j\omega)|_{max}$ (as shown in previous example!)

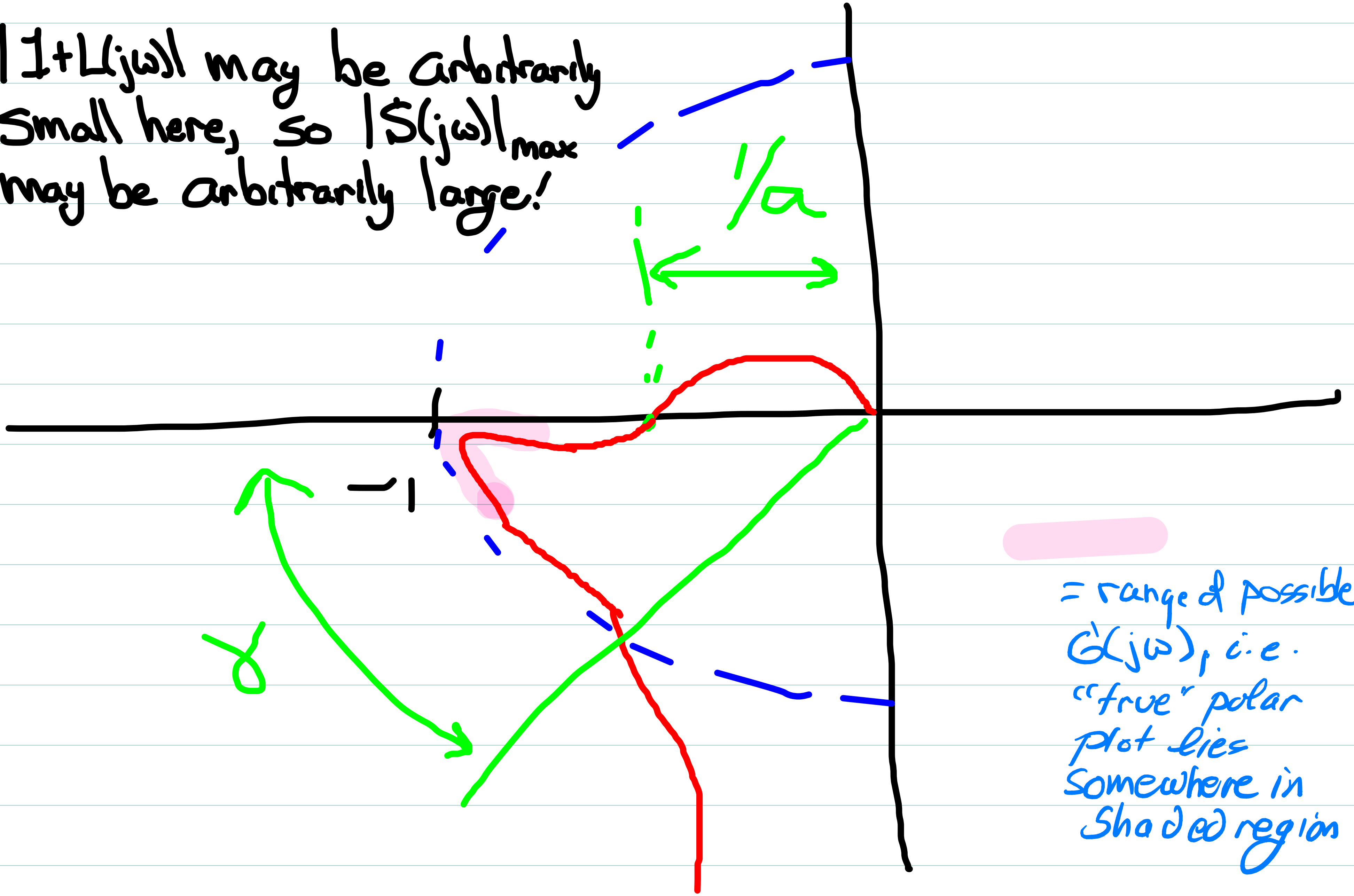
$\Rightarrow |S(j\omega)|_{max}$ (peak of sensitivity diagram) is a superior measure of robustness, and $|S(j\omega)|_{max} \lesssim +6$ dB is a good nominal target.



$$x = \max_{\omega} |S(j\omega)|$$

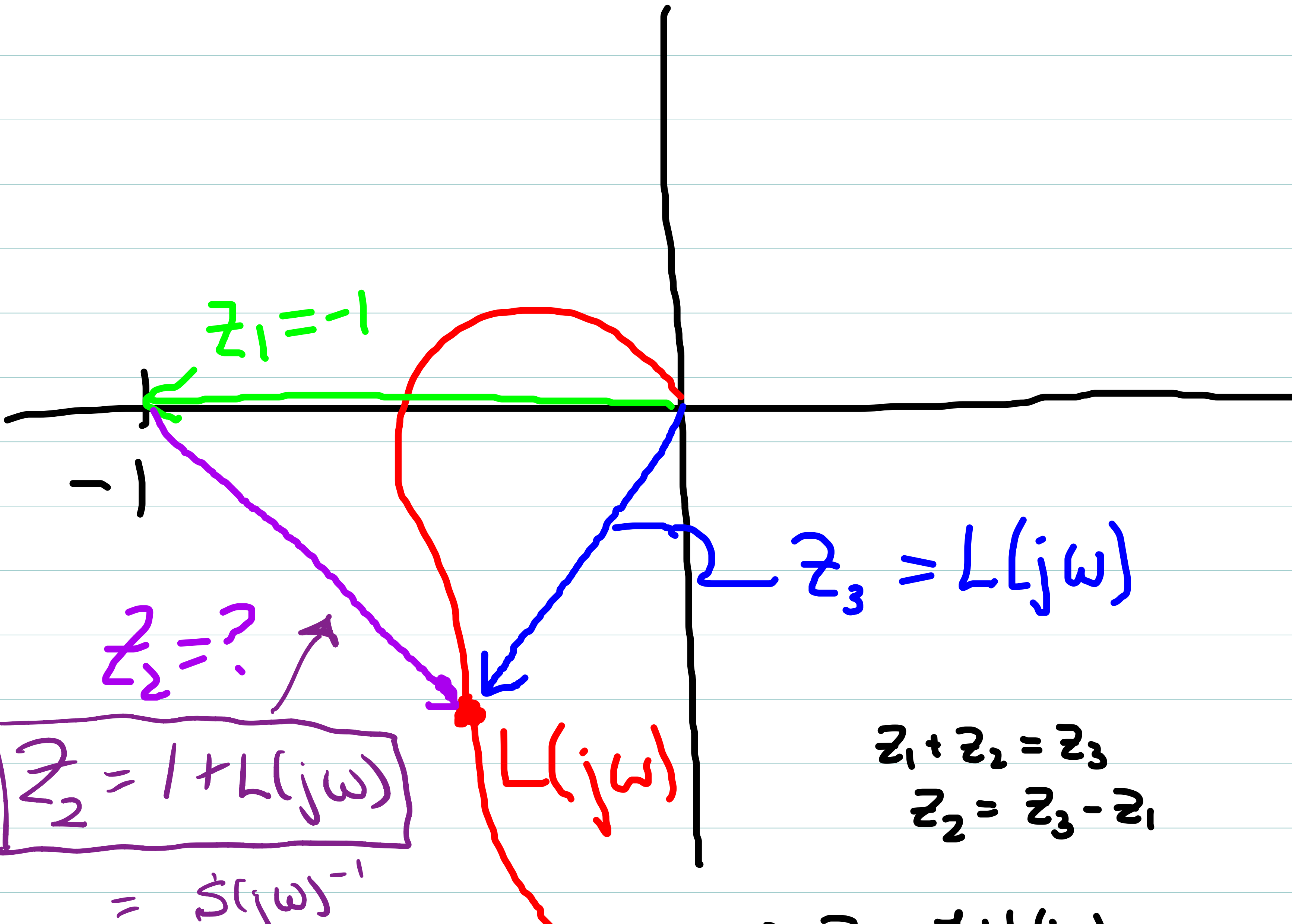
$$\frac{1}{x} = \min_{\omega} |1 + L(j\omega)|$$

$|1+L(j\omega)|$ may be arbitrarily small here, so $|S(j\omega)|_{max}$ may be arbitrarily large!



[Not a lot of room to tolerate model error if peak of $|S(j\omega)|$ is large]

Important Application



$$\boxed{Z_2 = 1 + L(j\omega)}$$

$$= S(j\omega)^{-1}$$

$$z_1 + z_2 = z_3$$

$$z_2 = z_3 - z_1$$

$$\Rightarrow z_2 = 1 + L(j\omega)$$

$|1+L(j\omega)|$ may be arbitrarily small here, so $|S(j\omega)|_{max}$ may be arbitrarily large!

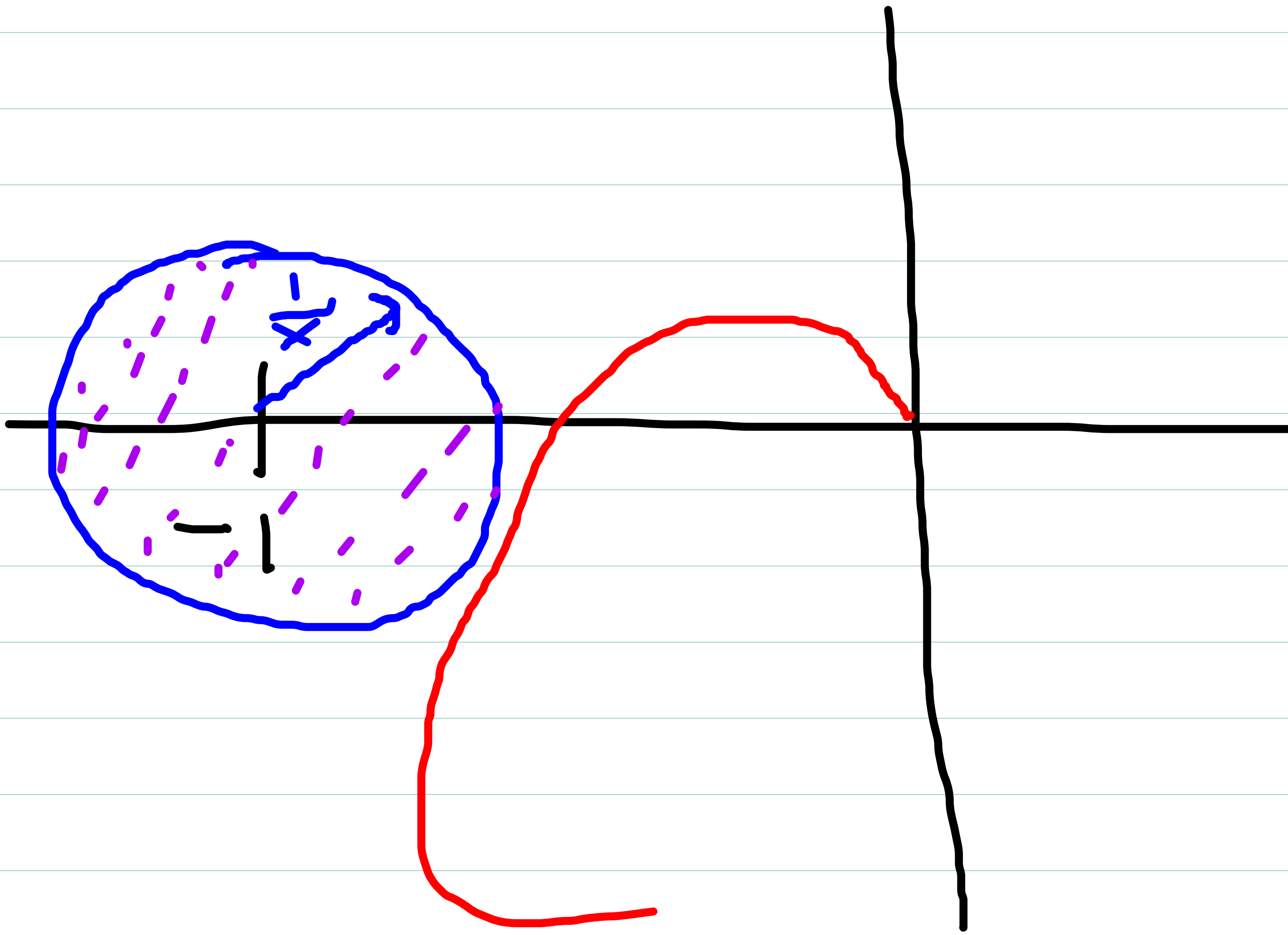
= range of possible $G(j\omega)$, i.e. "true" polar plot lies somewhere in shaded region.

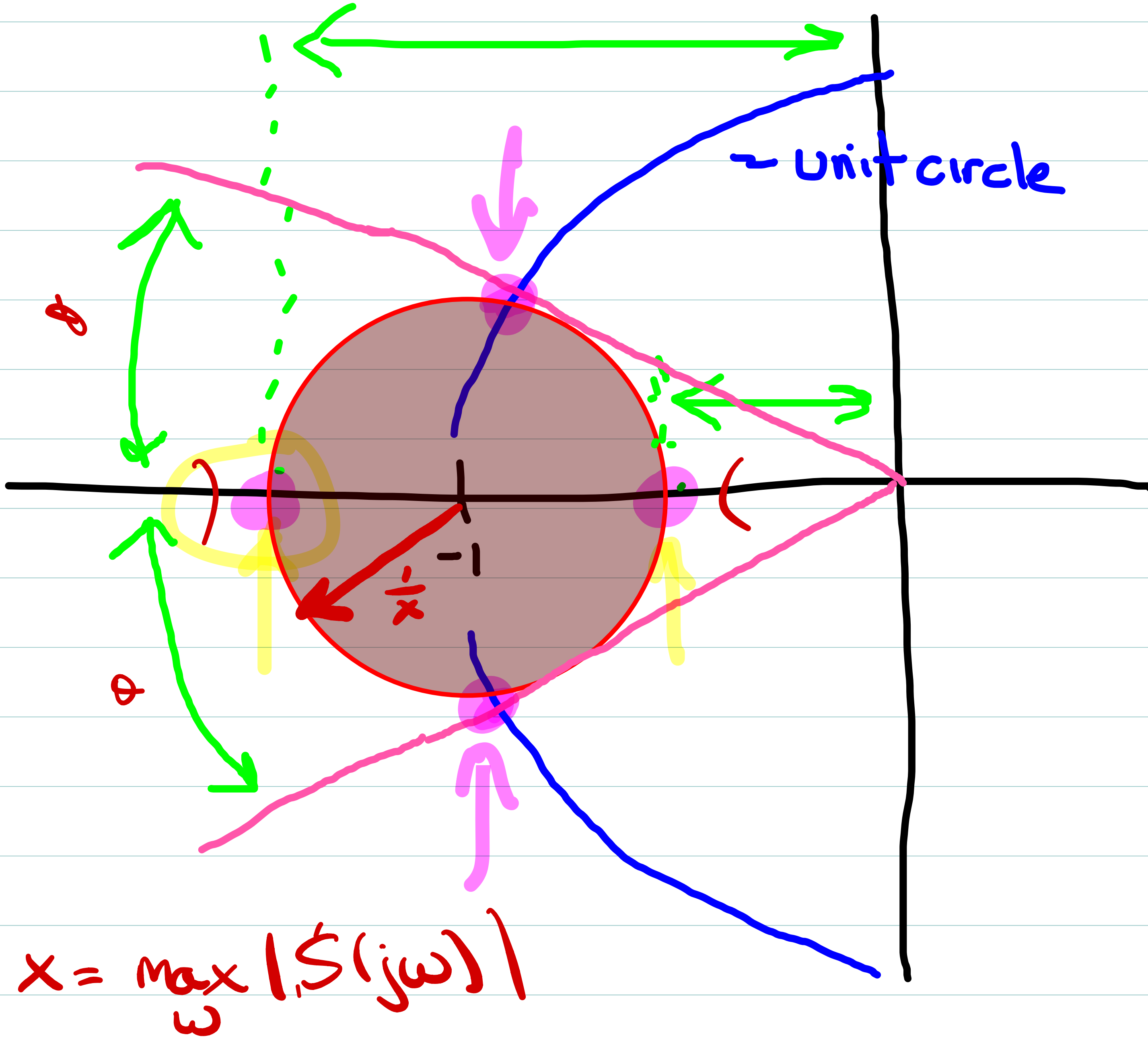
[Not a lot of room to tolerate model error if peak of $|S(j\omega)|$ is large]

Now:

$$|S(j\omega)|_{max} < X \Rightarrow |1 + L(j\omega)| > \frac{1}{X} \text{ for all } \omega \geq 0$$

$\Rightarrow$ Polar (Nyquist) diagram of $L(j\omega)$ cannot enter a disk of radius $\frac{1}{X}$ centered at -1





unit circle
ϕ
ϕ
-1
$\frac{1}{x}$
$x = \max_{\omega} |S(j\omega)|$

We can do much more with this idea!

Let $G_o(s)$ be our nominal plant model (what we use in Matlab)

Let $G(s)$ be the "true" plant TF (unknown)

Define:

$$\Delta(s) = \left[\frac{G(s) - G_o(s)}{G_o(s)} \right] = \left[\frac{G(s)}{G_o(s)} - 1 \right]$$

A Normalized measure of error in nominal model

We don't know what $\Delta(s)$ is, but may be able to place bounds on how "big" it can be to still ensure stability of feedback system.

Let:

$L_o(s) = G_o(s)\,H(s)$ Nominal OL TF

$L(s) = G(s)\,H(s)$ True OL TF "multiplicative uncertainty model"

The def'n of $\Delta(s)$ implies $G(s) = G_o(s)\,[1 + \Delta(s)]$

Hence $L(s) = G_o(s)\,H(s)\,[1 + \Delta(s)]$

$= G_o(s)\,H(s) + G_o(s)\,H(s)\,\Delta(s)$

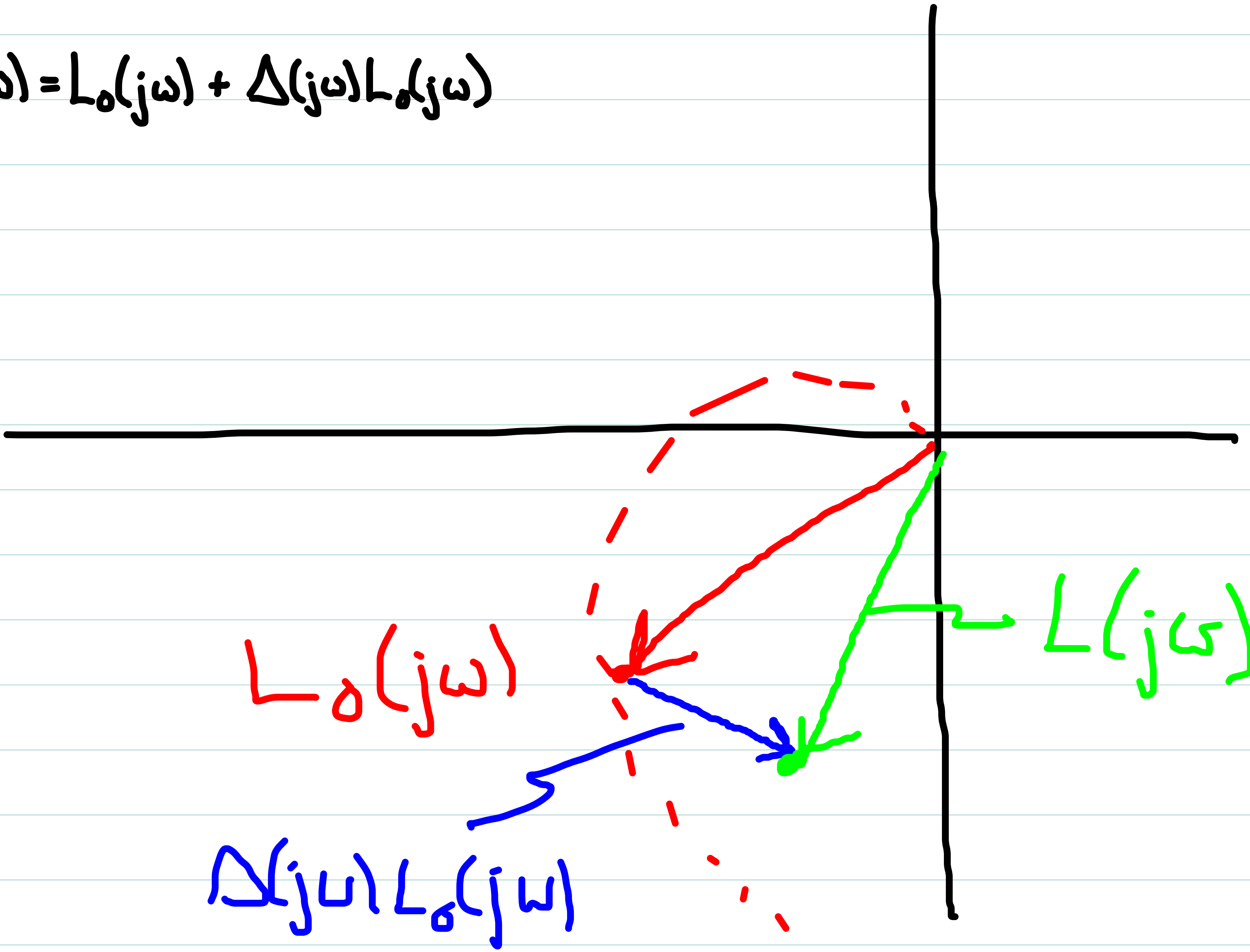
Or: $L(s) = L_o(s) + L_o(s)\,\Delta(s)$

and for each $\omega \geq 0$:

$$\underbrace{L(j\omega)}_{\text{true polar plot}} = \underbrace{L_o(j\omega)}_{\text{Nominal Polar Plot}} + \overbrace{L_o(j\omega)\,\Delta(j\omega)}^{\text{effect of model error on polar plot}}$$

Phasor Interpretation

$L(j\omega) = L_o(j\omega) + \Delta(j\omega)L_o(j\omega)$



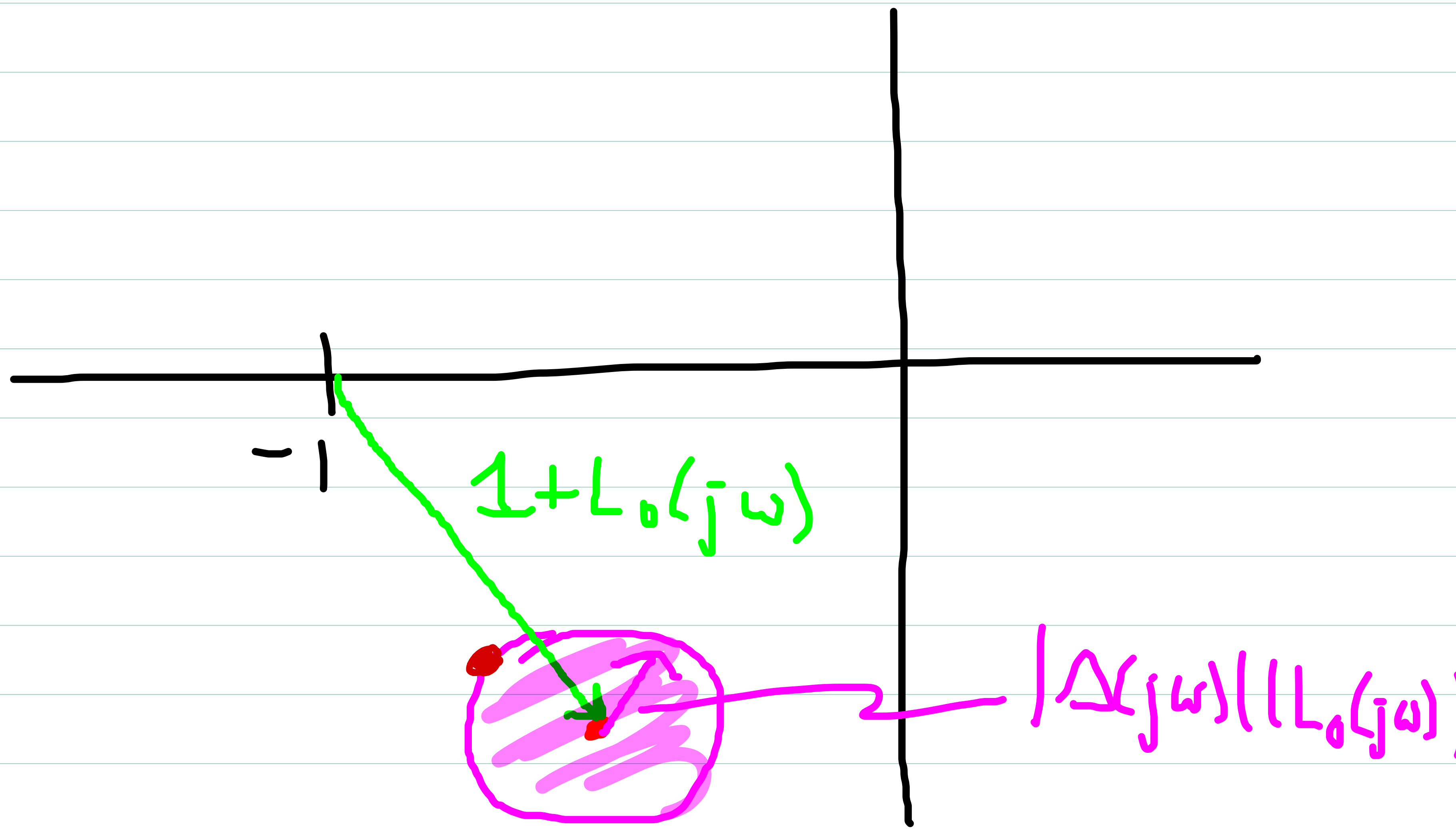
<u>Note</u>: $\Delta(j\omega)$ has unknown magnitude and direction

Assume: $\Delta(j\omega)$ can have <u>any</u> direction (worst case).

$\Rightarrow$ $L(j\omega)$ can lie anywhere in a <u>disk</u> of radius $|\Delta(j\omega)||L_o(j\omega)|$ centered at $L_o(j\omega)$

In order to ensure $\Delta(s)$ cannot change number of encirclements:

Each disk of radius $|\Delta(j\omega)||L_o(j\omega)|$ centered at $L_o(j\omega)$ must not extend to -1 point



This can be ensured if:

$$\underbrace{|\Delta(j\omega)||L_o(j\omega)|}_{\text{Radius of Disk}} < \underbrace{|1+L_o(j\omega)|}_{\text{Distance from } -1 \text{ to center of disk}} \quad \text{for all } \omega \geq 0$$

Re-arranging:

$$\frac{|L_o(j\omega)|}{|1+L_o(j\omega)|} < |\Delta(j\omega)|^{-1} \quad \text{for all } \omega \geq 0$$

Note that

$$T_o(s) = \frac{L_o(s)}{1+L_o(s)} \quad \text{is the } \underline{\text{nominal}} \text{ CL TF}$$

So the required condition is:

$$\boxed{|T_o(j\omega)| < |\Delta(j\omega)|^{-1} \quad \text{for all } \omega \geq 0}$$

Uncertainty robustness test

Graphical Interpretation

The Bode magnitude plot $|T_o(j\omega)|$ must lie <u>below</u> the graph of $|\Delta(j\omega)|^{-1}$ <u>at every frequency</u>.

"Multiplictive" Uncertainty Robustness Test

with

$$\Delta(s) = \left[\frac{G(s)}{G_o(s)} - 1\right]$$

test is:

$$|T_o(j\omega)| < |\Delta(j\omega)|^{-1} \quad \text{for every } \omega$$

Guarantees closed-loop stability only. Performance will generally suffer

Given an assumed bound on magnitude $|\Delta(j\omega)|$

Note: Simultaneous gain/phase uncertainty easily handled in this framework. If plant gain uncertain and time delay present, then

$$\Delta(s) = \left[\frac{K_p}{K_o} e^{-sT_s} - 1\right]$$

Where K_p is true gain of plant, K_o is assumed gain, and T is delay length. Can graph $|\Delta(j\omega)|^{-1}$ given bounds on T and (K_p/K_o).

Note: Test is inherently conservative. If it fails, $T'(s)$ may be unstable, but not necessarily.

For example, with pure time delay uncertainty

$$|\Delta(j\omega)| = |e^{-j\omega T_s} - 1|$$

The above test yields predictions for T_{max} which are about 5-10% shorter than phase margin analysis gives

In this case, the phase margin analysis is exact.

Discrepancy with Δ test is because there exist $\Delta(s)$ with the same magnitude bound as $|e^{-j\omega T_s} - 1|$ which would result in an unstable $T'(s)$. However, these $\Delta(s)$ would include other terms than pure delay.

But only Δ test lets us look at impact of <u>simultaneous</u> gain/phase changes, including effects of

⇒ <u>uncertain</u> pole/zero locations in $G(s)$

⇒ <u>neglected</u> pole/zero locations in $G(s)$

<u>Typically</u>: $|\Delta(j\omega)|$ is small at low frequencies, increases at higher freqs.

⇒ Effects of model errors on freq. response accumulate as freq. increases

<u>Then</u>: Bound on $|T_0(j\omega)|$ is large at low freqs, small at high freqs.

<u>Example</u>: Suppose $G_0(s)$ neglects a pole in $G(s)$, but is otherwise identical:

Then: $\Delta(s) = \left[\frac{1}{\tau s+1} - 1\right] = \frac{-\tau s}{\tau s+1} \Rightarrow \Delta^{-1}(s) = \frac{\tau s+1}{-\tau s}$

Now look at "typical" shapes for $|T_o(j\omega)|$

$$T_o(s) = \frac{L(s)}{1+L(s)}, \quad |T_o(j\omega)| = \frac{|L_o(j\omega)|}{|1+L_o(j\omega)|}$$

Typically, $|L_o(j\omega)| \gg 1$ for small ω (especially if $L_o(s)$ has at least 1 pole at origin)

$\Rightarrow |T_o(j\omega)| \approx 1$ (0 dB) for small ω.

Since relative degree of $L_o(s)$ is positive for any physical system, $|L_o(j\omega)| \to \emptyset$ as $\omega \to \infty$, and thus

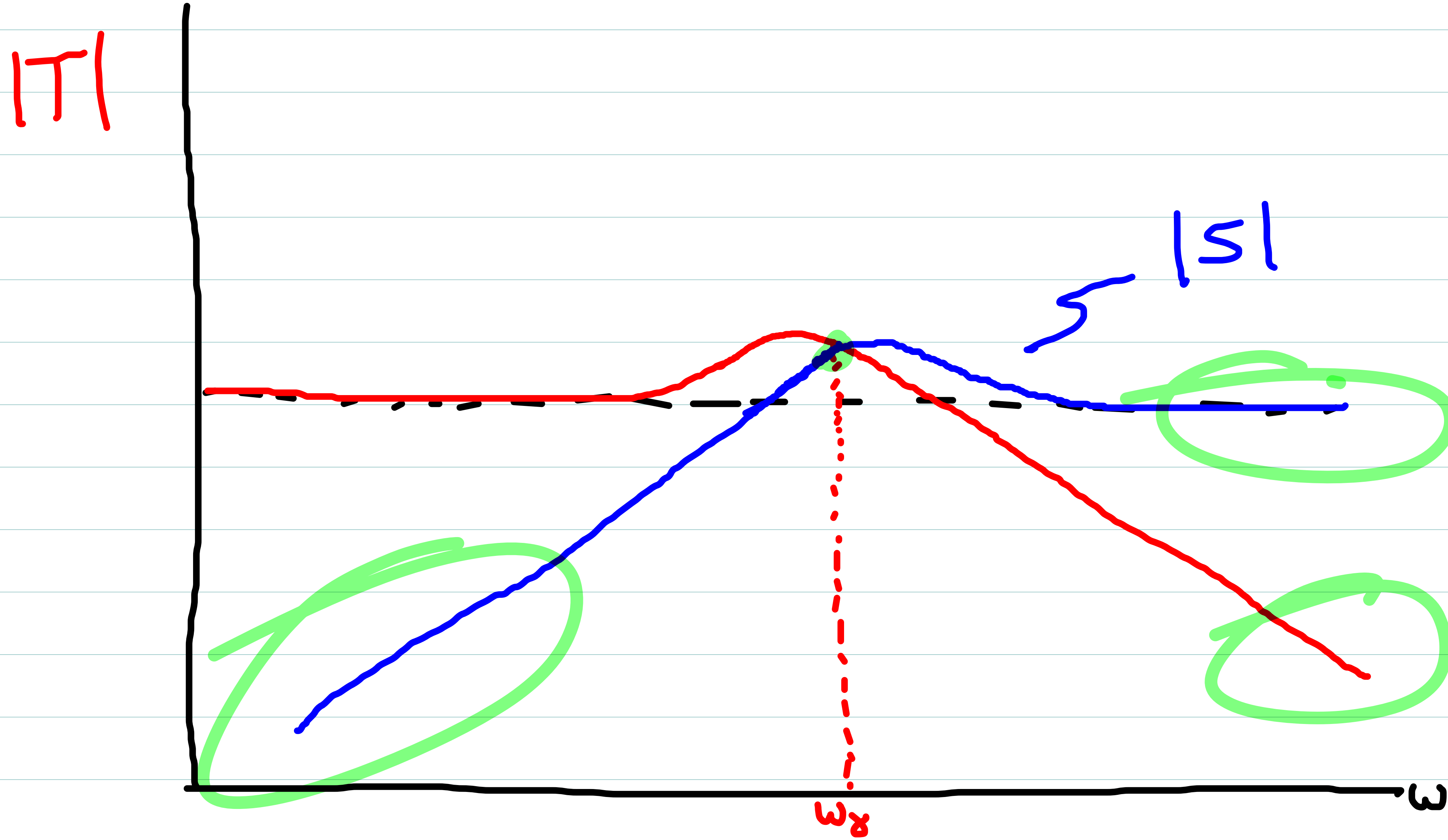
$|T_o(j\omega)| \approx |L_o(j\omega)|$ at high freq. and $|T_o(j\omega)| \to \emptyset$ also

Finally, note $$|T_o(j\omega_\gamma)| = \frac{|L_o(j\omega_\gamma)|}{|1+L_o(j\omega_\gamma)|} = \frac{1}{|1+L_o(j\omega_\gamma)|}$$

So $$|T_o(j\omega_\gamma)| = |S(j\omega_\gamma)| = \frac{1}{2\sin(\gamma/2)}$$

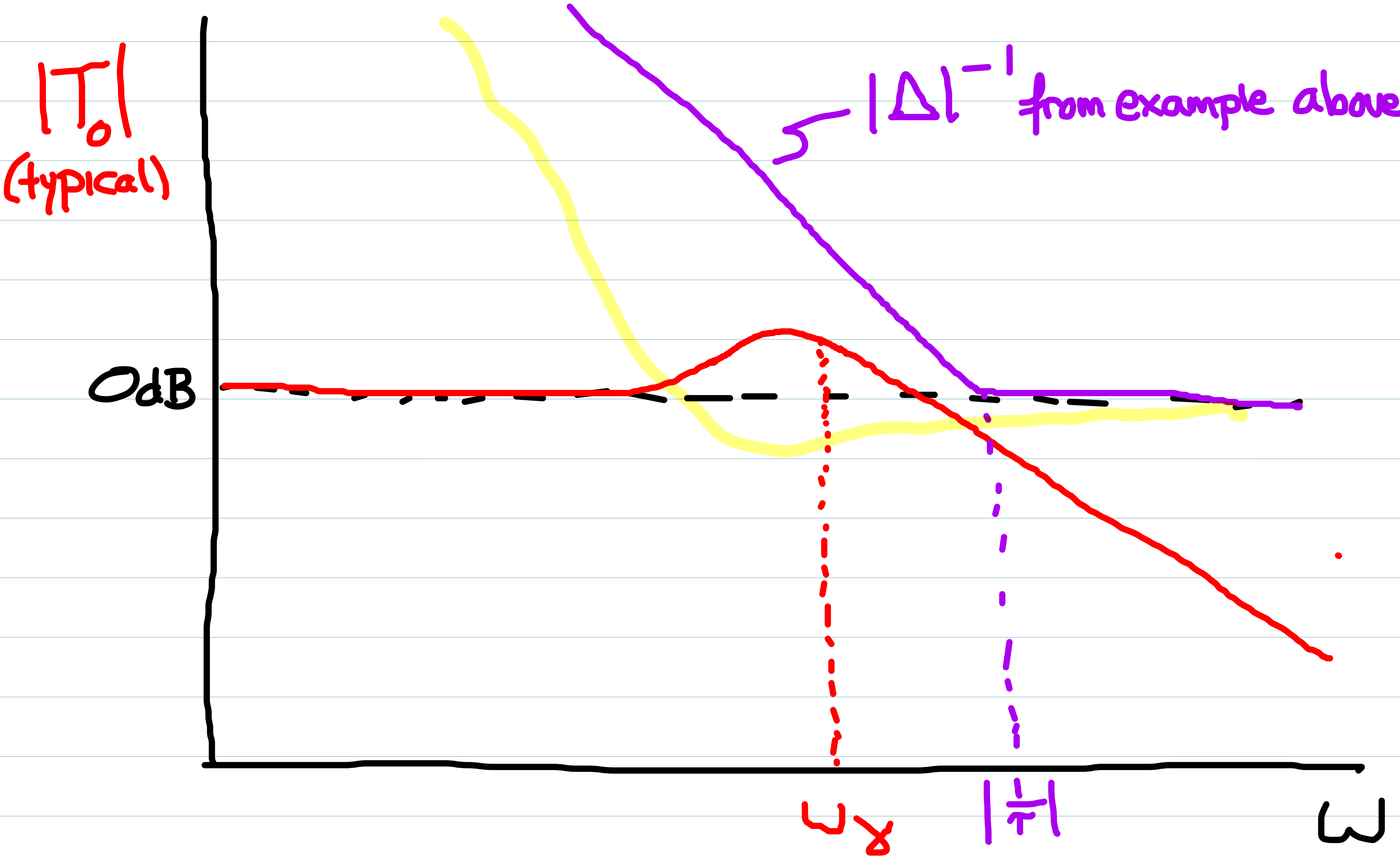
hence $|T_o|$ is also peaking near ω_γ.

$|T_0|$ (typical)
0dB
$\left|\frac{1}{2\sin(\gamma/2)}\right|_{dB}$
$\approx |L_0|$
ω_γ
ω



<u>Note</u>: $|T_o|$ and $|S_o|$ "complementary" in sense that $|S_o| \approx \phi$ when $|T_o| \approx 1$ and vice-versa.

Reflects algebraic <u>identity</u> $\boxed{S(s) + T(s) = 1}$ from def'ns.



Remember: must keep graph of $|T_o(j\omega)|$ <u>below</u> $|\Delta(j\omega)|^{-1}$ at every frequency

Design Implication of robustness

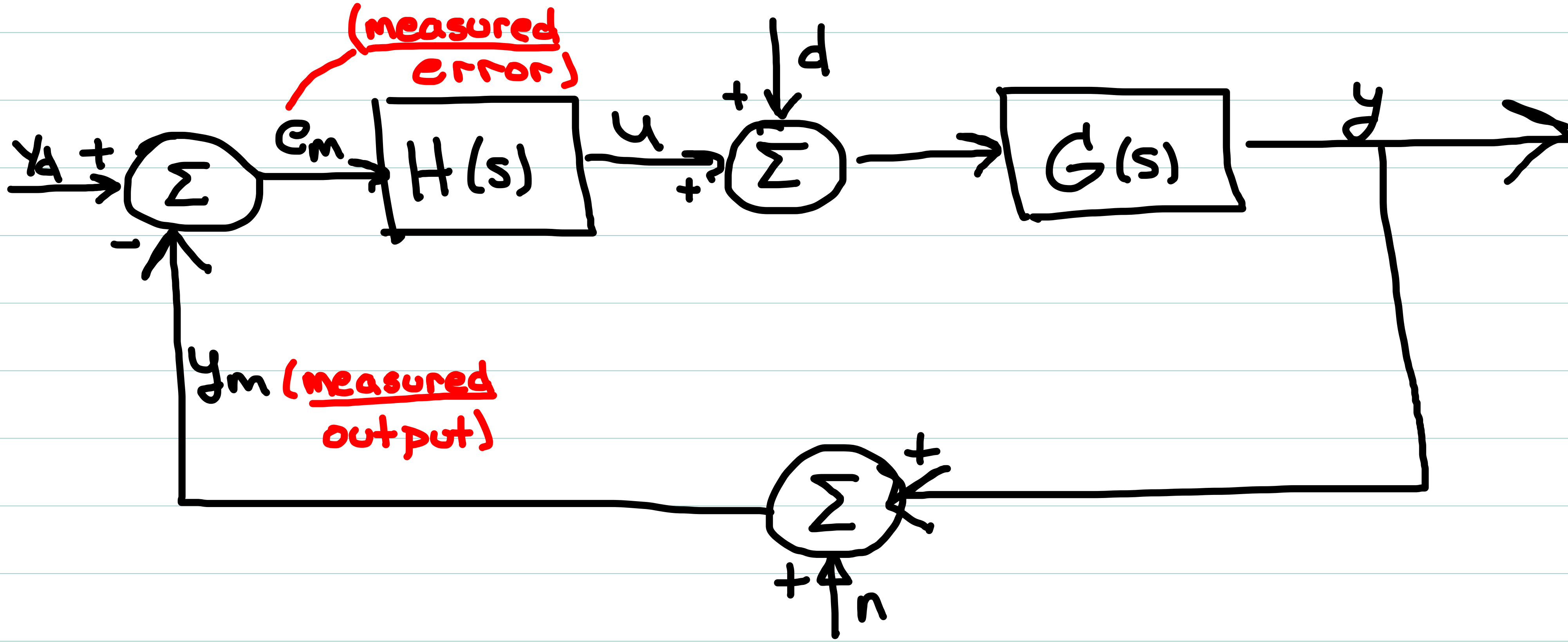
Uncertainty constrains size of ω_γ!

In specific example above, we'd need ω_γ significantly less than freq. $(\frac{1}{\tau})$ of neglected pole.

When $G(s)$ has "unmodeled dynamics" (i.e. poles/zeros neglected in nominal model $G_0(s)$), usually want ω_γ a decade below suspected freq. of neglected poles.

Recall, ω_γ is correlated w/ closed-loop settling time. Above observation means this should be slow compared to neglected poles. We need to avoid control actions so sharp and quick they might "excite" the unmodeled dynamics.

Effect of sensor noise



Now: $Y = G[U+D]$, $U = HE_m = H[Y_d - (Y+N)]$

So: $Y = GHY_d - GHY + GD - GHN$

Or: $\boxed{Y = TY_d - S_i D - TN}$ ← Bad

$$Y_d \longrightarrow \boxed{T} \longrightarrow Y$$

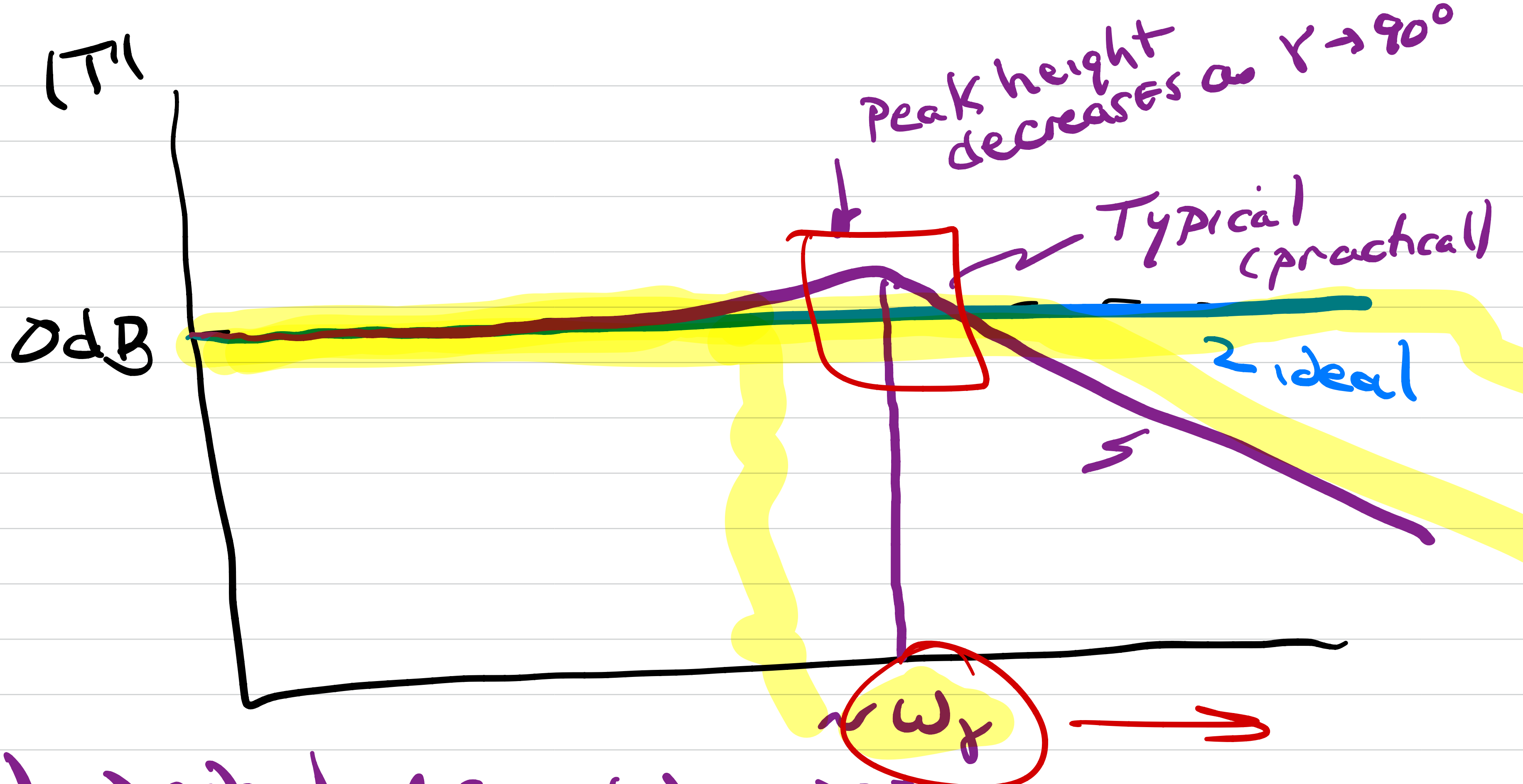
ideally (for tracking) $\boxed{Y = Y_d}$

$\Rightarrow$ ideally, $T(s) = 1$

$\Rightarrow |T(j\omega)| = 0\ \mathrm{dB}$

$\angle T(j\omega) = 0\ \mathrm{deg}$

for all $\omega \geq 0$



Typical → ideal As: $\omega_\gamma \to \infty$
$\gamma \geq 60°$

But $\omega_\gamma \to \infty$ means
- infinite phase loss from delay
- No robustness to model uncertainty
- impractically large $u(t)$
- high Noise sensitivity

and hence: $E = Y_d - Y$ satisfies:

$$\boxed{E = (1-T)Y_d - S_i D + TN}$$

New term!

or: $E = \underbrace{S Y_d}_{\text{Tracking error}} - \underbrace{S_i D}_{\text{error due to disturbance}} + \underbrace{TN}_{\text{Add'l error due to noise}}$

<u>Note</u>: TF from noise to Y is <u>same</u> as TF from Y_d to Y (both are $T(s)$)

<u>Implication</u>: $\Rightarrow$ feedback loop tries to "track the noise"

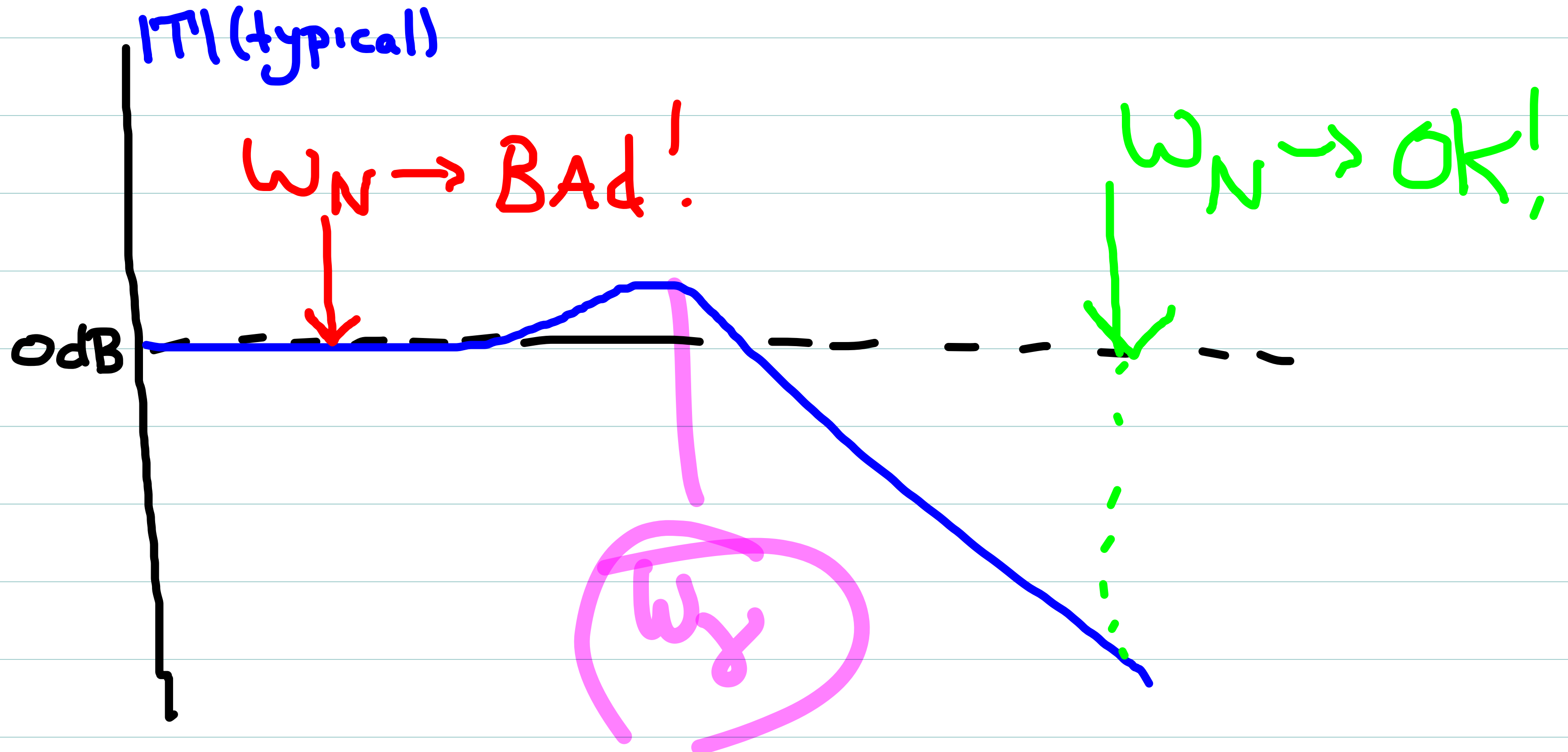
<u>Equivalently</u>: $\Rightarrow$ noise is indistinguishable from "signal" $y(t)$ loop is trying to control!

Impact of Noise

Assume for simplicity noise is "tonal": $n(t) = N\sin(\omega_N t)$
(it isn't really, but useful starting point!)

Then Added error is upper bounded by $N|T(j\omega_N)|$

$\Rightarrow$ Need $|T(j\omega)|$ small at noise frequency ω_N!

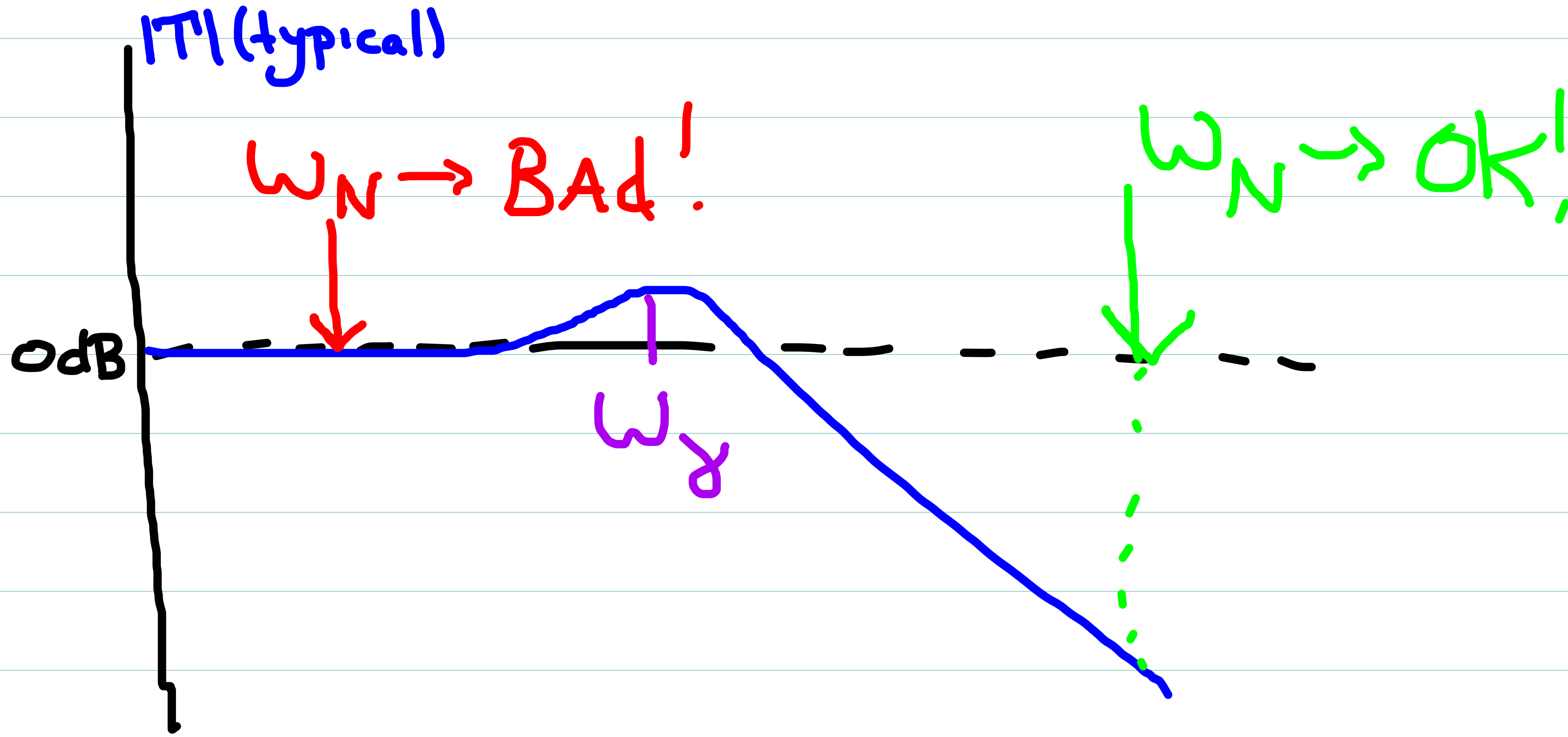


Impact of Noise

Assume for simplicity noise is "tonal": $n(t) = N \sin(\omega_N t)$
(it isn't really, but useful starting point!)

Then Added error is upper bounded by $\boxed{N|T(j\omega_N)|}$

$\Rightarrow$ Need $|T(j\omega)|$ small at noise frequencies!



Design Implications, I

$\Rightarrow$ Need $\omega_\gamma << \omega_N$

$\Rightarrow$ Constrains ω_γ / bandwidth

$\Rightarrow$ Conversely, designs with larger ω_γ will show worse performance due to increased noise impact!

Essentially, we need to make sure there is adequate separation between the frequencies we are trying to track (bandwidth), and the frequency of the noise.

$\Rightarrow$ Works against our desire for large ω_γ (fast settling)

Another perspective:

With noise, controller implementation equation is:

$$u(t) = c_0 e_m(t) + \sum c_k x_k(t)$$

$$\dot{x}_k(t) = a_k x_k(t) + e_m(t) \qquad [a_k \text{ poles of } H(s)]$$

Noise impacts $u(t)$:

⇒ directly if $c_0 \neq \emptyset$

⇒ indirectly through $x_k(t)$

$x_k(t)$ diff'l eq'ns have a "filtering" property (reduce magnitude of noise effects)

⇒ Designs with $c_0 = \emptyset$ have superior noise resistance

Design Implications, II

$C_0 = \emptyset \iff H(s)$ has more poles than zeros

$\Rightarrow$ Designs with this property have better noise resistance!

$\Rightarrow$ Works against our need to increase phase margin

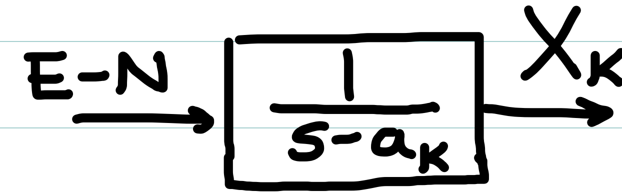
Most "Advanced" controller designs have 1 more pole than zeros to ensure good noise filtering.

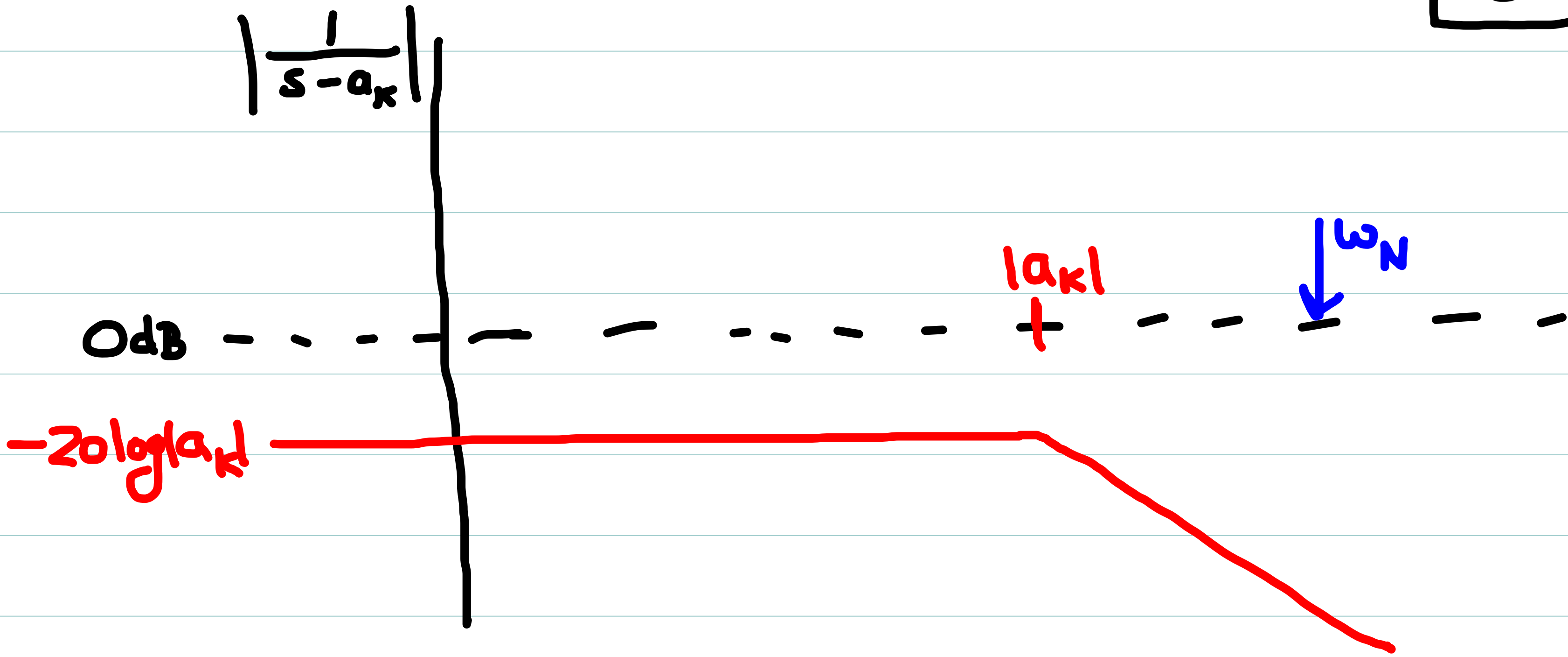
However, superior transient performance is achievable with $C_0 \neq \emptyset$ provided noise is not a significant issue.

"Filtering" by $x_K(t)$ States

$$\dot{x}_K(t) = a_K x_K(t) + e_m = a_K x_K(t) + \underbrace{e(t)}_{\text{true error}} - \underbrace{n(t)}_{\text{sensor noise}}$$

$$\Rightarrow X_K(s) = \left[\frac{1}{s+a_K}\right]\left[E(s) - N(s)\right]$$





Noise is attenuated in $x_K(t)$ if $\underbrace{|a_K|}_{\text{Compensator pole}} \ll \underbrace{\omega_N}_{\text{noise frequency}}$.

Design implication, III

For good noise rejection, compensator poles should be significantly lower frequency than the noise

⇒ Avoid excessively high frequency poles in $H(s)$ (i.e. poles very far from imag axis).

⇒ Another advantage of "minimum β" lead comp design:

By minimizing β (ratio of pole location to zero location in $H(s)$), we are bringing the pole as close to imag axis as possible while still providing necessary φ_{req} at desired ω_x.

Why it's bad to differentiate $y(t)$.

One is tempted to implement a $H(s)$ with only a zero (or more generally with 1 more zero than pole) by numerically differentiating $y(t)$ $\approx \dot{y}(t)$

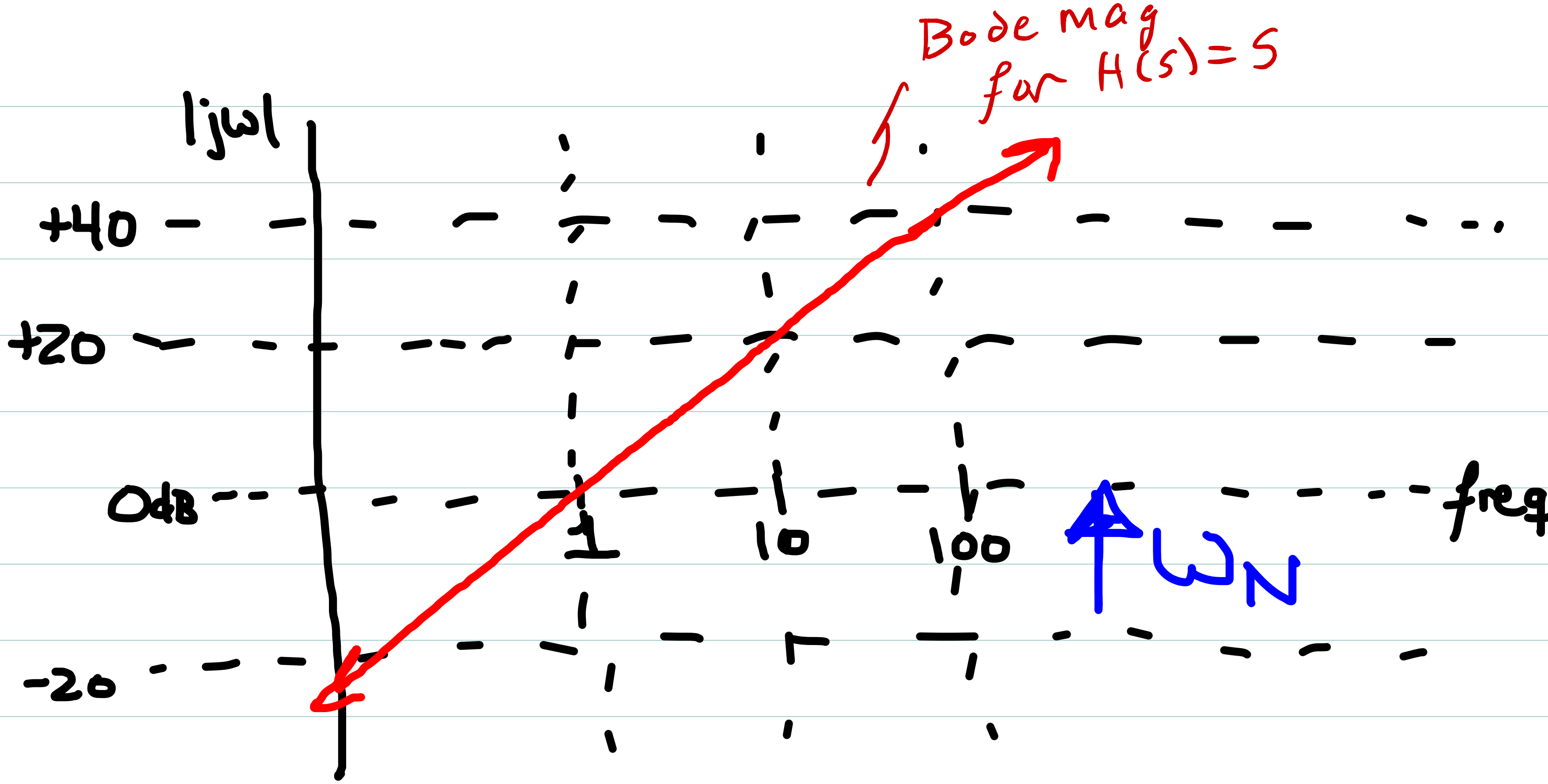
This would be needed since, as we've seen, such compensators will result in $u(t)$ having a term proportional to $\dot{e}(t)$ [hence $\dot{y}(t)$]

But with noise, we're really diff'ing $y_m(t) = y(t) + n(t)$.

Let $z(t) = \frac{d}{dt} y_m(t)$ be an estimate of $\dot{y}(t)$

$$\Rightarrow Z(s) = s\left[Y(s) + N(s)\right] \qquad \xrightarrow{Y+N} \boxed{s} \xrightarrow{Z}$$

Impact of noise depends on freq. response of s.



Differentiation amplifies the effect of noise

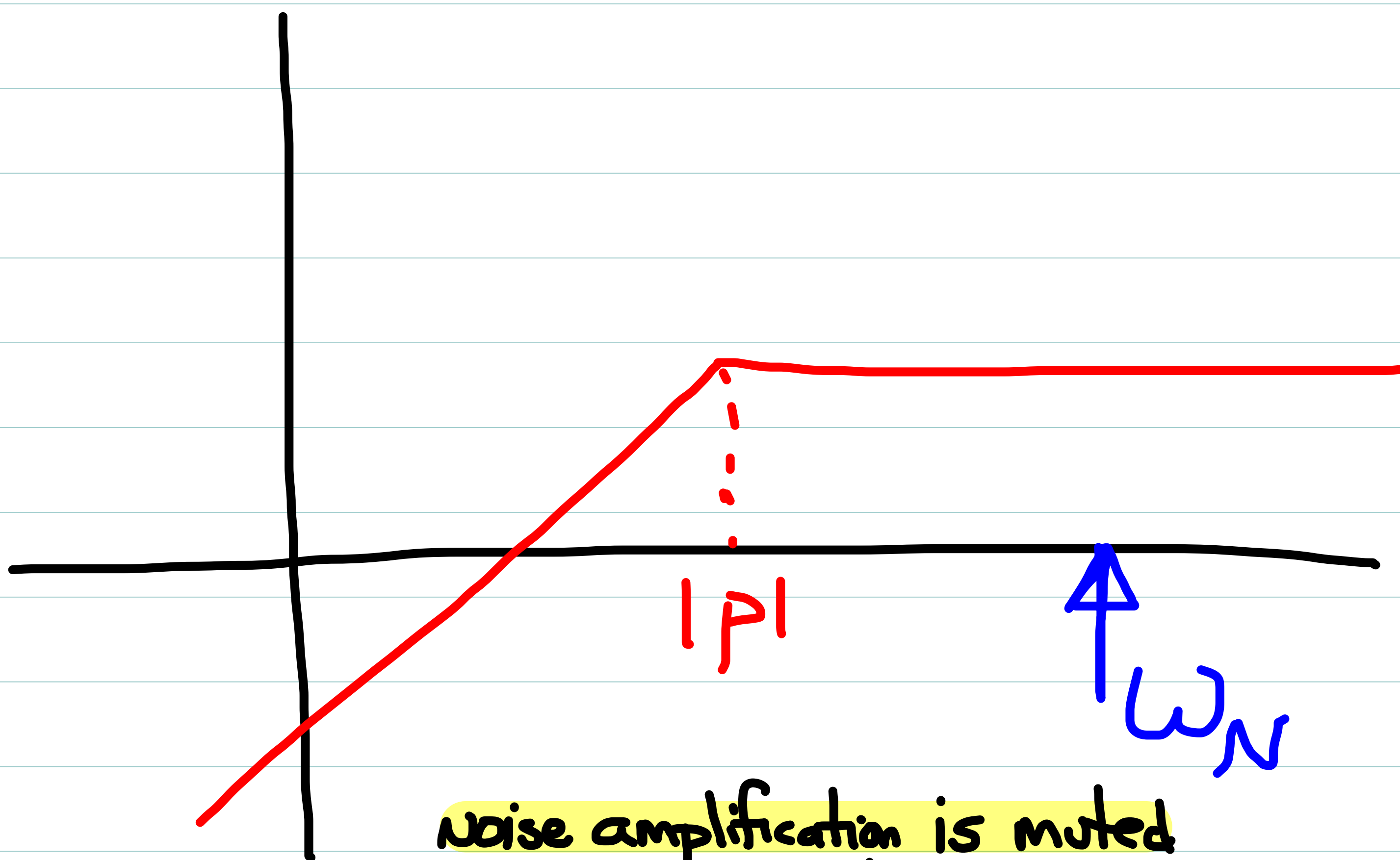
explicitly: if again $n(t) = \varepsilon \sin(\omega_N t)$, $\omega_N \gg 1$
then

$$z(t) = \frac{d}{dt}\left[y(t) + n(t)\right] = \dot{y}(t) + \underbrace{\varepsilon \omega_N}_{\text{Not small!}} \cos(\omega_N t)$$

(potentially larger than $\dot{y}$)

Note that if we added a pole to our derivative estimation scheme

$$Z(s) = \left[\frac{s}{s-p}\right] Y_m(s)$$



Noise amplification is muted and may be tolerable.

If we used this strategy to replace the derivative information needed for implementation an ideal zero:

$$H(s) = K(s-z) \Longrightarrow H(s) = K\left[\frac{s}{s-p} - z\right]$$

Then:

$$H(s) = K\left[\frac{(1-z)s + pz}{s-p}\right]$$

which is a <u>lead compensator</u> (for typical case $p<z$).

So really, a lead compensator is effectively a "practical" implementation of an ideal zero, which acknowledges the imperfect nature of the measurement process.

Alt: a lead comp is a PD with velocity measurements replaced by a low pass filtered <u>estimate</u> of velocity.

The most basic (and essential) task of the control engineer — achieving a stable closed-loop system with nominal performance characteristics — is straightforward to approach.

However, it is tricky to also incorporate and balance the competing constraints of

- Implementation constraints (relative degree of $H(s)$)
- Tracking accuracy
- Disturbance rejection
- Noise rejection
- Model uncertainty
- Sensor/Actuator/Computation delays
- Actuator Limits/Control saturation
- Power/weight/cost demands

The "best" design is one which achieves an acceptable trade-off among these competing factors.

There is no "one true design" which makes the "ideal" tradeoff — so don't waste time looking for it!

Find something that works acceptably well, and move on

Major, common families of Compensators

① $H(s) = K \Rightarrow u(t) = K e(t)$ "Proportional" control

② $H(s) = K_P + K_D s = K(s-z)$ $(K = K_D,\ z = -K_P/K_D)$

$\Rightarrow u(t) = K_P e(t) + K_D \dot{e}(t)$ "Prop. + Derivative (PD) control"

Note: implementable if both $y(t)$ and $\dot{y}(t)$ measured directly)

③ $H(s) = K_P + \frac{K_I}{s} = K\left[\frac{s-z}{s}\right]$ $(K = K_P,\ z = -K_I/K_P)$

$$\Rightarrow u(t) = K_P e(t) + K_I x_1(t)$$
$$\dot{x}_1(t) = e(t)$$

Equivalently: $u(t) = K_P e(t) + K_I \int_0^t e(\tau)\,d\tau$

"prop. + integral (PI) control")

4. $H(s) = K_P + K_D s + K_I/s = K\left[\frac{(s-z_1)(s-z_2)}{s}\right]$

$\left(K = K_D;\ z_1, z_2 \text{ roots of } K_D s^2 + K_P s + K_I\right)$

$\Rightarrow u(t) = K_P e(t) + K_D \dot{e}(t) + K_I \int_0^t e(\tau)\, d\tau$

"Prop/Int/Deriv (PID) control"

Notes: a.) Very popular. Special purpose chips which do this computation are commonly available

b.) 1)-3) above are special cases of this more general form.

c.) Provides 2 zeros to help meet margin/xover requirements, and pole at origin to help with tracking/dist. rejection requirements.

d.) Like PD, requires direct measurement of $\dot{y}(t)$

⑤ $H(s) = K\left[\frac{(s-z)}{s-p}\right]$, $|z|<|p|$

"Lead compensator"

Notes: a.) "Implementable" form of PD control when only $y(t)$ measured

b.) Using minimal values of $\beta = |p|/|z|$ helps with noise rejection and control saturation

⑥ $H(s) = K\left[\frac{(s-z_1)(s-z_2)}{s(s-p)}\right]$ $|p|>|z_1|,|z_2|$

"PI/Lead": "implementable" form of PID when only $y(t)$ measured

---

Of course, a designer is free to choose $H(s)$ as desired. These are common "go to" starting points which can be modified or added to as needed.

Case Study

Consider system from HW #8: $G(s) = \frac{1}{10s^2(s+1)}$

We had 2 designs

Case 1: "Low bandwidth" $\omega_r = 0.1,\ \gamma = 45^\circ$

$$H_1(s) = \frac{0.036(28s+1)}{(3.57s+1)}$$

Case 2: "Higher bandwidth" $\omega_r = 1,\ \gamma = 45^\circ$

$$H_2(s) = \frac{2.426(2.41s+1)^2}{(.41s+1)^2}$$

Step response characteristics:

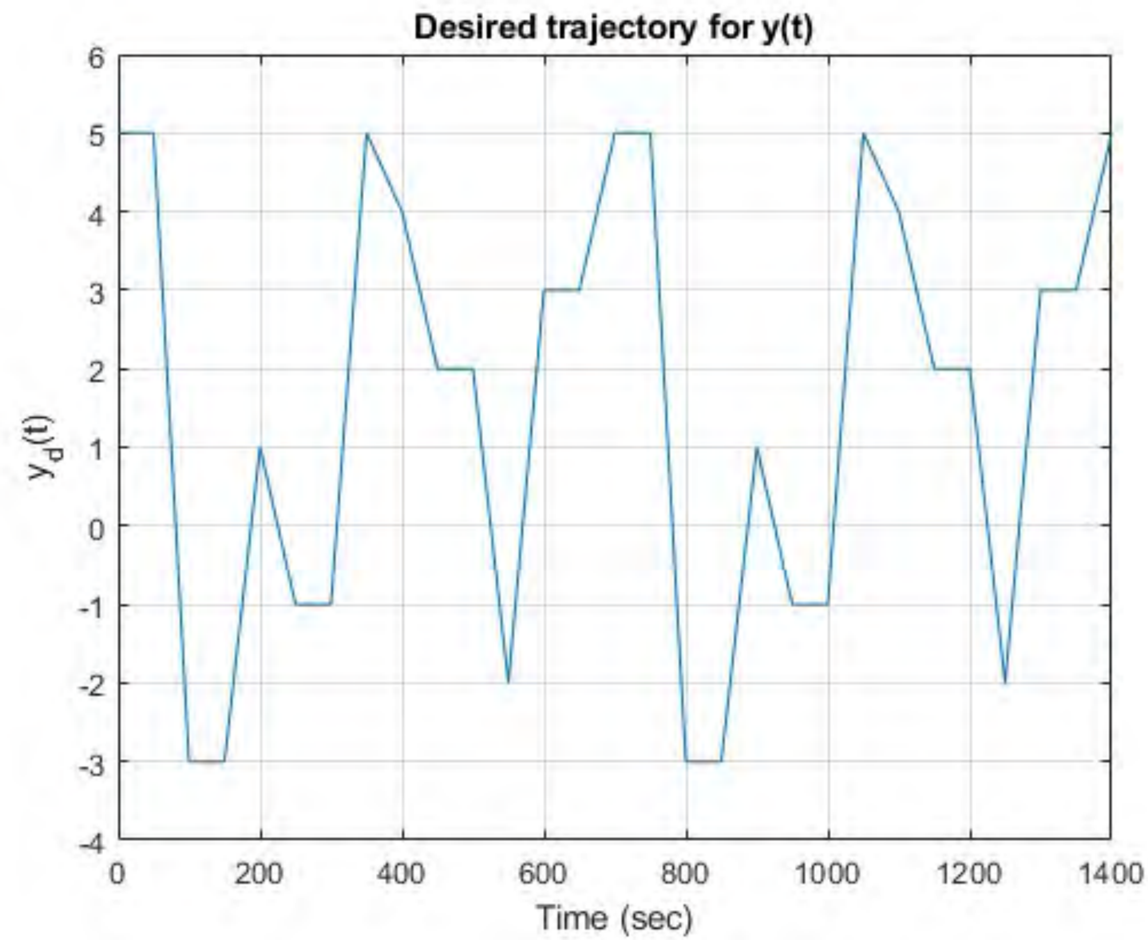
Case 1 gives about $t_s = 67.2$ sec, $\%OS = 33\%$

Case 2 gives about $t_s = 10.2$ sec, $\%OS = 34\%$

$e_{ss}(t) = 0$ in both cases (since $L(s)$ has 2 poles at 0)

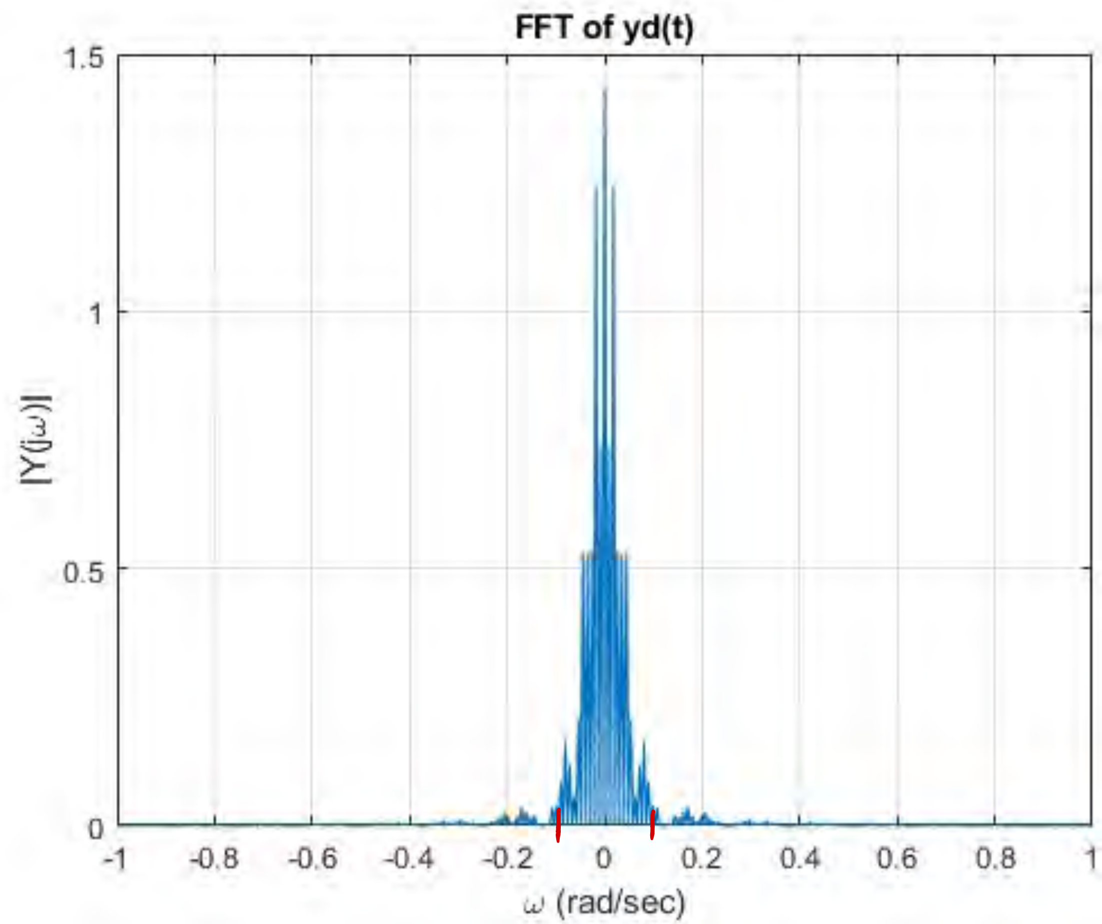
More complicated $y_d(t)$

Suppose instead $y_d(t)$ looks like:



Perhaps representative of a path followed by a drone navigating a cluttered environment.

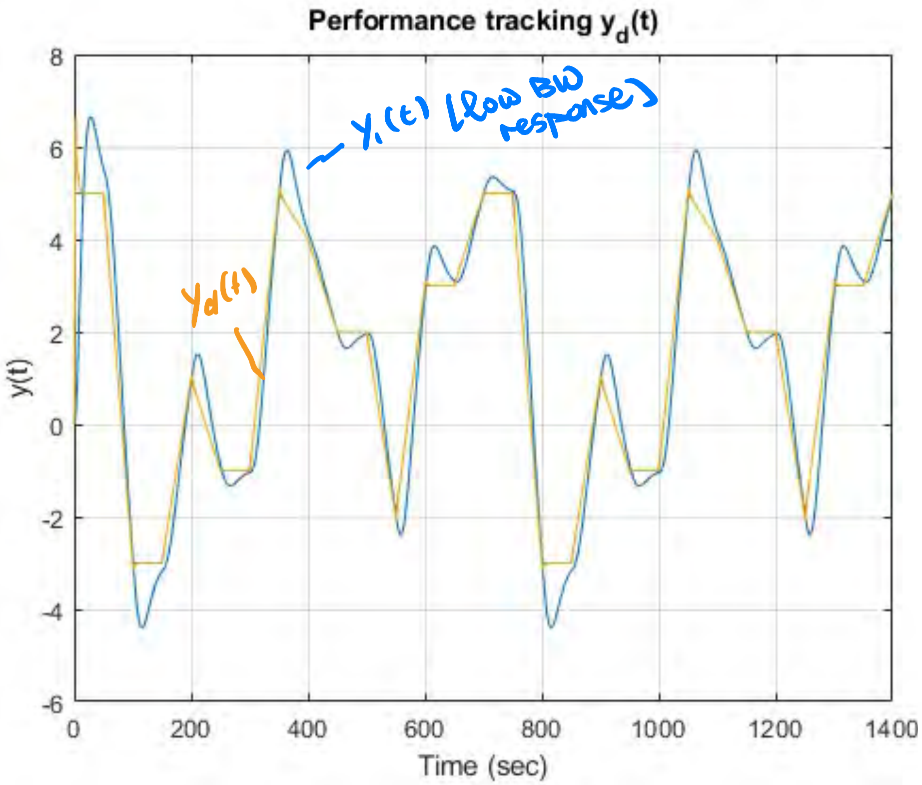
Frequency content of $y_d(t)$



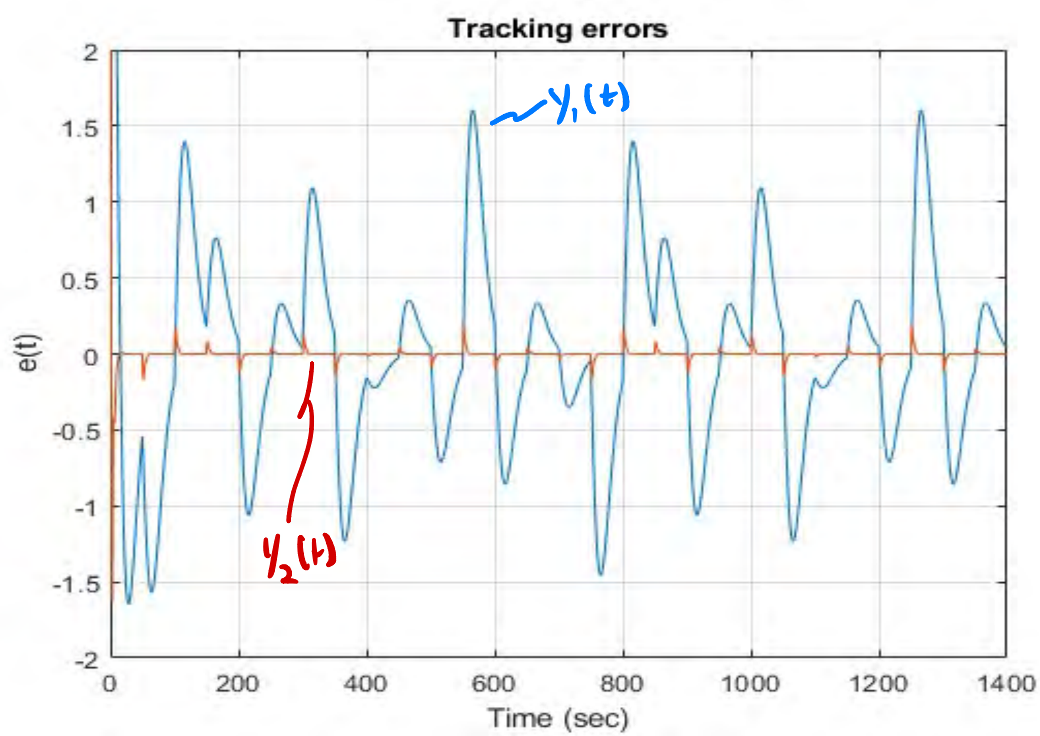
Note that $y_d(t)$ has significant freq. content near and below 0.1 rad/sec, but almost none near or above 0.5 rad/sec

$\Rightarrow$ Expect low BW ($\omega_r = 0.1$) design to have poor tracking, But high BW ($\omega_r = 1$) design to have good tracking.

Tracking performance

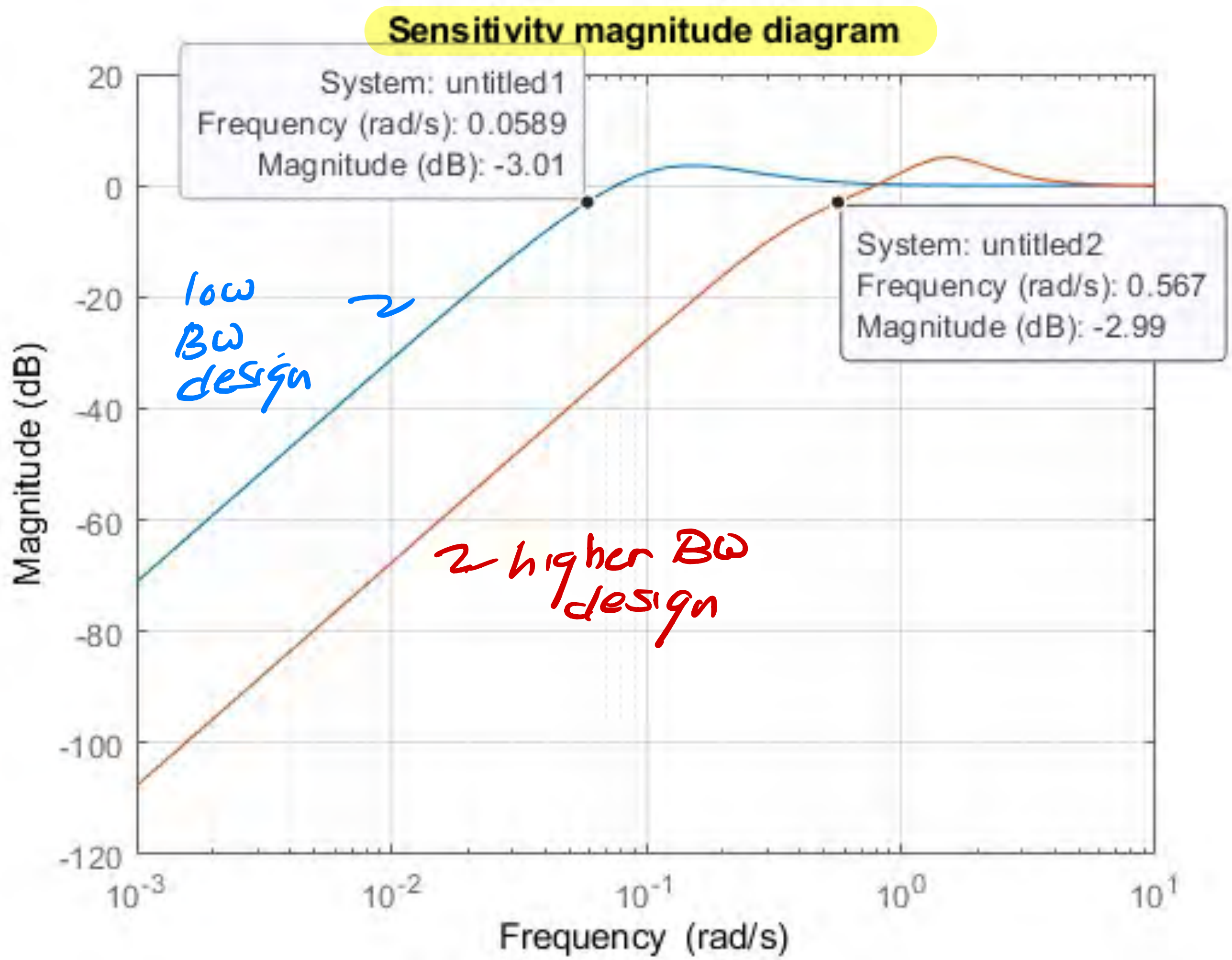


High BW response $y_2(t)$ almost identical to $y_d(t)$ at this scale



High BW response has atleast 10 times less tracking error!

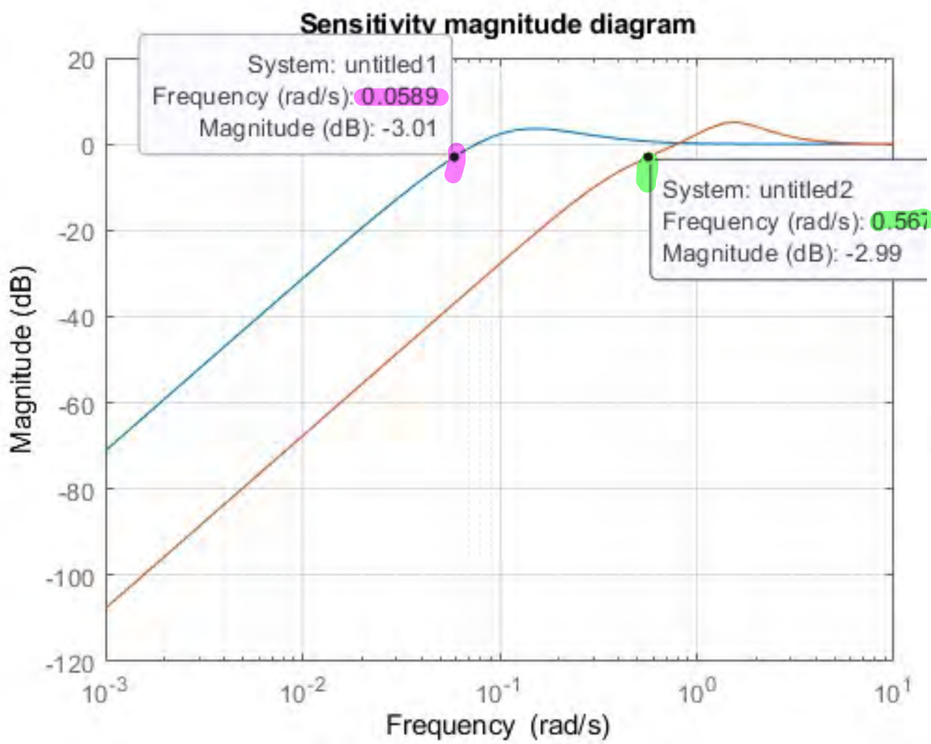
Importance of "bandwidth"

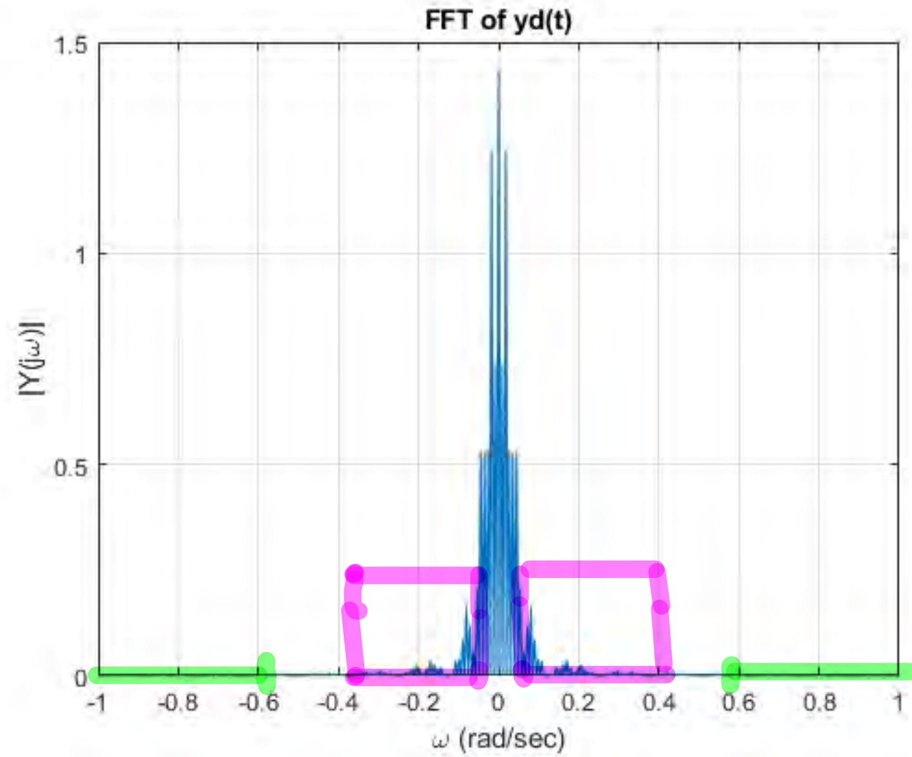


Recall $\quad \omega_\gamma/2.5 \le \omega_B \le \omega_\gamma \quad$ typically

here $\quad \omega_B \approx \omega_\gamma/1.7$ in both cases

Importance of "bandwidth"



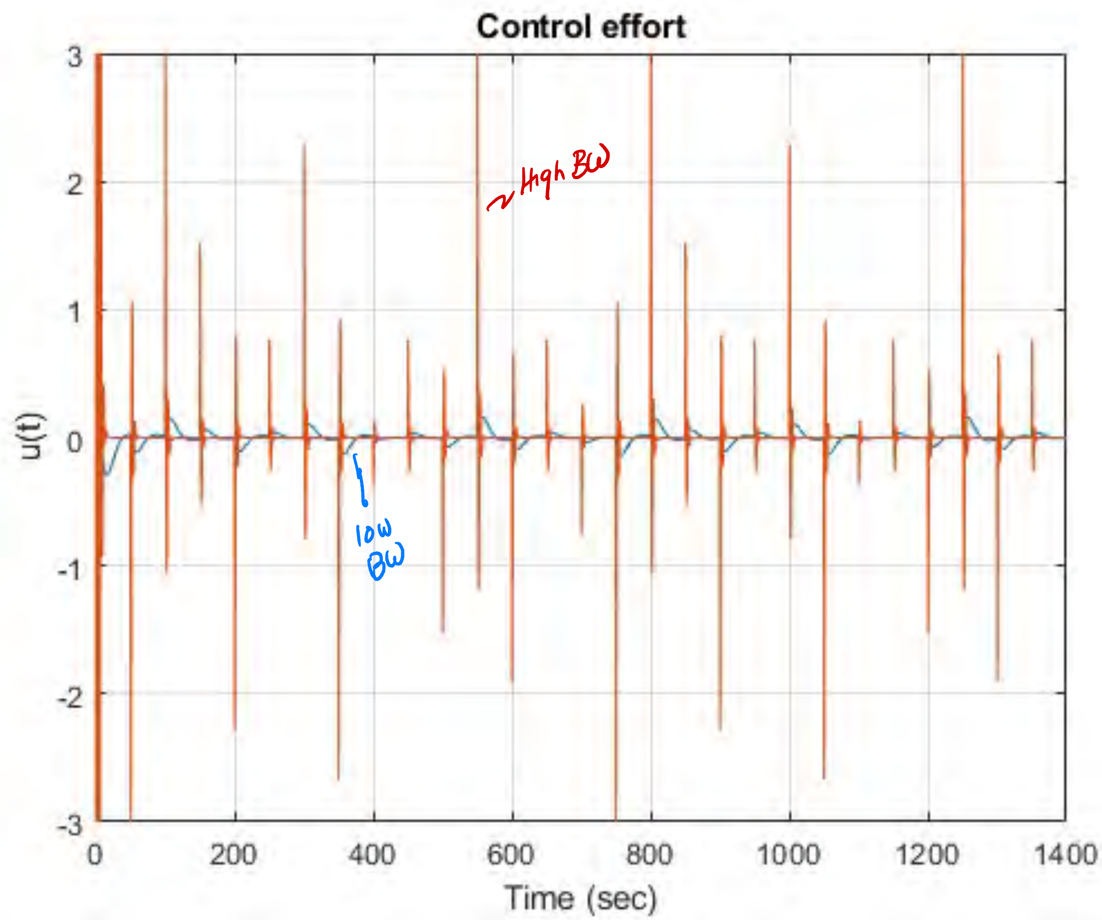


yd(t) has significant freq content between 0.06 and 0.4 rad/sec which low BW design fails to track well

However, yd(t) has very little freq content above 0.6 rad/sec, so high BW design does a good job tracking

(Remember: need $|S(j\omega)| \ll 0$ dB to track freq ω with minimal steady-state error)

Control Effort:

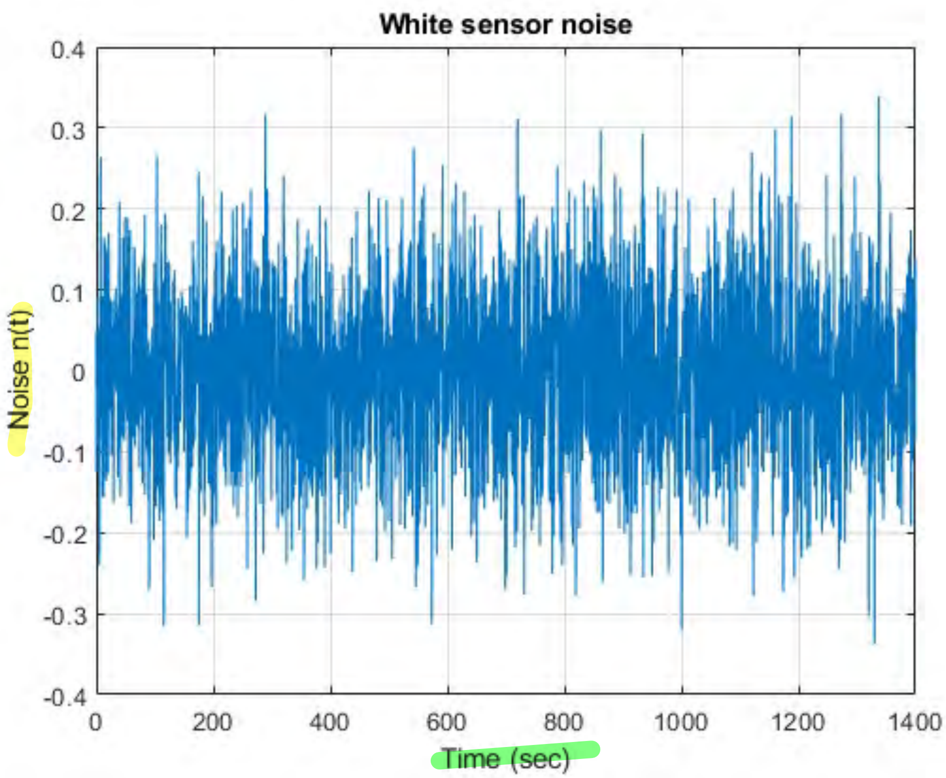


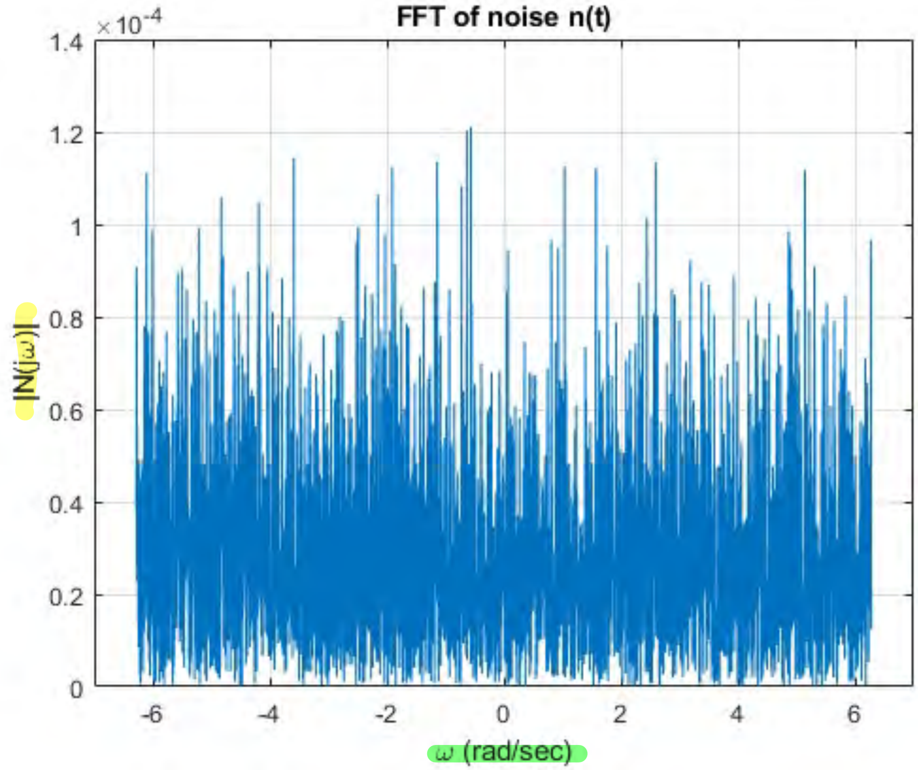
Of couse, significantly more force/moment must be applied to the vehicle to make it accurately execute those sharp corners in $y_d(t)$.

High BW inputs $u(t)$ about 20x larger than low BW.

Noise Sensitivity

Consider a "white" (extremely broadband) noise corrupting sensor measurements (worst case model): $y_m(t) = y(t) + n(t)$





Significant amplitudes across all frequencies.

Remember that each freq. in $n(t)$ will show up in $y(t)$ with amplitude multiplied by $|T(j\omega)|$

Recall: $Y(s) = T(s)Y_d(s) + S_i(s)D(s) + T(s)N(s)$

look at this term

If $n(t) = \sum_k A_k \sin(\omega_k t)$ (i.e. many diff't sinusoids in $n(t)$)

then $y(t)$ will contain

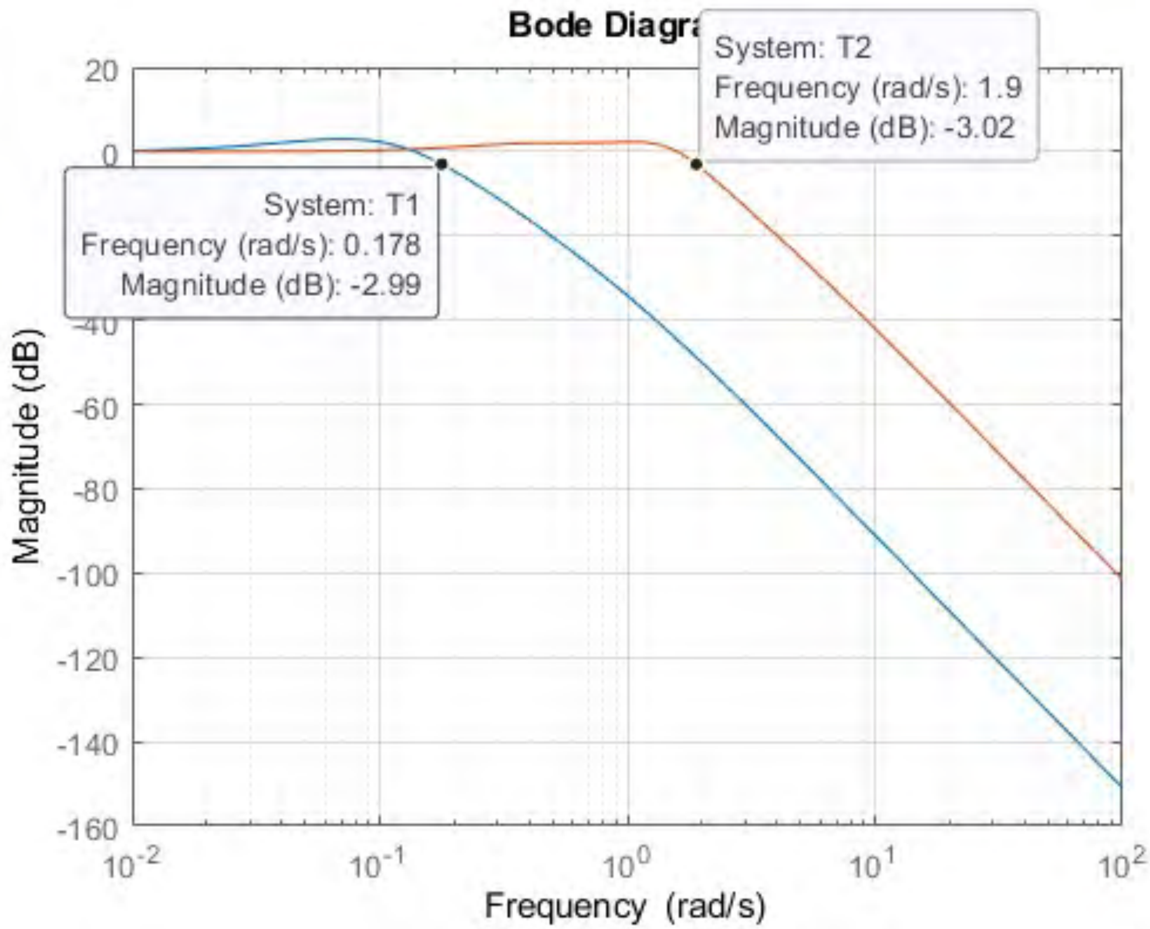
$$\sum_k A_k |T(j\omega_k)| \sin(\omega_k t + \angle T(j\omega_k))$$ (by linearity)

which will be non-negligible wherever $|T(j\omega)| \approx 1$

Want $|T(j\omega)| << 0\,dB$ to "reject" noise at freq ω

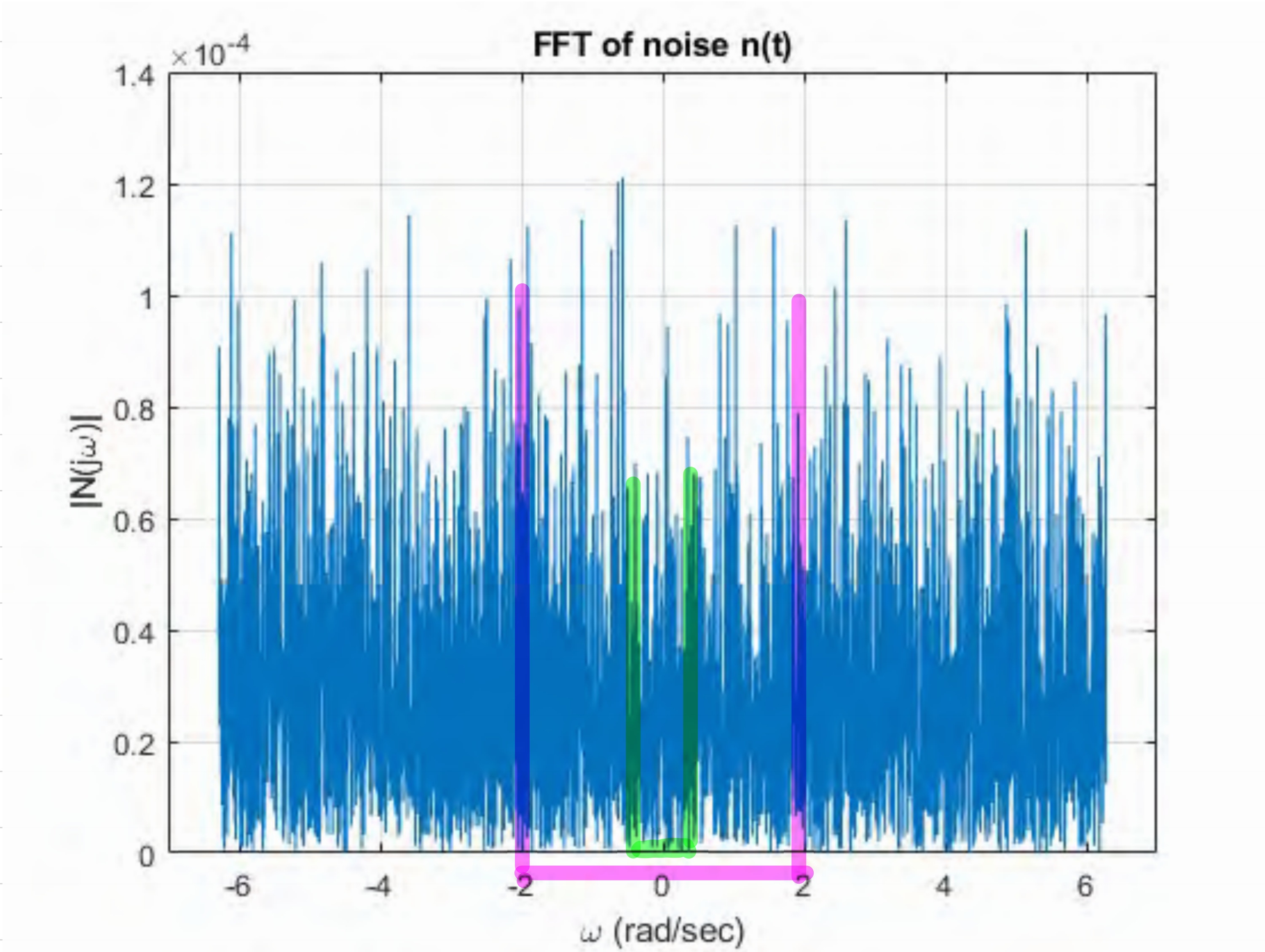
If $|T(j\omega)| \simeq 1$ at noise freqs ω_k, those components of noise will be "PASSED through" unfiltered to $y(t)$ creating significant unwanted motion.

CL Bode magnitudes



High BW design will "pass through" ($T(j\omega) \approx 1$) a larger range of freqs (up to ~2 rad/sec) compared to low BW design (which passes only up to ~0.2 rad/sec)

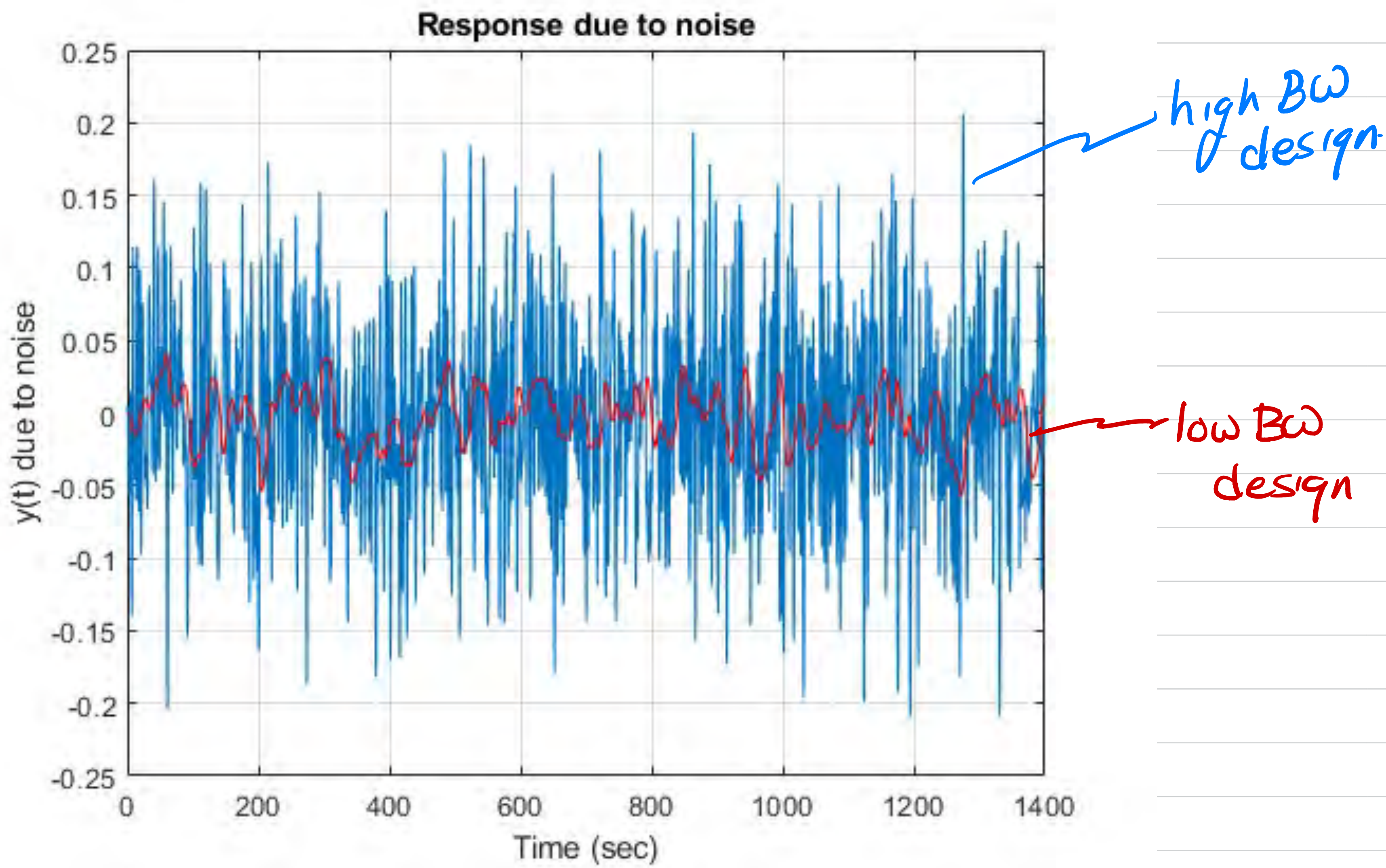
$\Rightarrow$ More effect of noise on high BW design



freqs passed through by low BW T(s)

freqs passed through by high BW T(s)

Noise effect on y(t)



Higher BW design shows 3x greater impact of Noise on y(t).

Alternate Design Perspectives

Our correlation between phase margin/crossover and the poles of $T(s)$ [hence its transient response characteristics] is approximate and tenuous at best.

It would be nice if we could specifically target the desired closed-loop poles, and design $H(s)$ to obtain them.

There are, in fact, techniques for this, although in using them we give up many of the insights afforded by the freq. response design methods . . .

(Everything is a trade-off! There are no magic bullets in this game!)

Recall the Characteristic Equation:

All closed-loop poles satisfy: $1 + L(s) = \emptyset$

$$\boxed{\Rightarrow L(s) = -1}$$

Let $L_o(s) = \left[L(s)\right]_{K=1}$ (K = compensator gain – real!)

Then s is a CL pole if: $K L_o(s) = -1$

which requires $\boxed{L_o(s) \text{ to be real.}}$

For any such s: $K = \dfrac{-1}{L_o(s)}$

is the gain which would make this s a CL pole

Moreover: $L_0(s)$ is real (hence s a possible CL pole)

if: $\measuredangle L_0(s) = \ell(180°)$ $\quad$ (ℓ = any integer)

In particular:

$$L_0(s) = (1+2\ell)180° \quad \text{(odd multiple of } 180°)$$

$\Rightarrow$ $L_0(s)$ is a <u>negative</u> real number

$\Rightarrow$ corresponding $K = \dfrac{-1}{L_0(s)}$ is positive

and:

$$L_0(s) = 2\ell(180°) \quad \text{(even multiple of } 180°)$$

$\Rightarrow$ corresponding K is <u>negative</u>

"Angle condition", $K > 0$

If we restrict ourself initially to $K > 0$, we need

$$\boxed{\measuredangle L(s) = \measuredangle L_0(s) = (1+2\ell)180^\circ}$$

for s to be a CL pole. This is the "angle condition".

Any value of s satisfying this condition will be a CL pole for an appropriate positive value of K.

Suppose that we want a specific CL pole, s_{Des}.

We need $\measuredangle L(s_{Des}) = (1+2\ell)180^\circ$

But recall: $\measuredangle L(s) = \measuredangle G(s) + \measuredangle H(s)$ for any $s \in \mathbb{C}$

Thus, to make s_{Des} a CL pole we need

$$(1+2\ell)180^\circ = \measuredangle G(s_{Des}) + \measuredangle H(s_{Des})$$

Hence, must design compensator $H(s)$ so that:

$$\boxed{\measuredangle H(s_{Des}) = (1+2\ell)180^\circ - \measuredangle G(s_{Des})}$$

Similar to Bode design approach, define:

$$\boxed{\varphi_{req} = (1+2\ell)180^\circ - \measuredangle G(s_{Des})}$$

(choose ℓ to get φ_{req} in range $[-180^\circ, +180^\circ]$)

Then choose poles/zeros in $H(s)$ so that

$$\boxed{\measuredangle H(s_{Des}) = \varphi_{req}}$$

Example:

Suppose $G(s) = \frac{3}{s(s+2)}$

and we want $T(s)$ to have a pole at $s_{Des} = -3+3j$

$$\measuredangle G(s_{Des}) = 116.56^\circ$$

(In Matlab: angle(evalfr(G, -3+3j)))

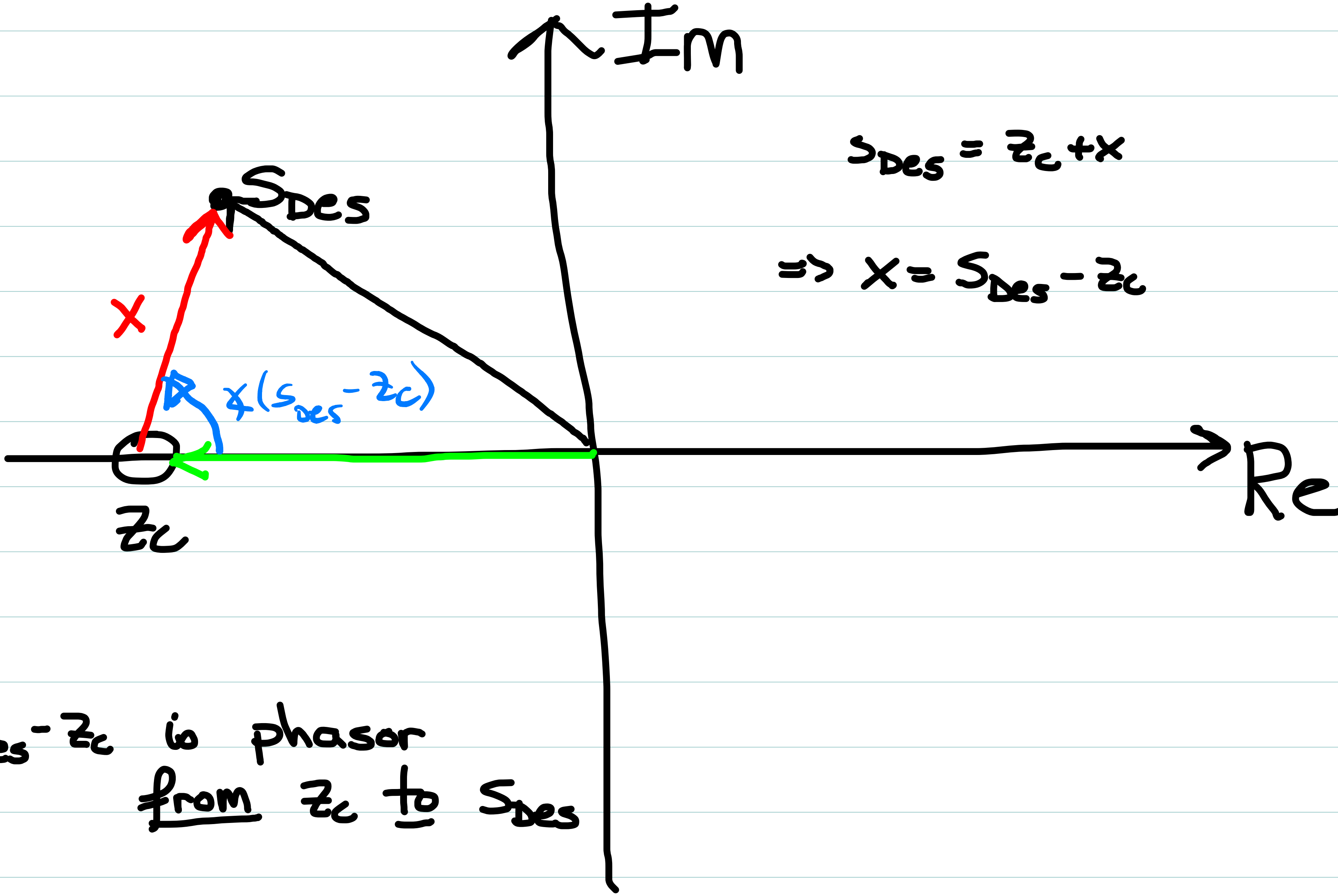
$$\Rightarrow \varphi_{req} = 180^\circ - 116.56^\circ = 63.43^\circ \text{ here}$$

Assume initially: $H(s) = K(s - z_c)$, $K > \emptyset$

(Not implementable: for illustration only!)

Then we need $\measuredangle(s_{Des} - z_c) = 64.43^\circ$

Visualization – Phasor Interpretation

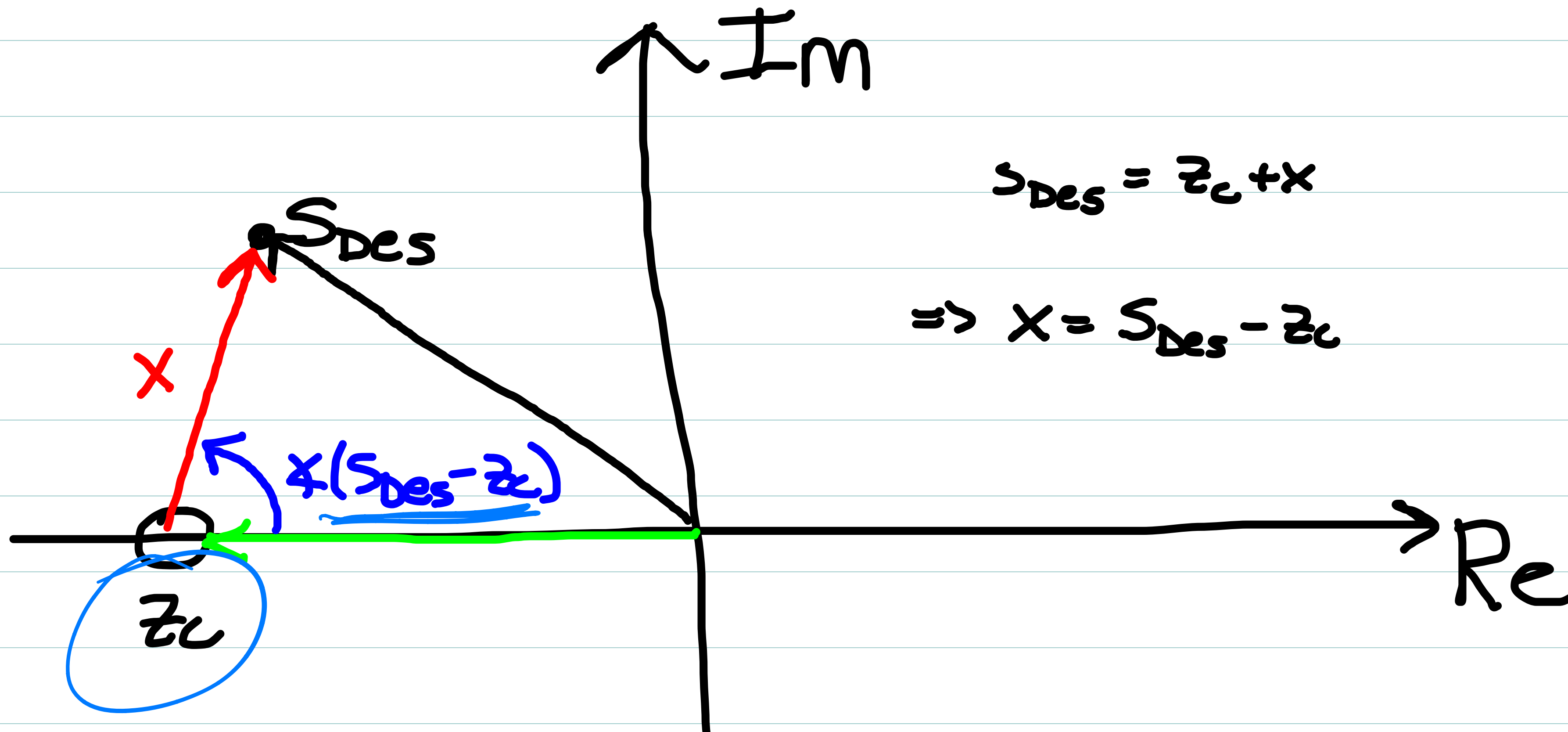


$$s_{Des} = z_c + x$$

$$\Rightarrow x = S_{Des} - z_c$$

$S_{Des} - z_c$ is phasor <u>from</u> z_c <u>to</u> S_{Des}

Visualization – Phasor Interpretation

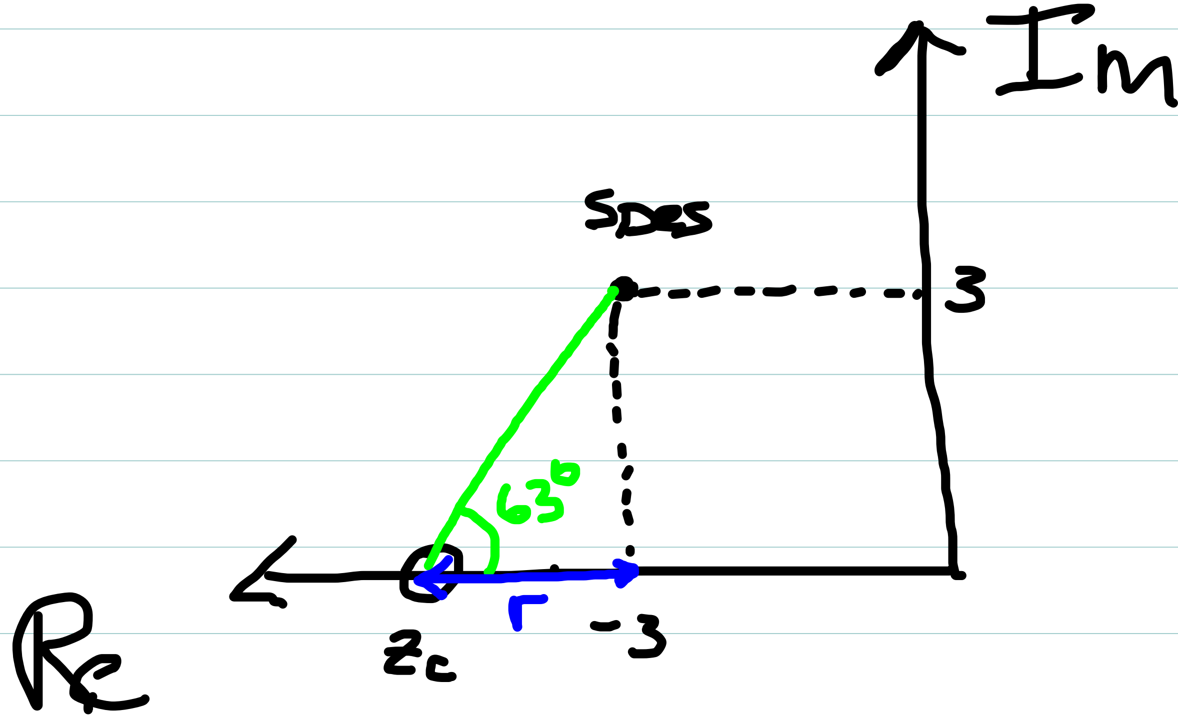


$s_{Des} - z_c$ is phasor
<u>from</u> z_c <u>to</u> s_{Des}

<u>Note</u>: unlike Bode designs we can get <u>up to +180°</u> at s_{Des} from a <u>single</u> zero.

Example cont'd

If we need $\measuredangle(s_{Des} - z_c) = 63.43°$ at $s_{Des} = -3+3j$:



$$\tan(63.43°) = \frac{3}{r} \implies r = 1.5$$

$$\implies z_c = -4.5$$

Example, cont'd

So $H(s) = K(s+4.5)$ and then

$$L_o(s) = \frac{3(s+4.5)}{s(s+2)}$$

$|L_o(s_{Des})| = 3/4$ (Matlab: abs(evalfr(Lo,-3+3j)))

so $K = 4/3$.

Check:

$$T(s) = \frac{4(s+4.5)}{s^2+2s+4(s+4.5)} = \frac{4(s+4.5)}{s^2+6s+18}$$ ✓

roots at $-3 \pm 3j$

Notes

1.) To be implementable $H(s)$ needs a pole. Choose pole p_c so that

$$\measuredangle(s_{Des} - p_c) \approx 5^\circ$$

Then $\measuredangle H(s_{Des}) = \measuredangle(s_{Des} - z_c) - \measuredangle(s_{Des} - p_c) = \measuredangle(s_{Des} - z_c) - 5^\circ$

Add $+5^\circ$ to φ_{req} to account for required pole.

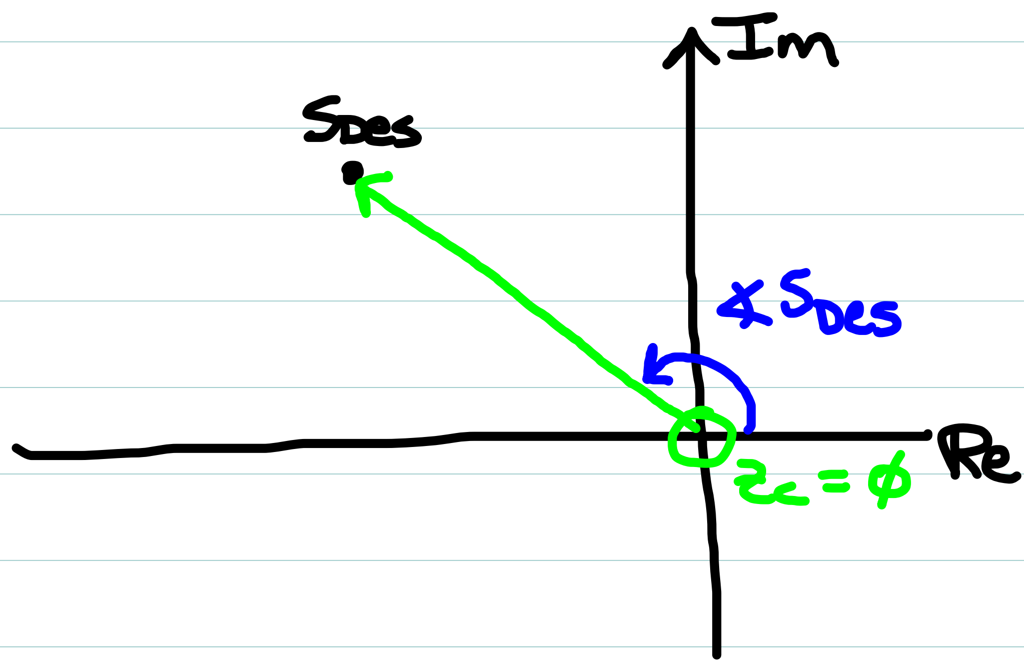
("β-minimizing" principle is quite messy here).

2.) Keep $\varphi_{req} < 90^\circ$, pref. below $60^\circ - 70^\circ$, or else zero will be closer to imag axis than s_{Des}, creating substantial additional overshoot. "Split" large φ_{req} over multiple zeros if necessary.

Notes (cont).

3.) Do not choose z_c in RHP! (We'll see why later)

$\Rightarrow$ places practical limit on maximum angle contribution from a zero



(effective max for $z_c \le \emptyset$)

Notes (cont')

4.) Design method guarantees s_{Des} is a CL pole, but says nothing about location of other CL poles.

$\boxed{\text{These might actually be unstable!}}$

Suppose: $G(s) = \dfrac{2}{s^2(s+1)}$, $s_{Des} = -2 + 0j$

$G(-2) = -1/2 \Rightarrow \varphi_{req} = \emptyset \Rightarrow H(s) = K > \emptyset$ sufficient

$K = \dfrac{-1}{-1/2} = 2$ and here

$$T(s) = \frac{4}{s^3 + s^2 + 4}$$

With poles at $\boxed{-2}$, $\boxed{\tfrac{1}{2} \pm \tfrac{4}{3}j}$ ← unstable!

To use these ideas effectively as a design tool, we must have *some* idea where the *other* poles of $T(s)$ will be; i.e. at least if they are stable.

Requires us to more generally understand *all* *possible* solutions of $1 + L(s) = \emptyset$

Or, equivalently, the "locus" of points in the complex plane which satisfy the angle condition(s):

$$\measuredangle L_o(s) = (1+2\ell)180^\circ \qquad (\text{if } K > \emptyset)$$

$$\measuredangle L_o(s) = (2\ell)180^\circ \qquad (\text{if } K > \emptyset).$$

"Root Locus" Method for CL pole prediction

Set up: $L(s) = K\left[\frac{N(s)}{D(s)}\right]$

$\Rightarrow$ $\mathrm{Deg}\{N(s)\} = m$; m zeros z_i such that $N(z_i) = \emptyset$

$$N(s) = (s - z_1)(s - z_2)\cdots(s - z_m) = \prod_{i=1}^{m}(s - z_i)$$

$\Rightarrow$ $\mathrm{Deg}\{D(s)\} = n$; n poles p_k such that $D(p_k) = \emptyset$

$$D(s) = (s - p_1)(s - p_2)\cdots(s - p_n) = \prod_{k=1}^{n}(s - p_k)$$

$\Rightarrow$ $n \geq m$: no more zeros than poles

$\Rightarrow$ Characteristic equation: s is a CL pole if

$$1 + L(s) = \emptyset$$

Basic Observations

$$1+L(s)=\phi \Rightarrow 1+K\left[\frac{N(s)}{D(s)}\right]=\phi$$

$$\Rightarrow D(s)+KN(s)=\phi$$

This is an n^{th} order polynomial equation to define CL poles:

$\boxed{\Rightarrow \text{There are } n \text{ CL poles, same number as OL poles}}$

Consider limit as $K\to\phi$. Then CE becomes: $D(s)=\phi$

$\Rightarrow$ same eq'n as defines OL poles.

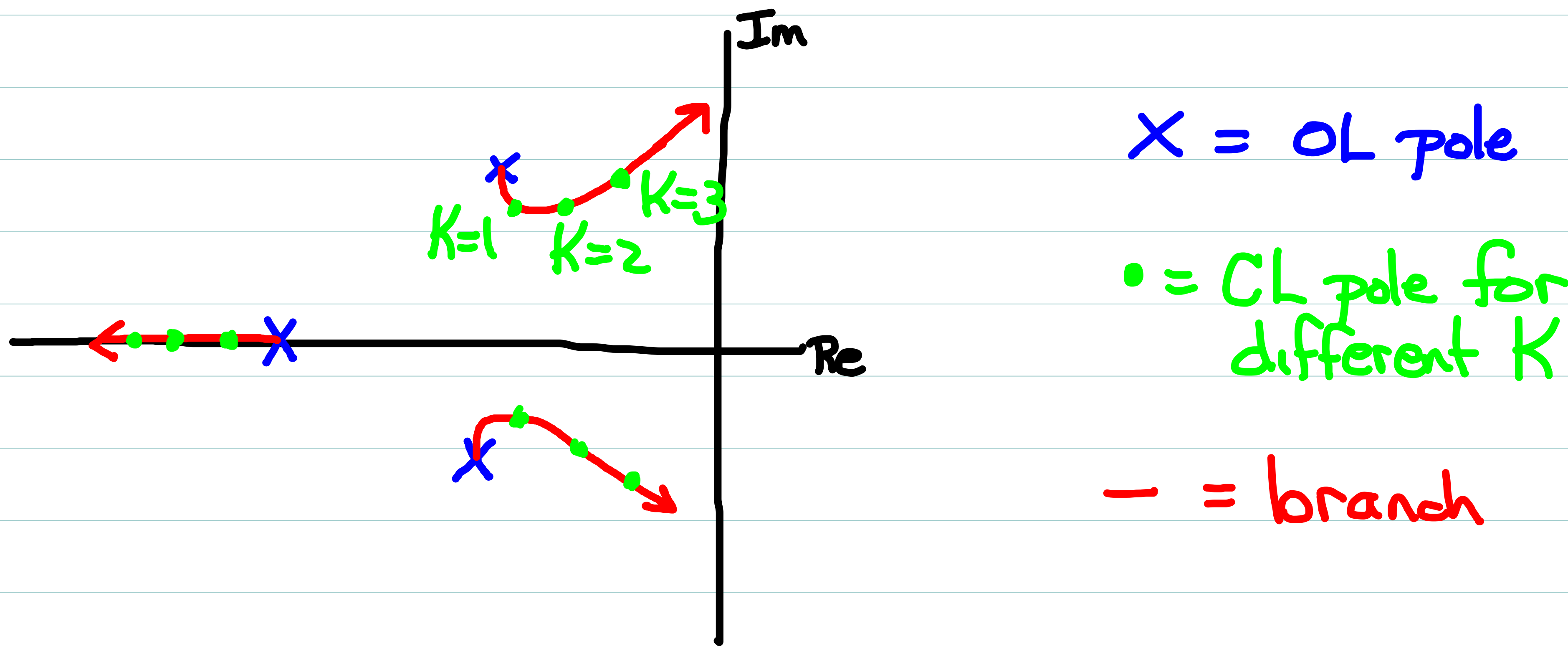
$\boxed{\Rightarrow \text{In low gain limit, } K\to\phi\text{, the CL poles are same as OL poles}}$

Varying K

⇒ As K changes, the CL pole locations migrate away from OL poles

⇒ Each CL pole location traces out a continuous curve starting at an OL pole. These curves are called **branches**.

⇒ Since there are n CL poles, there are n branches



Symmetry

⇒ Recall that complex roots of polynomial equations occur in <u>conjugate pairs</u>.

⇒ If $s \in \mathbb{C}$ satisfies $1+L(s)=\emptyset$, so also $\bar{s}$ satisfies $1+L(\bar{s})=\emptyset$.

⇒ CL pole locations are symmetric about <u>real</u> Axis.

⇒ Branches of CL pole loci are symmetric ("mirror image") about real Axis.

⇒ Can we predict branch behavior as $|K|$ increases?

High gain limit: $|K| \to \infty$

Recall CL poles satisfy $D(s) + K N(s) = \emptyset$

Equivalently if $K \neq \emptyset$:

$$N(s) + \left[\frac{1}{K}\right] D(s) = \emptyset$$

and as $|K| \to \infty$ we have: $N(s) = \emptyset$

$\Rightarrow$ As $|K| \to \infty$, the CL poles coincide with OL zeros!

$\Rightarrow$ Branches terminate at OL zeros!

$\Rightarrow$ OL zeros "attract" CL poles to them in high gain limit

$\Rightarrow$ RHP zeros in $L(s)$ are dangerous!

High gain limit, cont

$\Rightarrow$ n CL poles (branches), but only $m \leq n$ OL zeros.

$\Rightarrow$ What happens to other $n-m$ CL poles (branches)?

$\Rightarrow$ *The remaining $n-m$ branches asymptote to infinity*

$\Rightarrow$ But how? Depends on sign of K. Suppose for simplicity we take $K > 0$.

$\Rightarrow$ Recall "angle condition" for $K > 0$:

if s is a possible CL pole, then

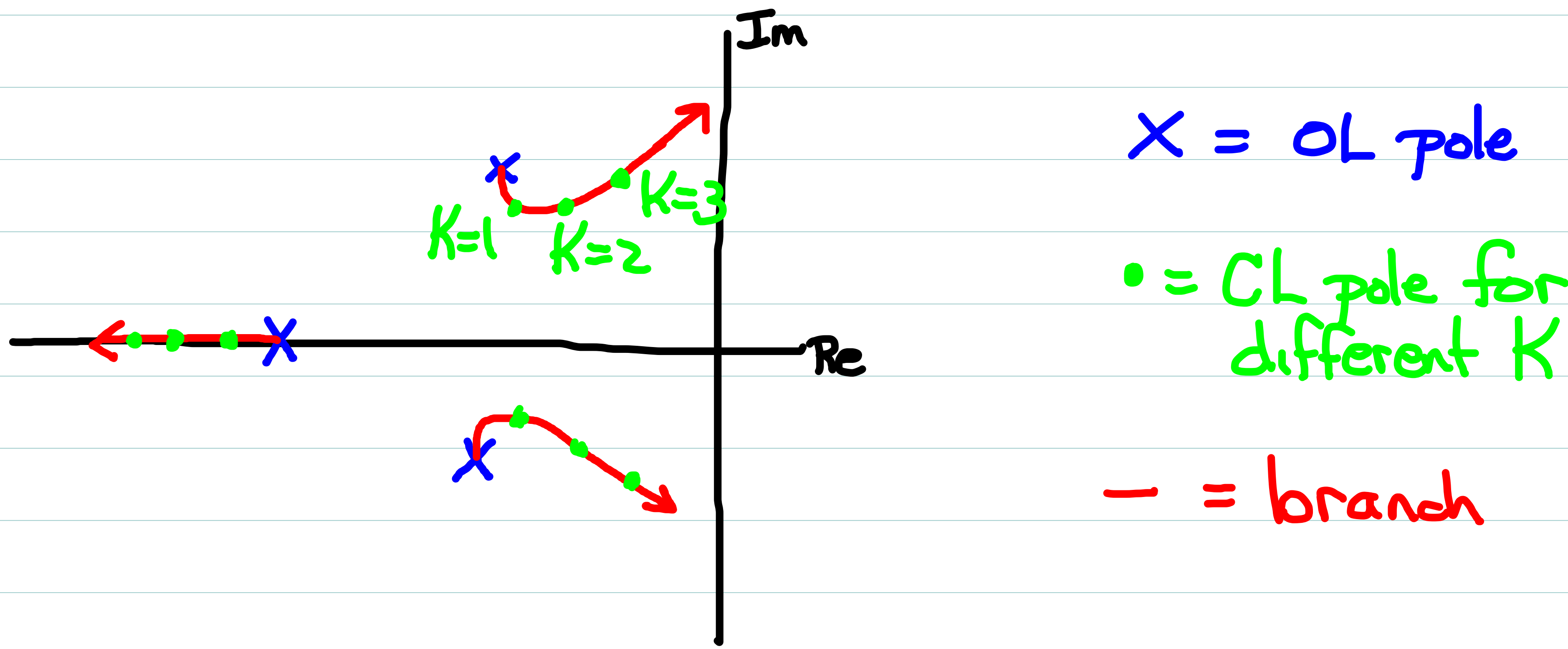
$$\angle L(s) = (1+2\ell)180^\circ \quad \text{(odd multiple of } 180^\circ\text{)}.$$

Varying K

⇒ As K changes, the CL pole locations migrate away from OL poles

⇒ Each CL pole location traces out a continuous curve starting at an OL pole. These curves are called branches.

⇒ Since there are n CL poles, there are n branches



Symmetry

⇒ Recall that complex roots of polynomial equations occur in conjugate pairs.

⇒ If $s \in \mathbb{C}$ satisfies $1 + L(s) = 0$, so also $\bar{s}$ satisfies $1 + L(\bar{s}) = 0$.

⇒ CL pole locations are symmetric about real Axis.

⇒ Branches of CL pole loci are symmetric ("mirror image") about real Axis.

⇒ Can we predict branch behavior as $|K|$ increases?

High gain limit : $|K| \to \infty$

Recall CL poles satisfy $D(s) + KN(s) = 0$

Equivalently if $K \neq 0$:

$$N(s) + \left[\frac{1}{K}\right] D(s) = 0$$

and as $|K| \to \infty$ we have : $N(s) = 0$

$\Rightarrow$ As $|K| \to \infty$, the CL poles coincide with OL zeros!

$\Rightarrow$ Branches terminate at OL zeros!

$\Rightarrow$ OL zeros "attract" CL poles to them in high gain limit

$\Rightarrow$ RHP zeros in $L(s)$ are dangerous!

High gain limit, cont

$\Rightarrow$ n CL poles (branches), but only $m \leq n$ OL zeros.

$\Rightarrow$ What happens to other $n-m$ CL poles (branches)?

$\Rightarrow$ *The remaining $n-m$ branches asymptote to infinity*

$\Rightarrow$ But how? Depends on sign of K. Suppose for simplicity we take $K > \emptyset$.

$\Rightarrow$ Recall "angle condition" for $K > \emptyset$:

if s is a possible CL pole, then

$$\angle L(s) = (1+2\ell)180^\circ \quad \text{(odd multiple of } 180^\circ\text{).}$$

Interpretation of Angle Condition

Just like in Bode, for any $s \in \mathbb{C}$:

$$\measuredangle L(s) = \sum_{i=1}^{m} \measuredangle (s - z_i) - \sum_{k=1}^{n} \measuredangle (s - p_k)$$

More compactly:

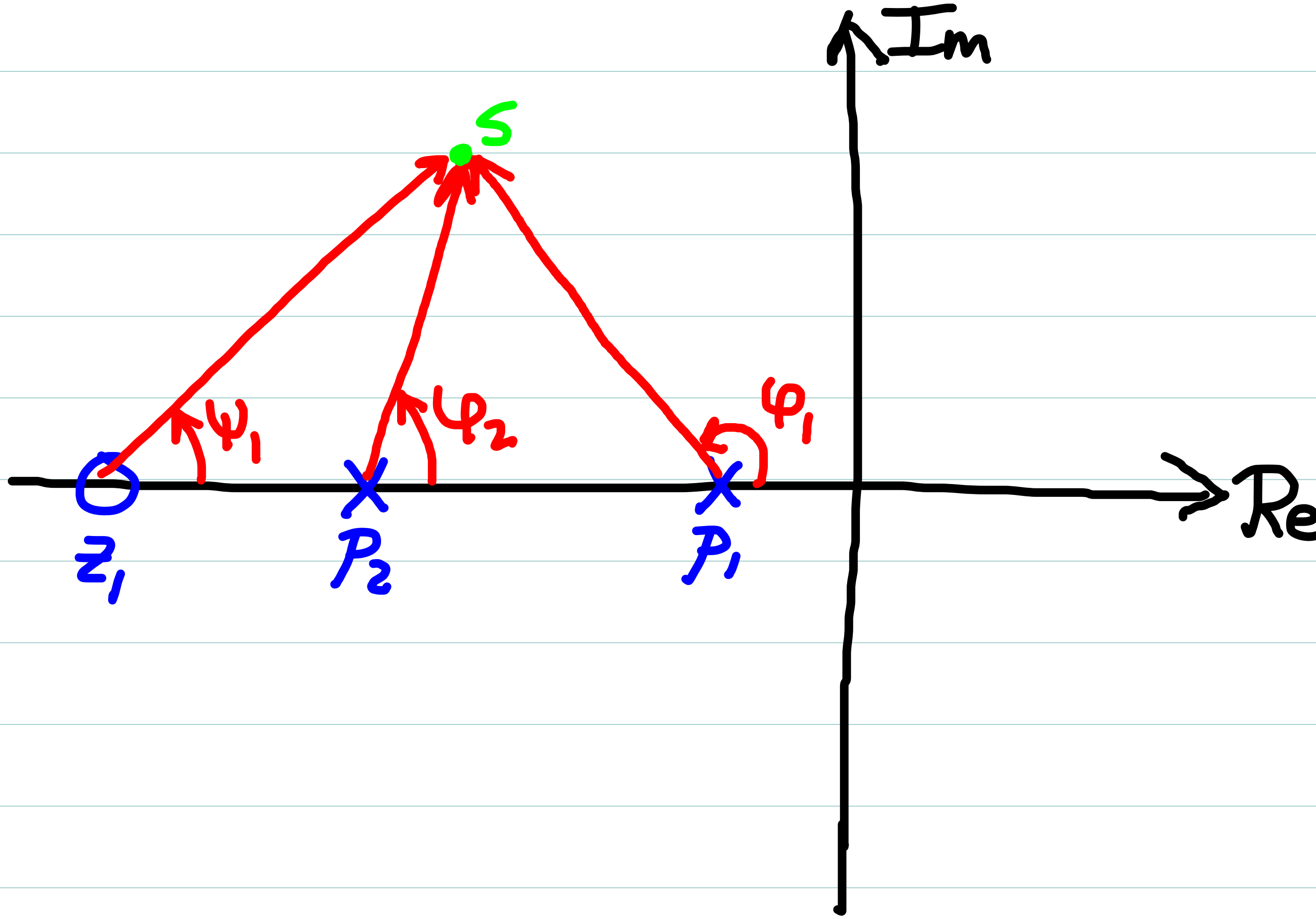
$$\measuredangle L(s) = \sum_{i=1}^{m} \psi_i - \sum_{k=1}^{n} \varphi_k$$

where.

$\psi_i = \measuredangle (s - z_i)$ (Contribution of each zero)

$\varphi_k = \measuredangle (s - p_k)$ (Contribution of each pole)

Graphical (Phasor) Interpretation



$\Rightarrow$ s is a possible CL pole (hence lies on a branch of the locus) if:

$$\psi_1 - \varphi_1 - \varphi_2 = (1+2\ell)180^\circ$$

For high gain limit, look for s with $|s| \gg 1$ which satisfy this

Note that for such s, the phasors become indistinguishable!

$$\varphi_1 \approx \varphi_1 \approx \varphi_2 = \alpha$$

Thus, for $|s| \gg 1$, the angle condition becomes:

$$(1+2l)180° = \sum_{i=1}^{m} \psi_i - \sum_{k=1}^{n} \varphi_k$$

$$= m\alpha - n\alpha$$

$$= (m-n)\alpha$$

Where α is common phasor angle from z_i or p_k to s.

We then have

$$\boxed{\alpha = \frac{(1+2l)180°}{n-m}}$$

is the angular direction in complex plane for s with $|s| \gg 1$ that satisfy angle condition

Asymptotes

$\Rightarrow$ The $n-m$ branches which diverge to infinity do so along asymptotes which make angles of

$$\boxed{\alpha_\ell = \frac{(1+2\ell)180^\circ}{n-m}}$$

with respect to real Axis.

$\Rightarrow$ A slightly messy additional derivation shows these asymptotes intersect at a common point on real Axis, given by

$$\boxed{\sigma_a = \frac{\sum_{k=1}^{n} Re\{p_k\} - \sum_{i=1}^{m} Re\{z_i\}}{n-m}}$$

where again z_i, p_k are zeros and poles of $L(s)$.

"Asymptote rule"

There will be $n-m$ unique asymptotes that result from this equation.

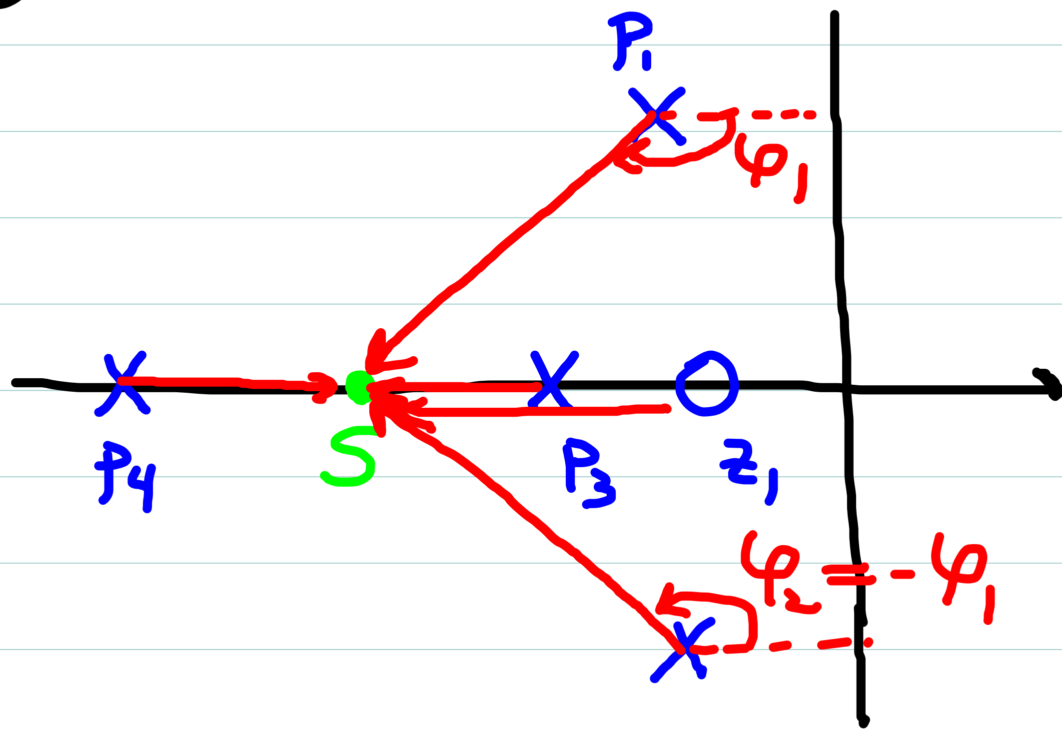
$n-m$	α_ℓ
1	-180°
2	$\pm 90^\circ$
3	$\pm 60^\circ, -180^\circ$
4	$\pm 45^\circ, \pm 135^\circ$

$\vdots$

etc

Branches on real Axis

Look at angle condition on real Axis



$\Longrightarrow$ Contribution to angle condition from complex conjugate pole or zero pairs will <u>cancel</u>.

$\Rightarrow$ Contribution from any real pole or zero to <u>left</u> of s will be zero

Thus, only the real poles and zeros lying to right of s will contribute to angle condition at a real s.

In particular:

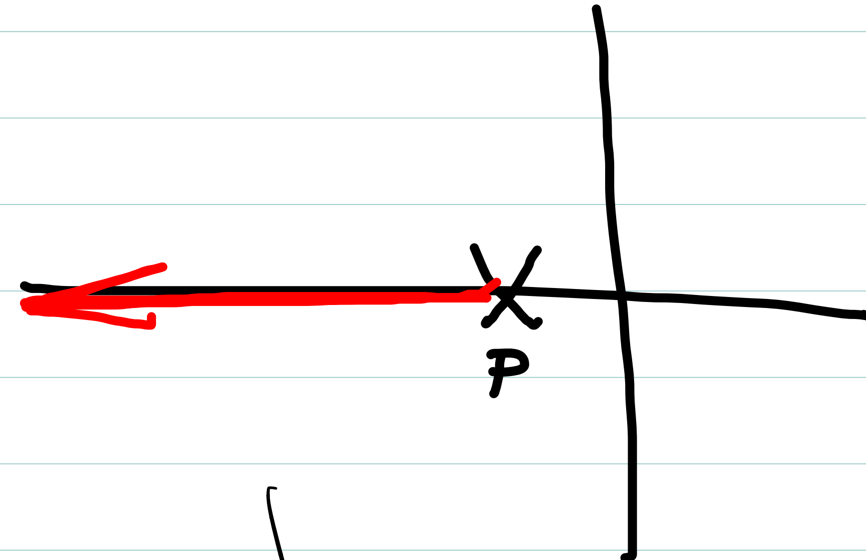
If the total number of real poles and zeros lying to the right of a point s on real Axis is odd, then that point satisfies the angle condition.

Portions of branches of the locus lie on segments of real Axis which satisfy this cond'n:

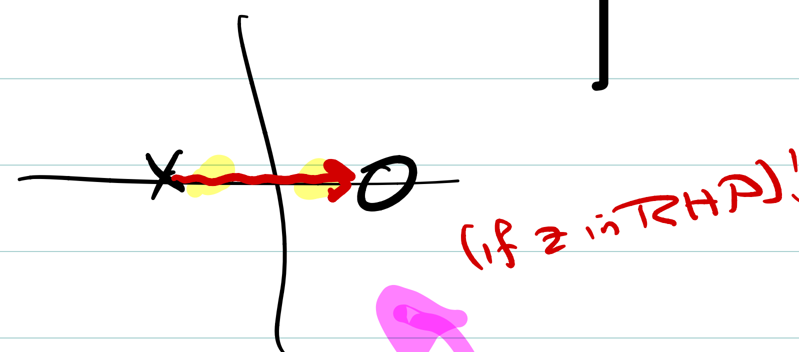
"real Axis rule"

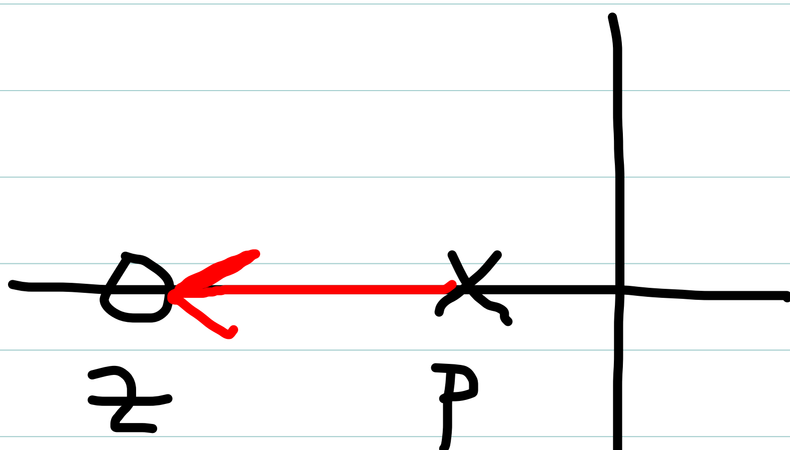
Simple Examples

#1) $L(s) = \dfrac{K}{s-p}$

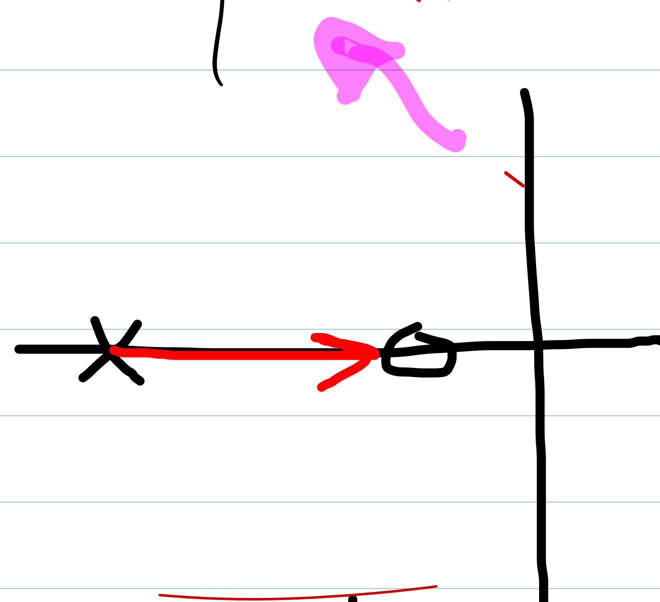


#2) $L(s) = K\left[\dfrac{(s-z)}{(s-p)}\right]$





$|z| > |p|$

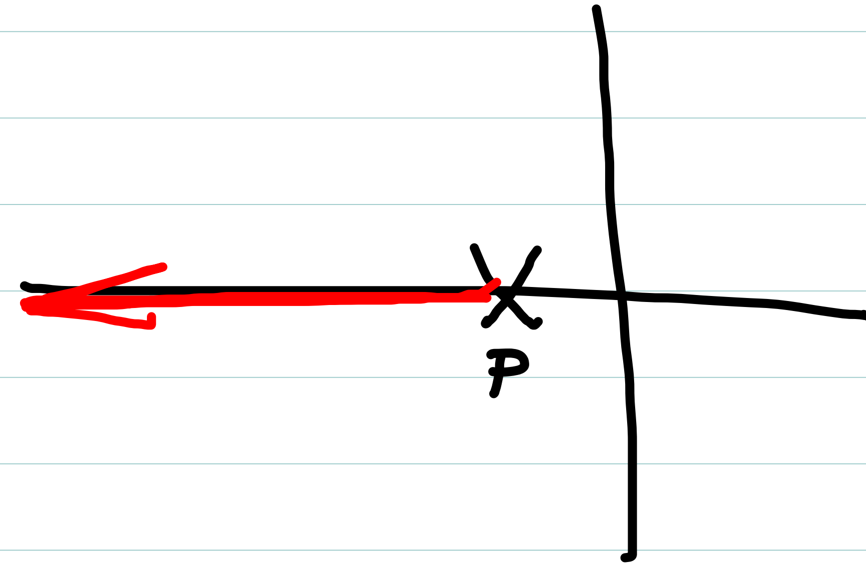


$|z| < |p|$

Simple Examples

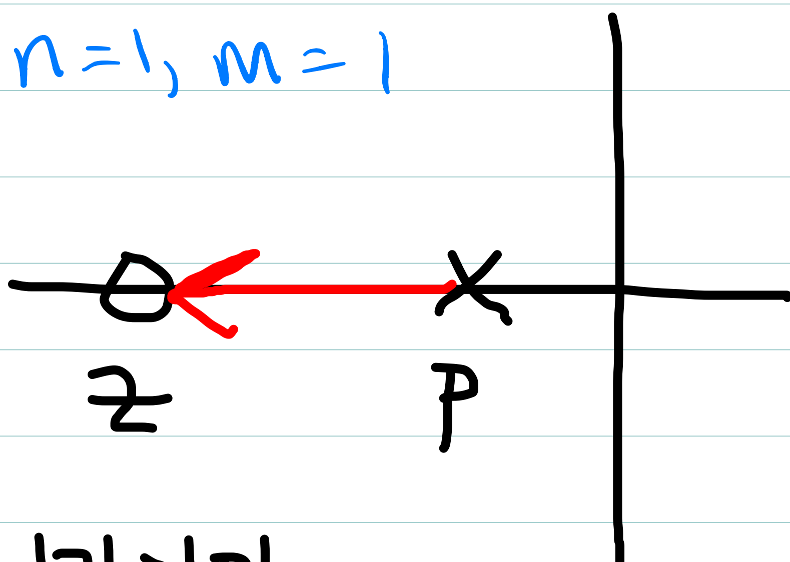
#1) $L(s) = \dfrac{K}{s-p}$

$n=1,\ m=0$

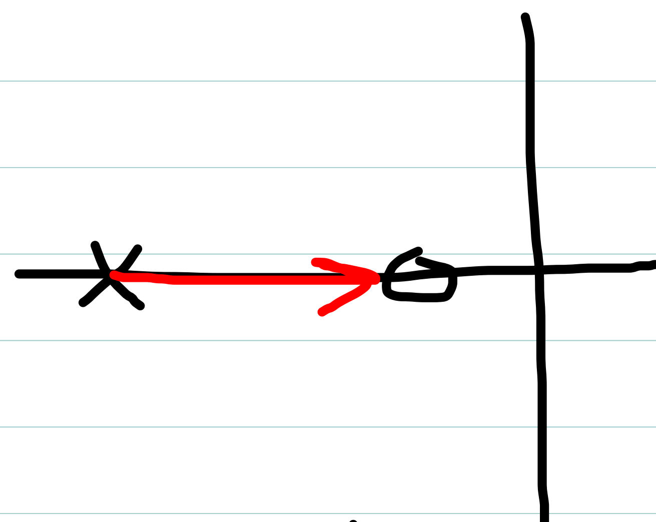


#2) $L(s) = K\left[\dfrac{(s-z)}{(s-p)}\right]$

$n=1,\ m=1$

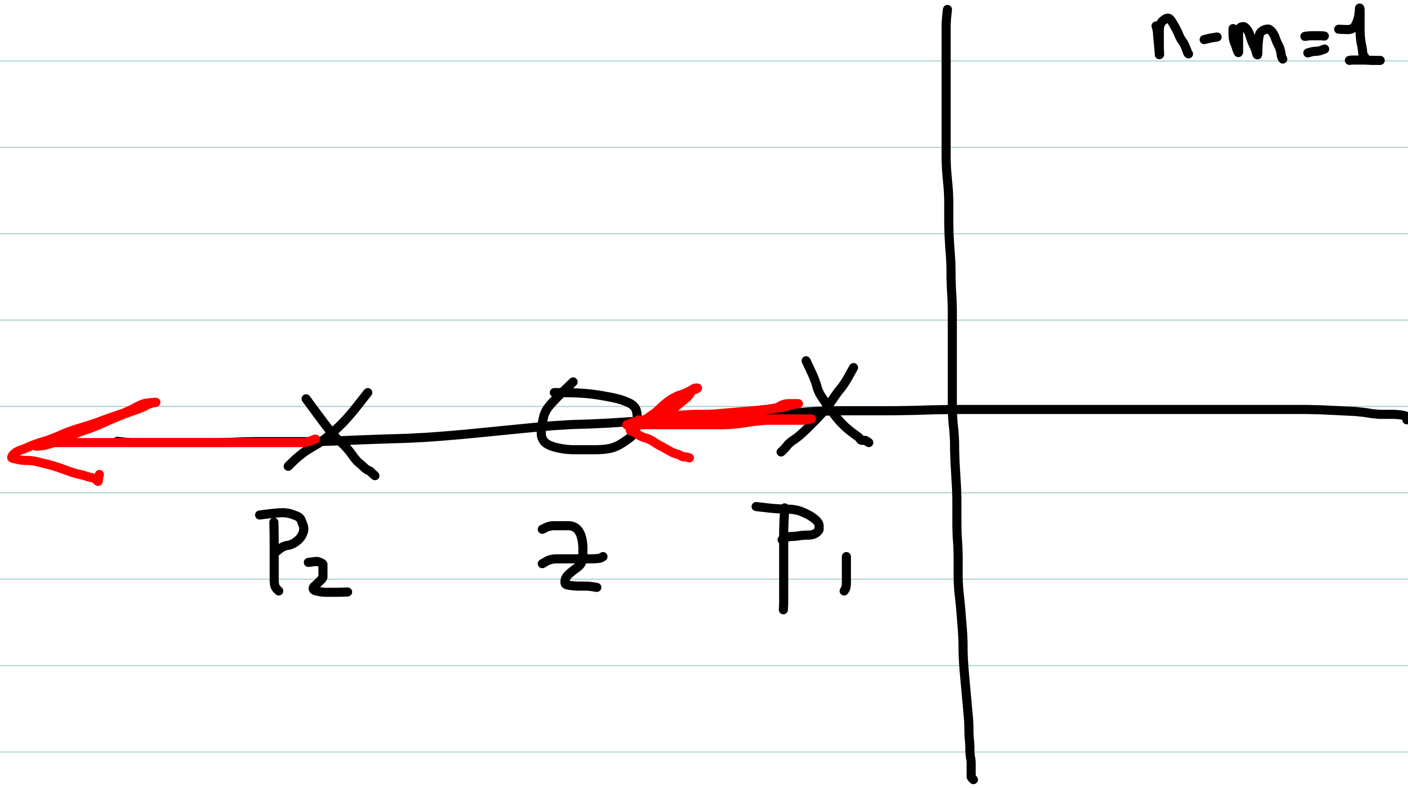


$|z|>|p|$



$|z|<|p|$

#3] $L(s) = K\left[\dfrac{(s-z)}{(s-p_1)(s-p_2)}\right]$



$n-m=1$: asymptote along neg real Axis

$|p_2| > |z| > |p_1|$

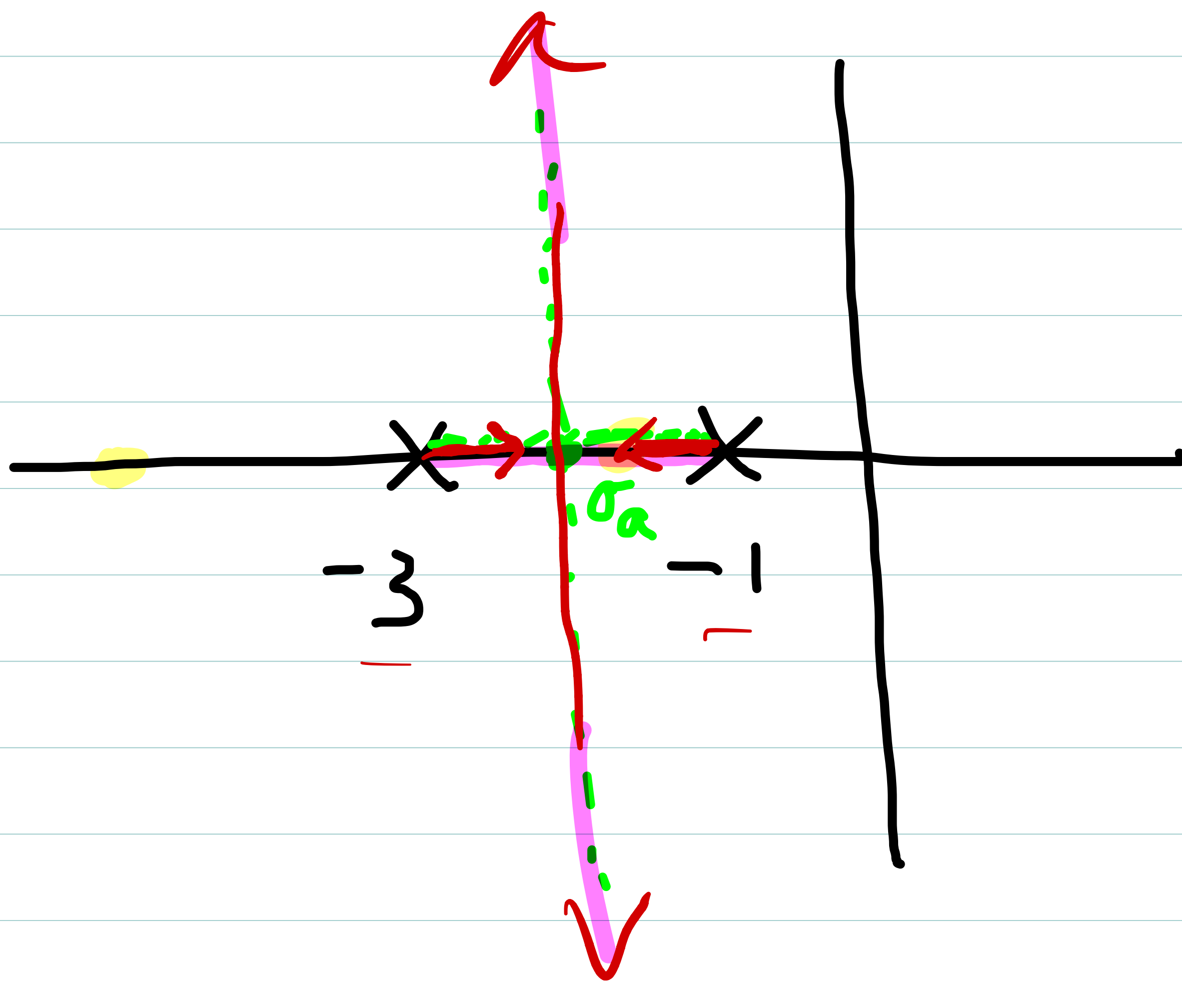
Case where $|z| < |p_1| \leq |p_2|$ is more complicated -- see below.

#4] $L(s) = \dfrac{K}{(s+1)(s+3)}$

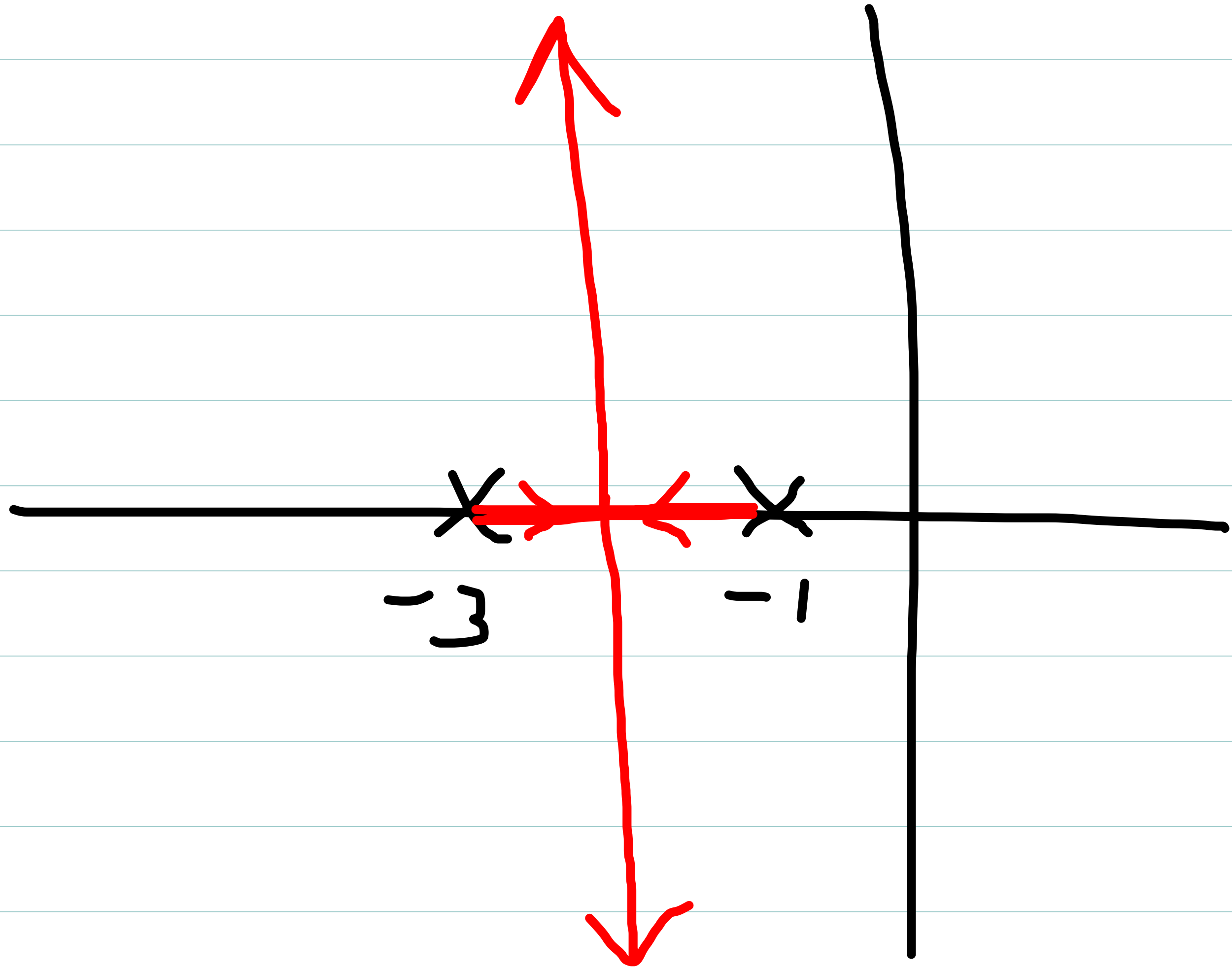
$n-m=2 \Rightarrow \alpha_\ell = \pm 90^\circ$

real axis locus from $-3 \rightarrow -1$

$\sigma_a = \dfrac{(-1)+(-3)}{2} = -2$



Actual locus:



Compare w/ exact sol'n for CL poles:

$$s^2+4s+(3+K)=0 \implies s = -2 \pm \frac{\sqrt{16-(3+K)^2}}{2}$$

Example #5

$$L(s) = \frac{K}{s(s+1)(s+2)}$$

$n=3$, $m=\emptyset$, $n-m=3 \Rightarrow$ 3 branches go to ∞ along

asymptotes: $\alpha_e = \pm 60°, 180°$

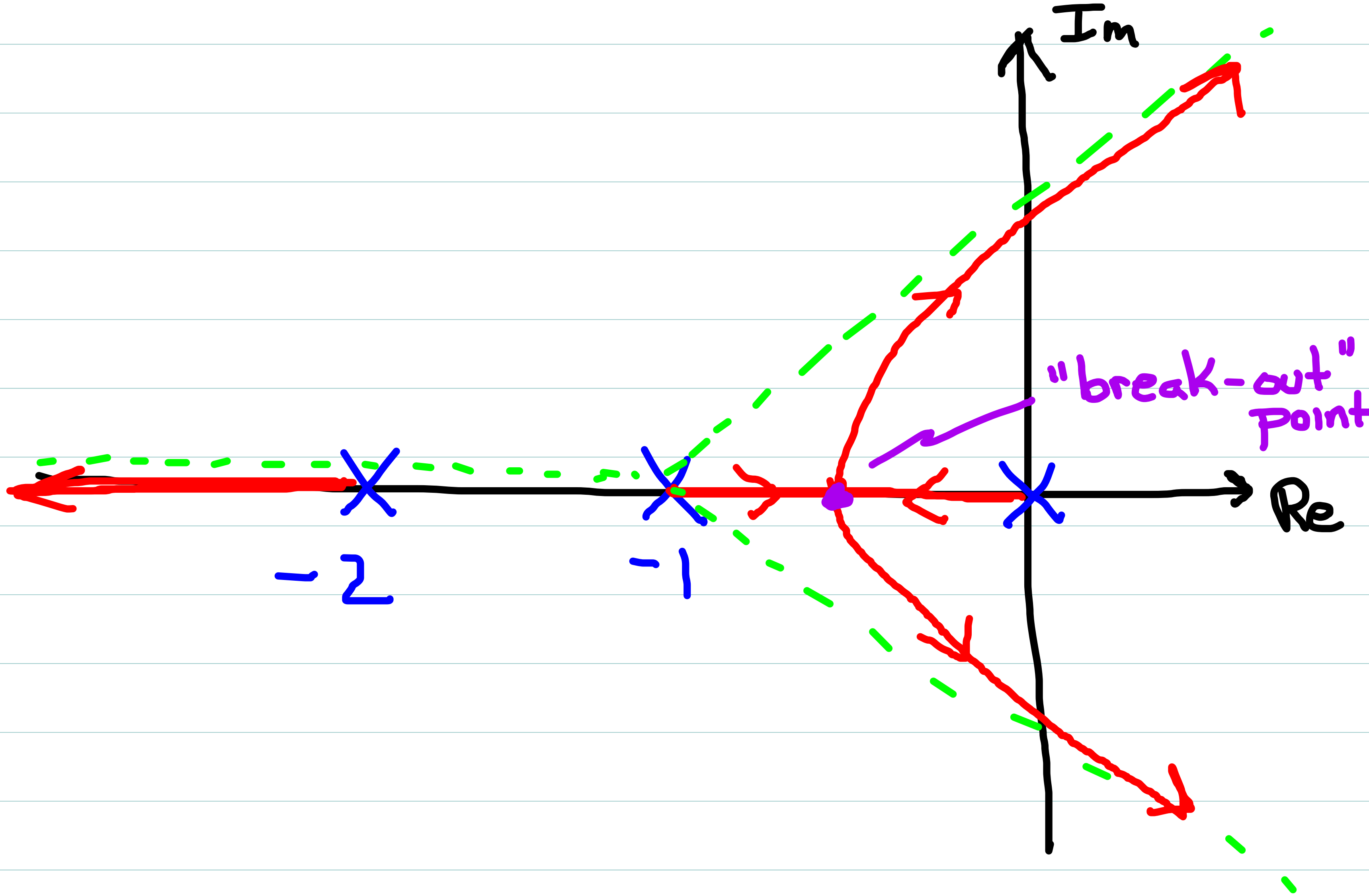
with intercept: $\sigma_a = \dfrac{0+(-1)+(-2)}{3} = -1$

Real Axis branch locations.

$\Rightarrow$ between -1 and $\emptyset$

$\Rightarrow$ left of -2

Example #5, cont



Break-out Points

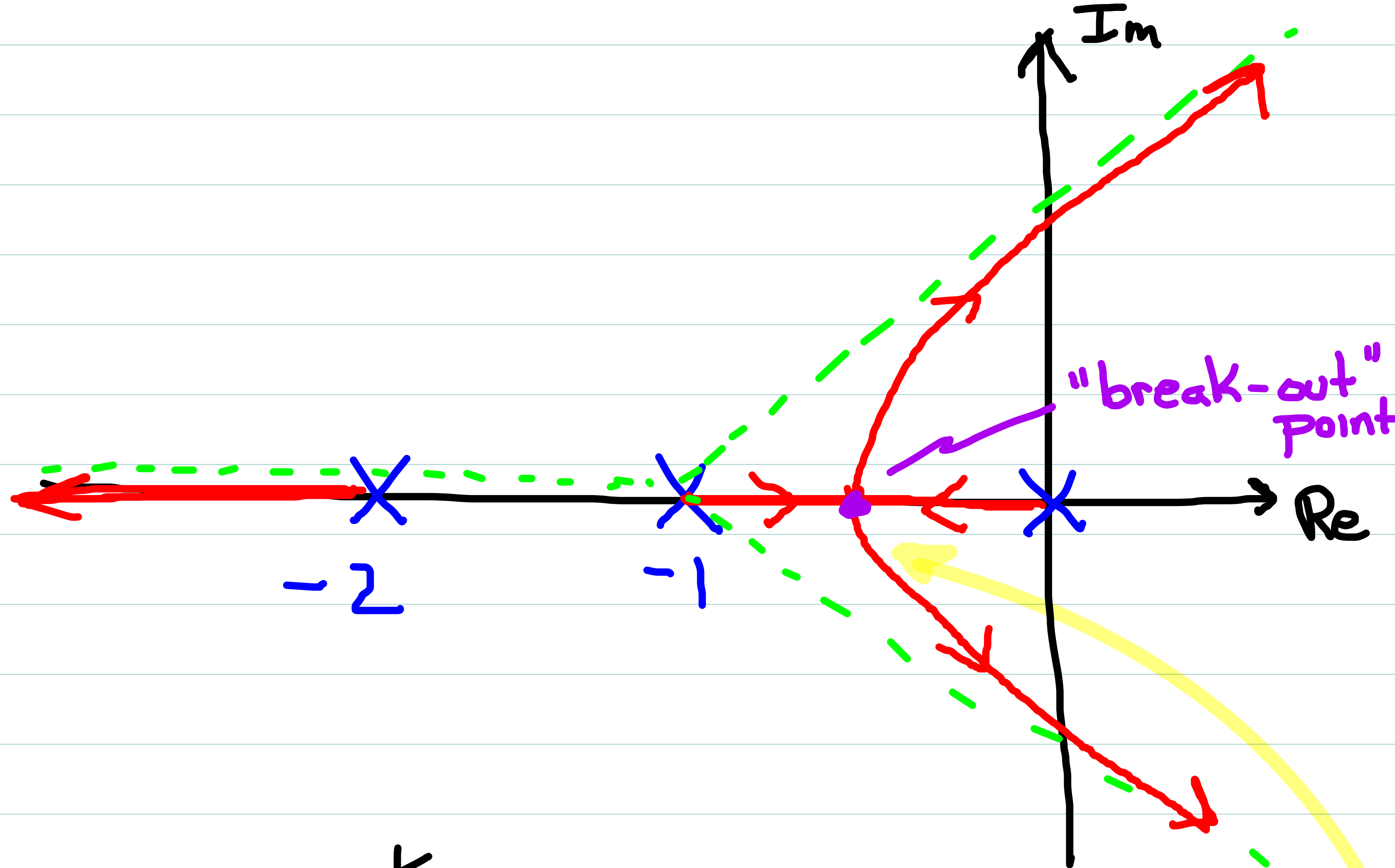
Break-out points occur for values of s satisfying

$$\frac{d}{ds}\left(L(s)\right) = 0$$

Since this (usually) leads to another high-order polynomial to factor, we often just approximate a break-out as occuring half-way along the branch

Use Matlab ("rlocus" command) to nail exact details when needed).

Example #5, cont

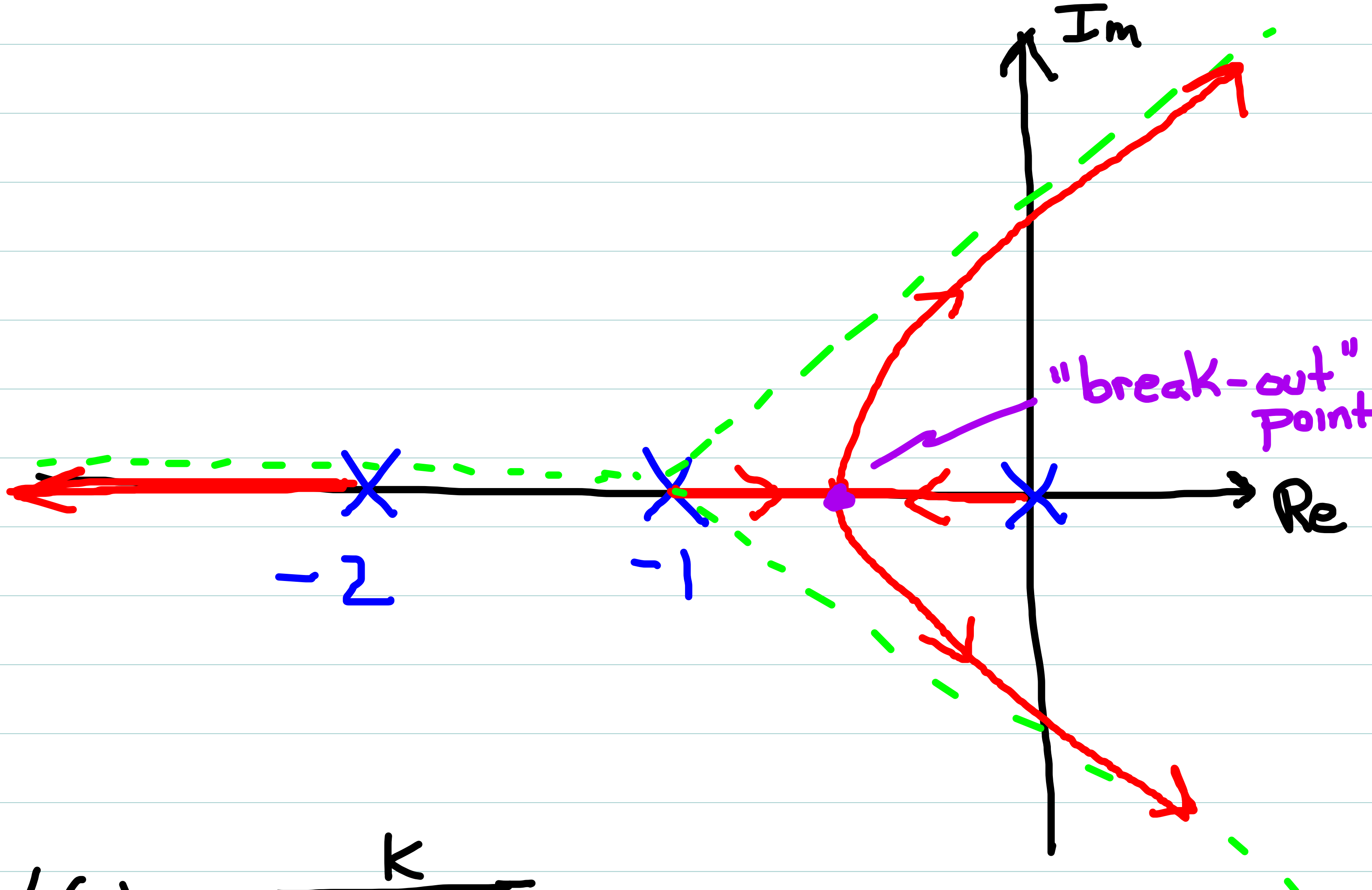


$$L(s) = \frac{K}{s(s+1)(s+2)}$$

$$\frac{d}{ds}L(s) = -\left[(s+1)(s+2) + s(s+2) + s(s+1)\right]$$
$$= -(3s^2+6s+2)$$

$$\frac{dL(s)}{ds} = 0$$
$$\Rightarrow s = \cancel{-1.57}, -.423$$

Example #5, cont



$$L(s) = \frac{K}{s(s+1)(s+2)}$$

High Gain Instability

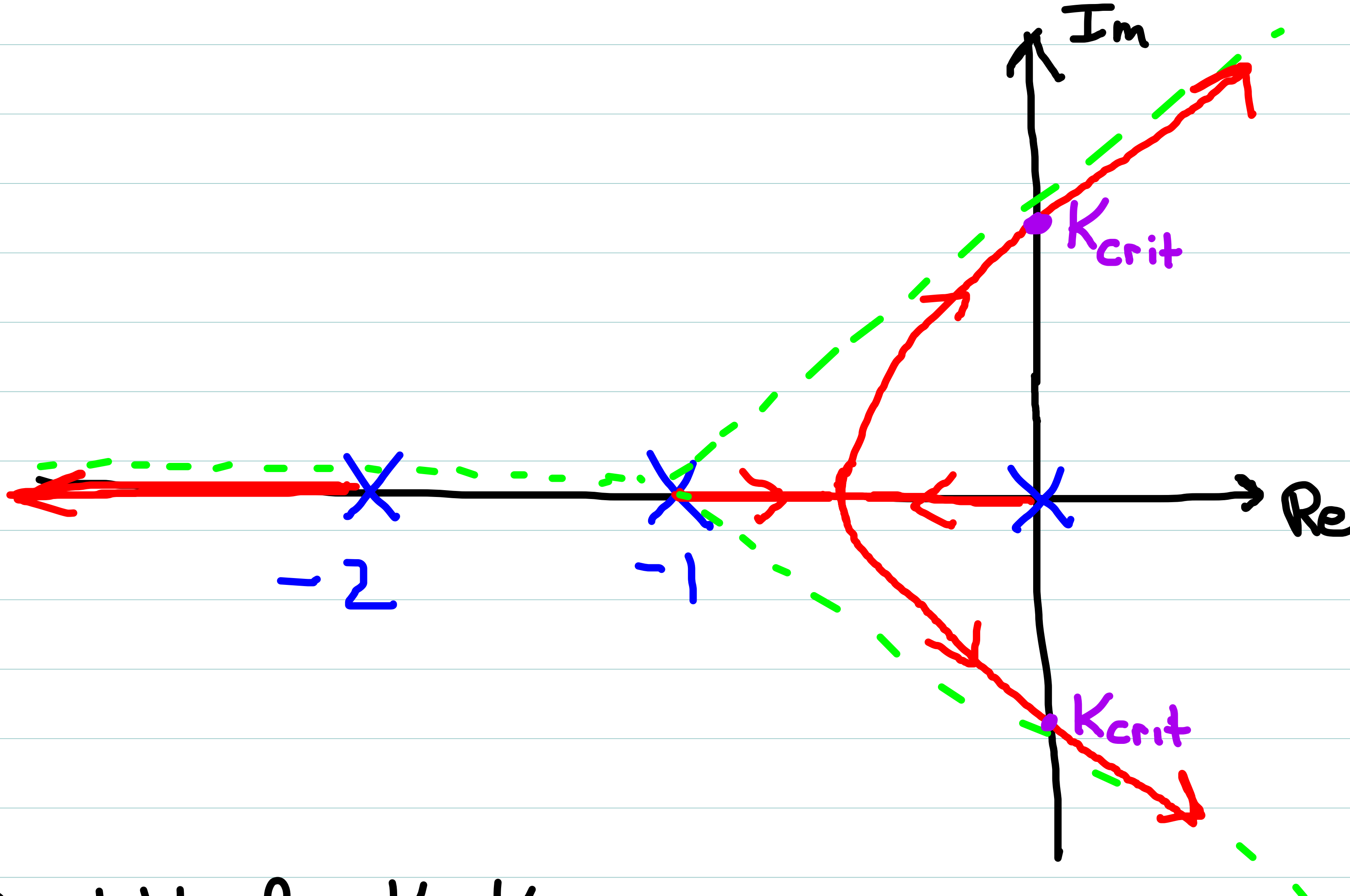
Since this example has Asymptotes in RHP, we can see the CL system will be unstable for sufficiently high gains K.

Whether this is a problem or not depends on the gain we want/need to get the desired CL poles

T(s) is Not automatically unstable b/c the root locus branches in RHP!

Such a locus only tells us T(s) will be unstable for some values of K.

Example #5, cont



Unstable for $K > K_{crit}$

Example #6

$$L(s) = \frac{K(s+3)}{(s+1)(s+2)}$$

$\Rightarrow n=2,\ m=1,\ n-m=1$

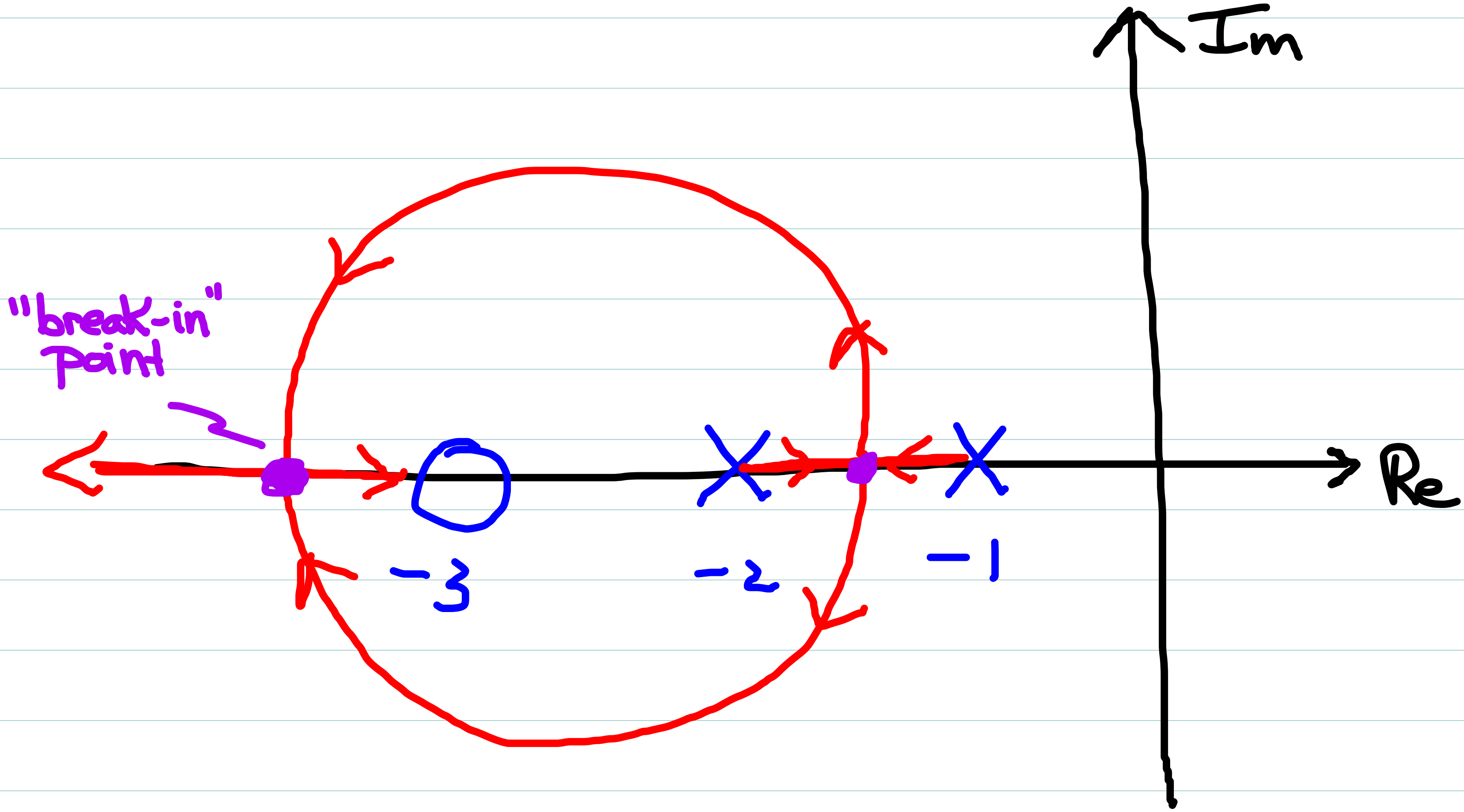
$\Rightarrow$ One branch ends at -3 (OL zero). One branch goes to ∞ along asymptote $\alpha = 180°$ (negative real Axis)

$\Rightarrow$ Segments of branches lie on real Axis:

$\Rightarrow$ Between -2 and -1

$\Rightarrow$ left of -3

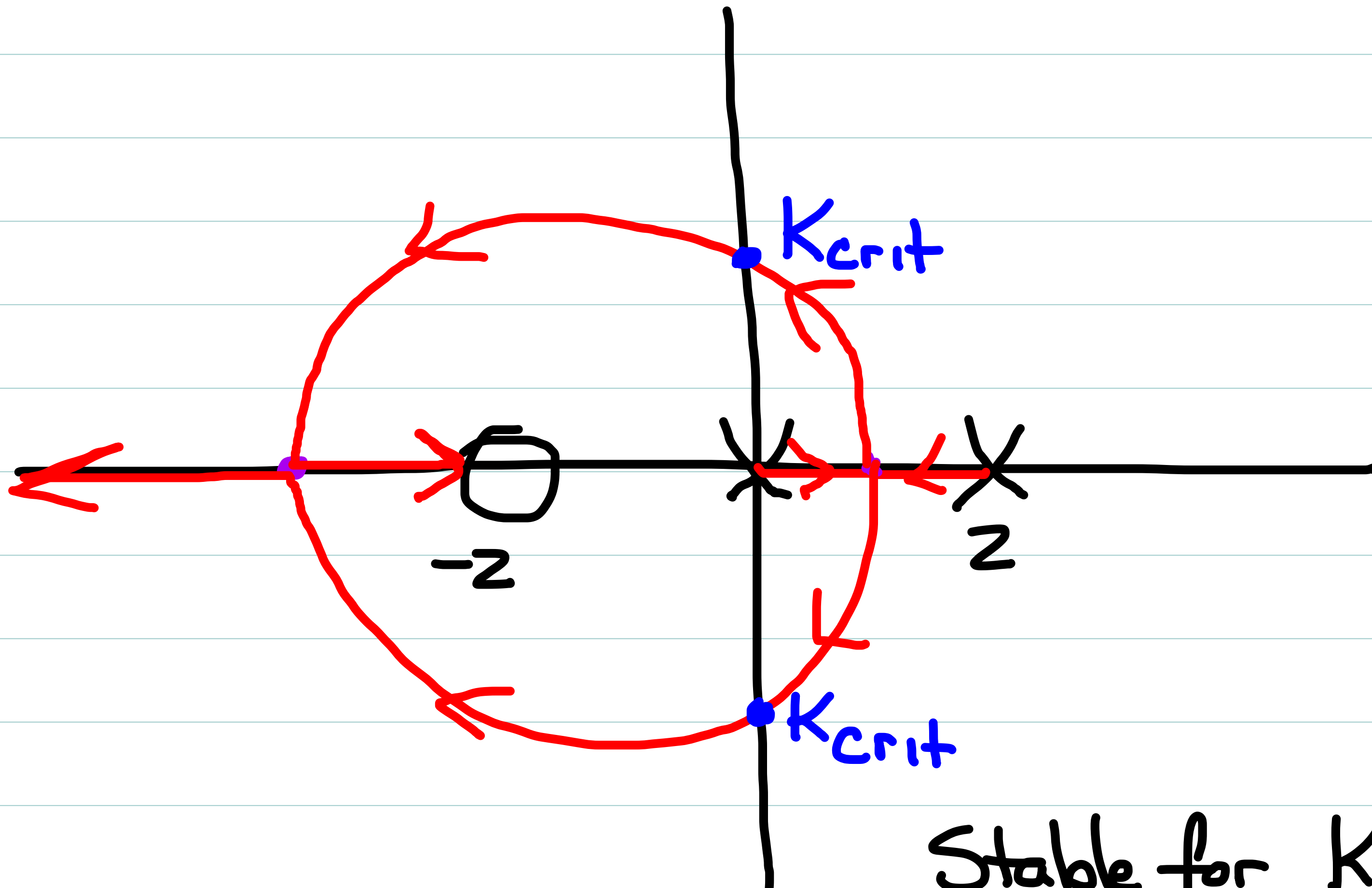
Example #6, cont



Example #7

$$L(s) = \frac{K(s+2)}{s(s-2)}$$

Similar analysis to above

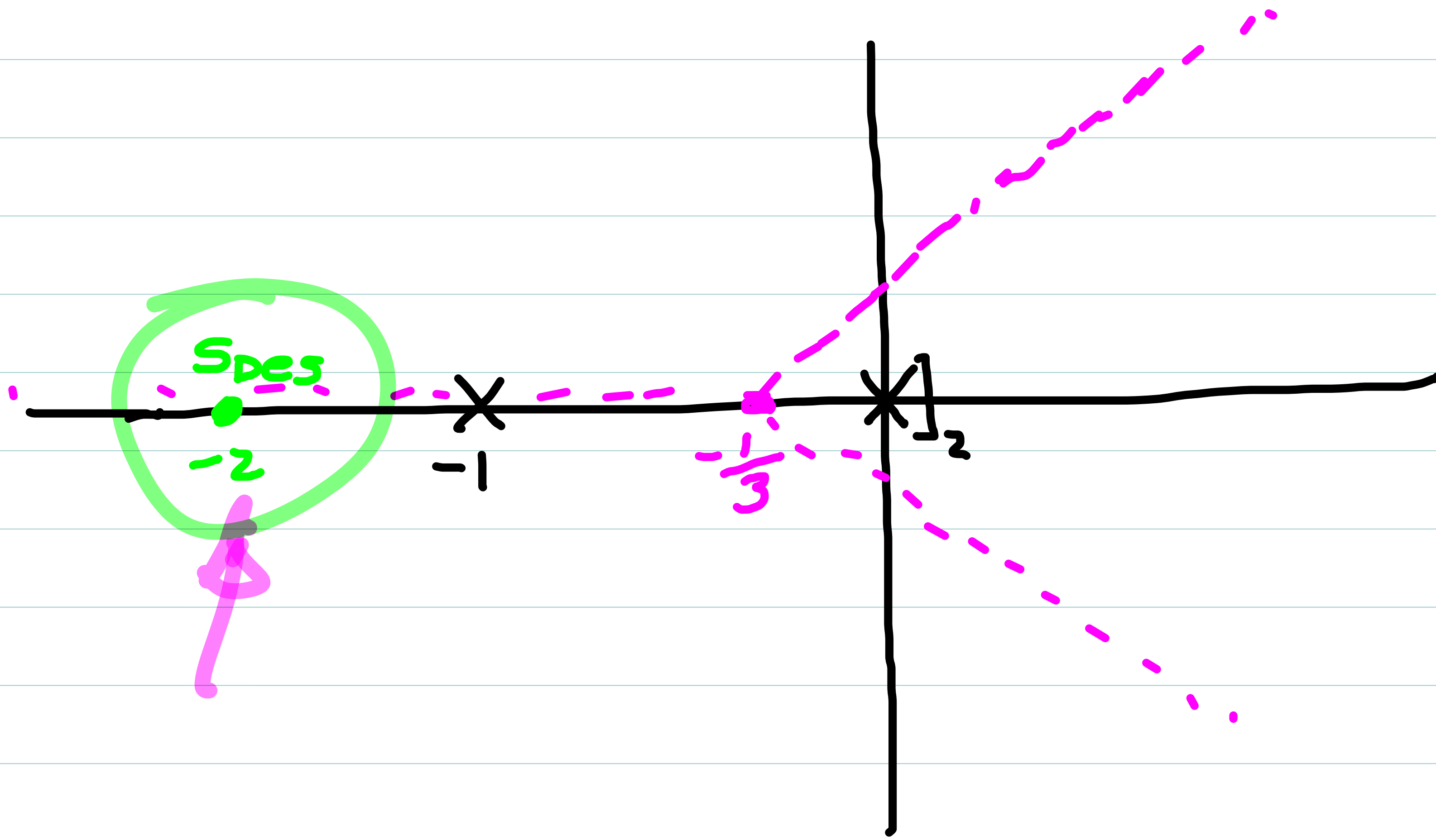


Stable for $K > K_{crit}$

Example #3

This is where we originally started our investigation

with $H(s) = K$, $L(s) = \dfrac{K}{s^2(s+1)}$



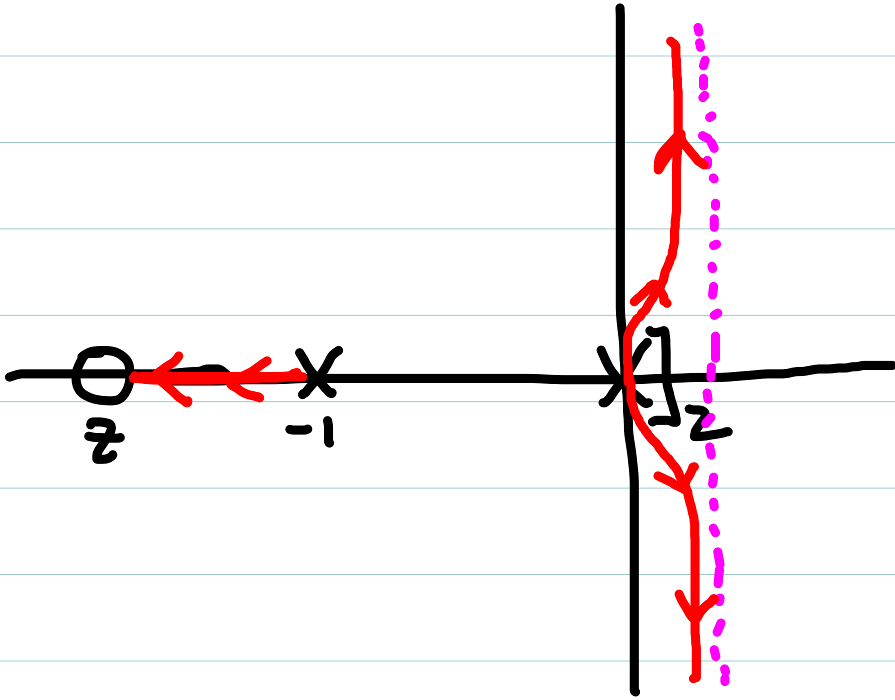
Example #3

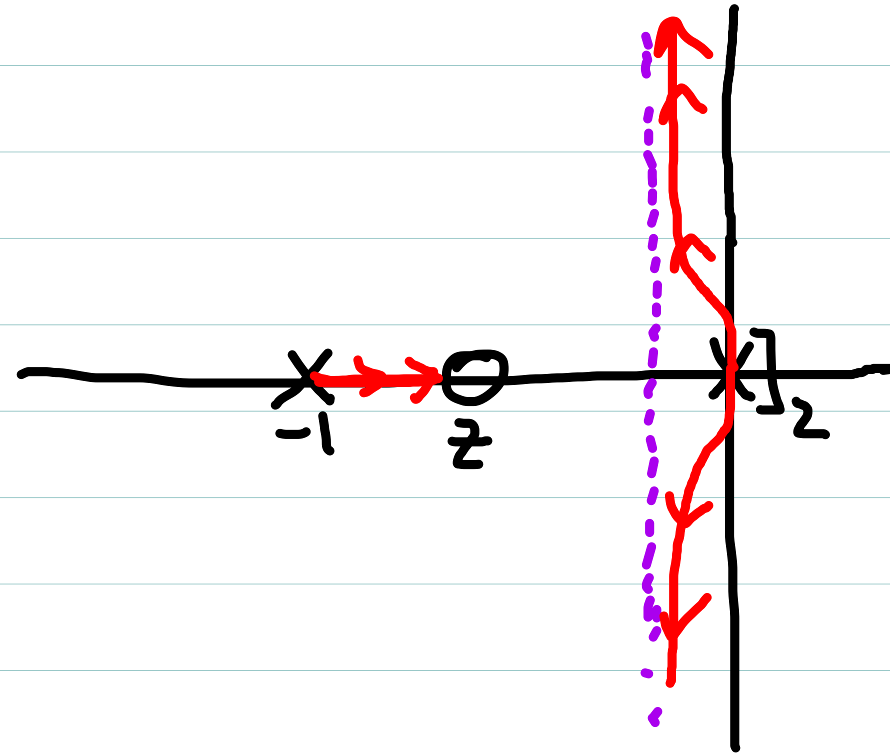
This is where we originally started our investigation

$$L(s) = \frac{K}{s^2(s+1)}$$

We can get the desired pole at -2, but will inevitably have poles of $T(s)$ in RHP

With instead $H(s) = K(s-z)$





$z < -1 \Rightarrow \alpha_e = \pm 90^\circ$

$\sigma_a = \frac{-1}{2}(1+z) > 0$

$0 < z < -1$

$\Rightarrow \sigma_a < 0$

So, with $H(s) = K(s-z)$ we can stabilize the system as long as $|z| < 1$ (which would agree with a Nyquist/phase margin analysis)

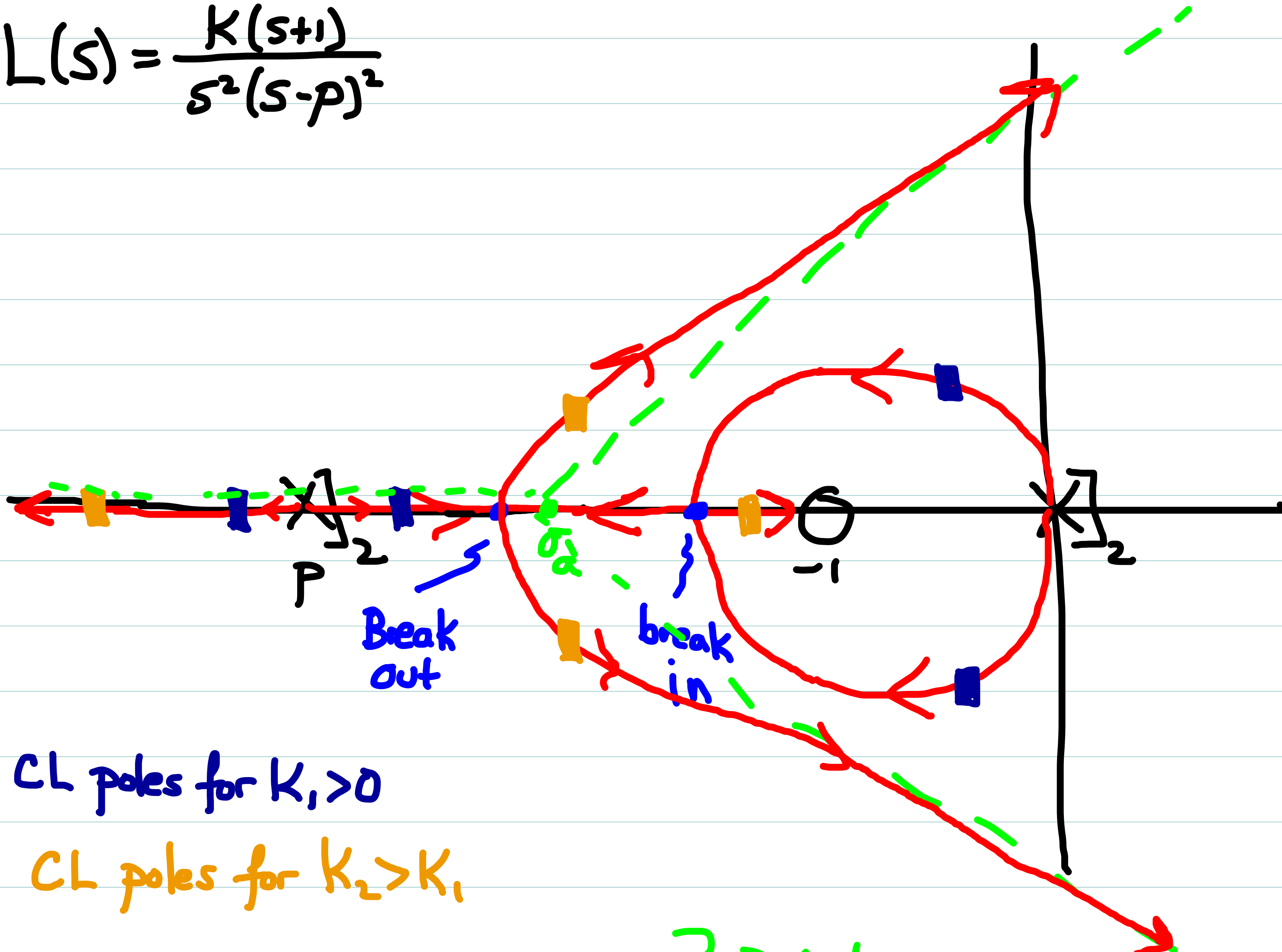
But we would have to accept a real pole > -1, and moreover this pole would not be dominant

An implementable compensator which could allow a real dominant CL pole near -2 would be

$$H(s) = K\left[\frac{(s+1)^2}{(s-p)^2}\right]$$

which has an interesting locus (next page)

$$L(s) = \frac{K(s+1)}{s^2(s-p)^2}$$



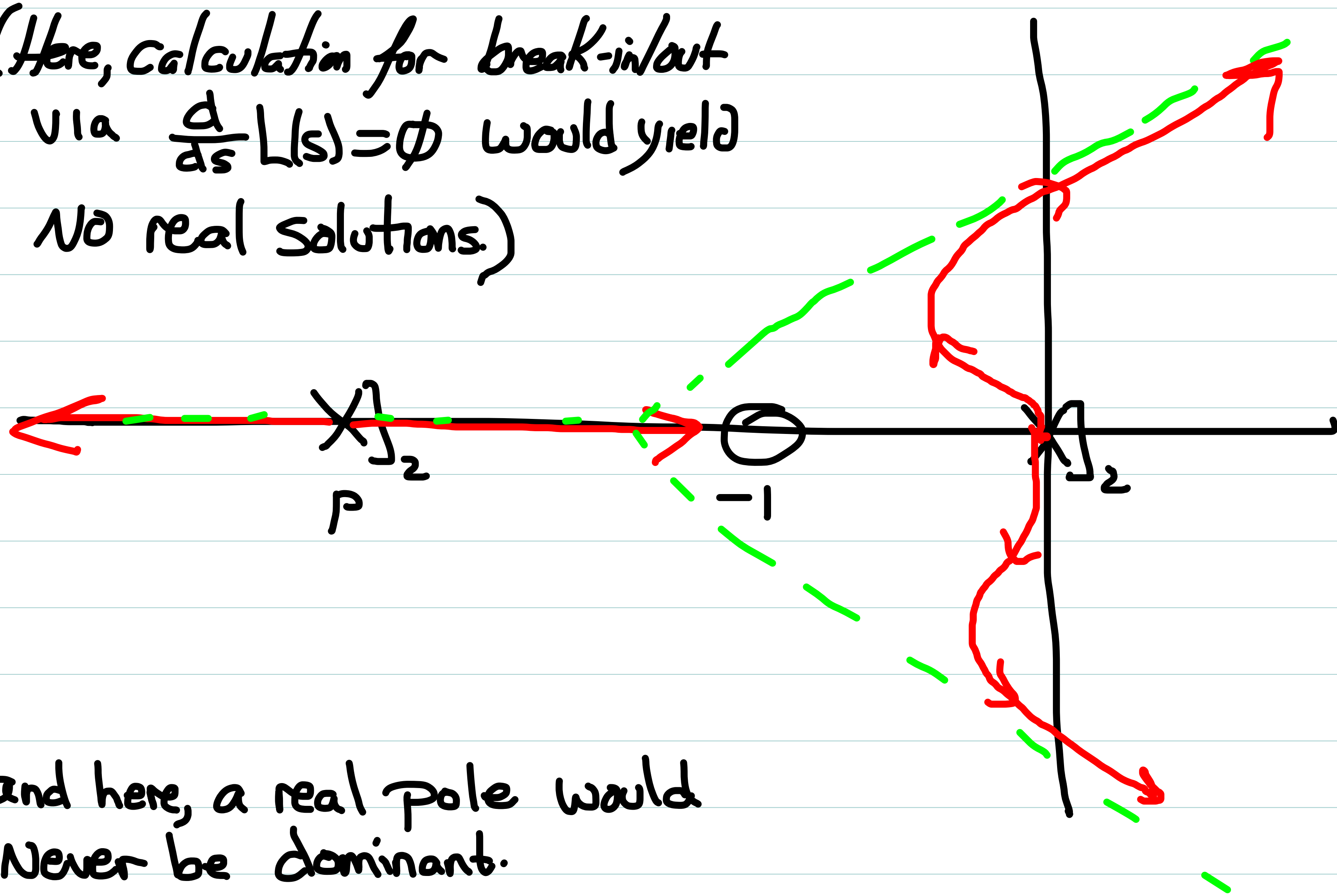
■ CL poles for $K_1 > 0$

■ CL poles for $K_2 > K_1$

$$\sigma_a = \frac{2p+1}{3}$$

However, depending on exact value of p, this is also possible:

(Here, calculation for break-in/out via $\frac{d}{ds}L(s)=\emptyset$ would yield No real solutions.)



and here, a real pole would Never be dominant.

Comments on root locus method

⇒ Rules are not determinative; there may be many locus shapes consistent with calculations (although Matlab rlocus command will show you an exact plot).

⇒ Cannot adapt method to account for effects of time delay

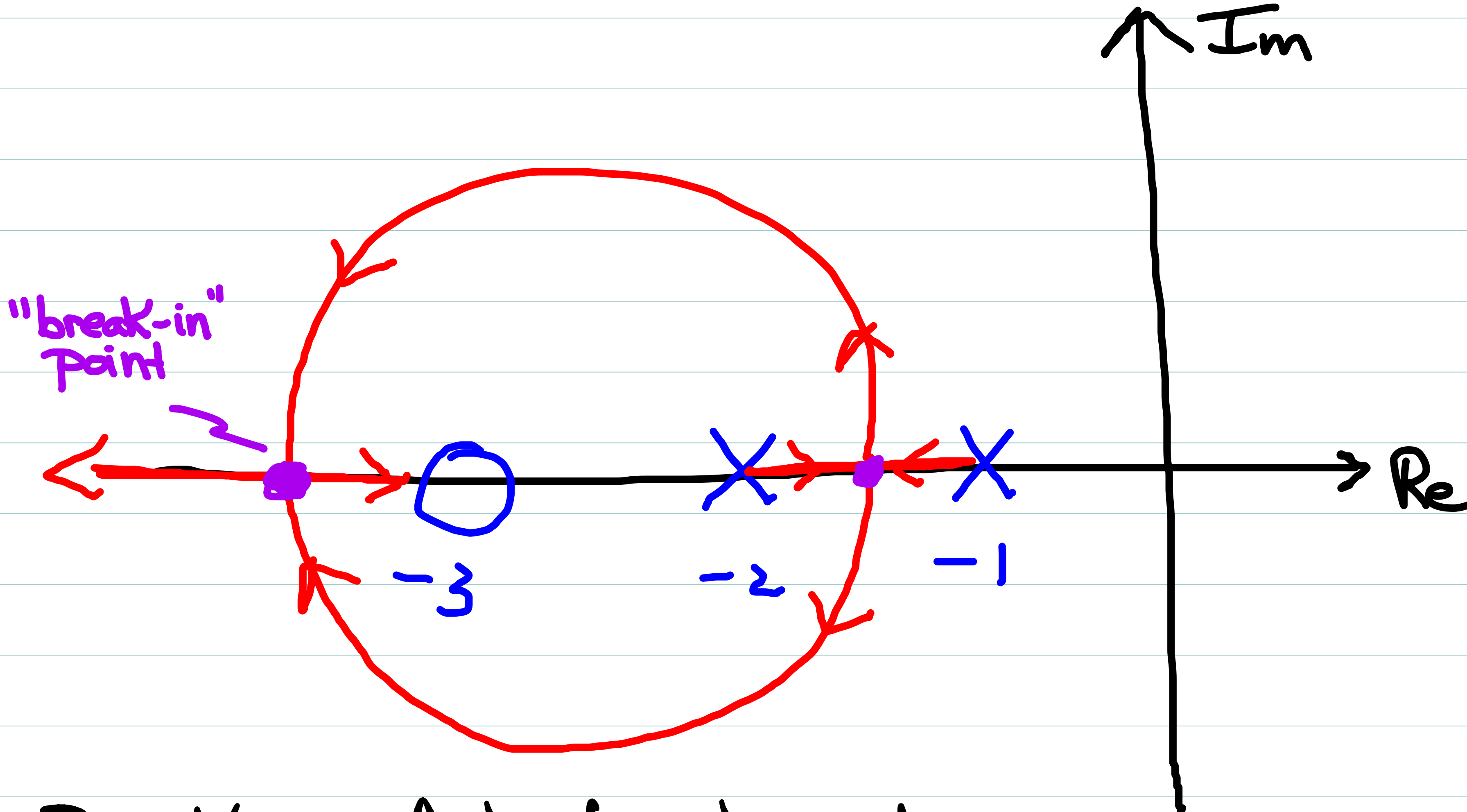
⇒ Can adapt method only for very simple kinds of robustness analysis.

⇒ Bode/Nyquist methods preferred in professional practice.

⇒ But root locus does provide useful additional insights which are not available using freq. methods

⇒ Familiarity with both gives "best of both worlds"

Example #6, cont



Break-in, like break-out satisfies

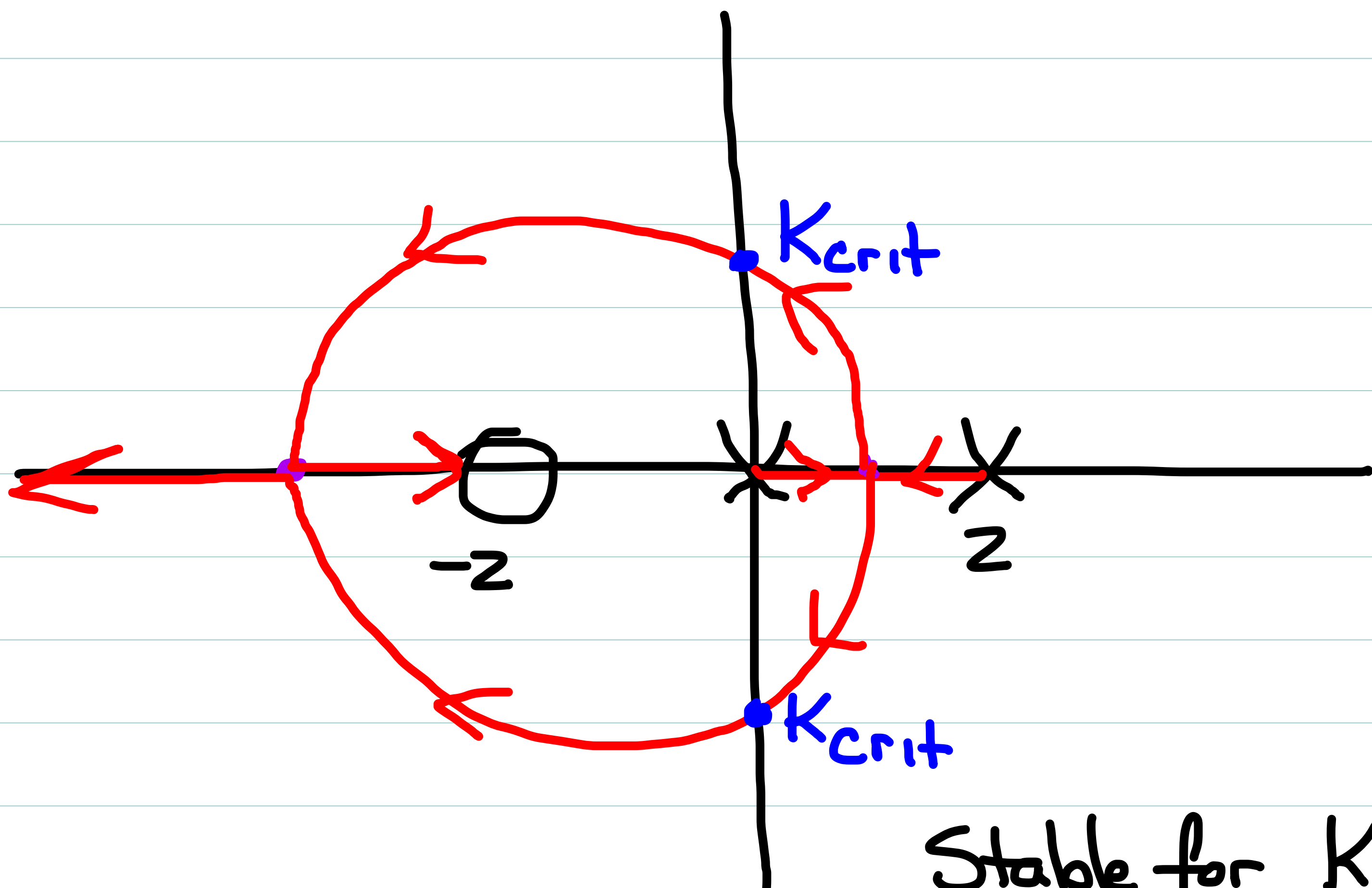
$$\frac{d}{ds} h(s) = 0$$

Example #7

$$L(s) = \frac{K(s+2)}{s(s-2)}$$

Similar analysis to above

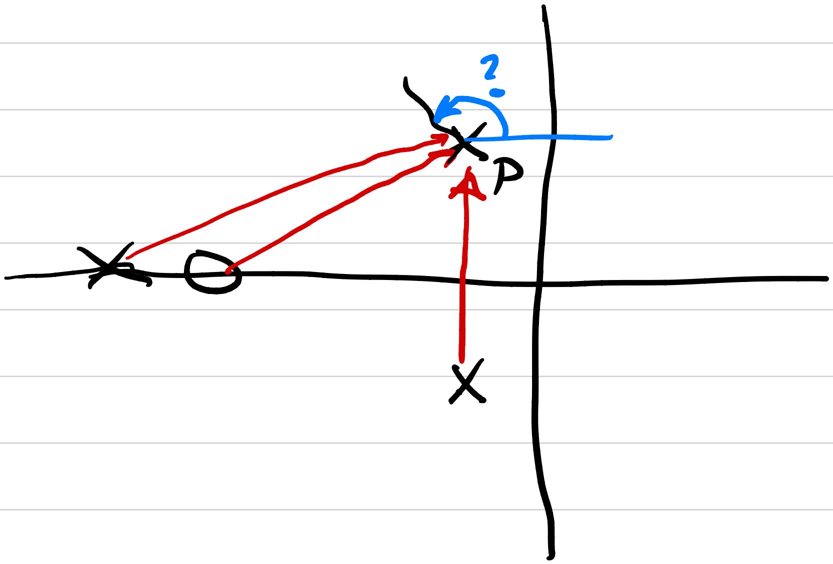
Same pole-zero pattern, shifted to the Right.



Stable for $K > K_{crit}$

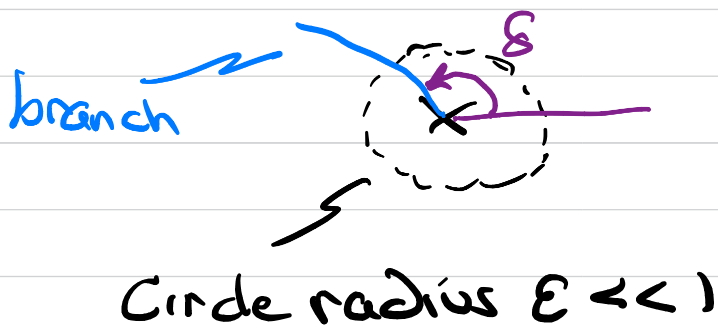
Root Locus: Add'l Rules

Angle of departure/arrival: complex pole/zero



We know branch will start at complex pole here, but what direction does this branch leave from the pole?

Consider a tiny circle around complex pole p



Branch leaving pole will puncture this tiny circle at angle δ

This is the "angle of departure"

What is departure angle δ?

Let $\quad L'(s) = [(s-p)L(s)]$ (remove pole being examined from $L(s)$)

Then

$$\measuredangle L(s) = \measuredangle L'(s) - \measuredangle(s-p) \quad \text{for any } s$$

$$= (1+2\ell)180^\circ \quad (\text{if } s \text{ is a CL pole})$$

Since ε-circle is tiny compared to distance to other poles/zeros in $L'(s)$

Then

$$\measuredangle L'(s)\Big|_{\varepsilon\text{-circle}} \approx \measuredangle L'(p)$$

hence

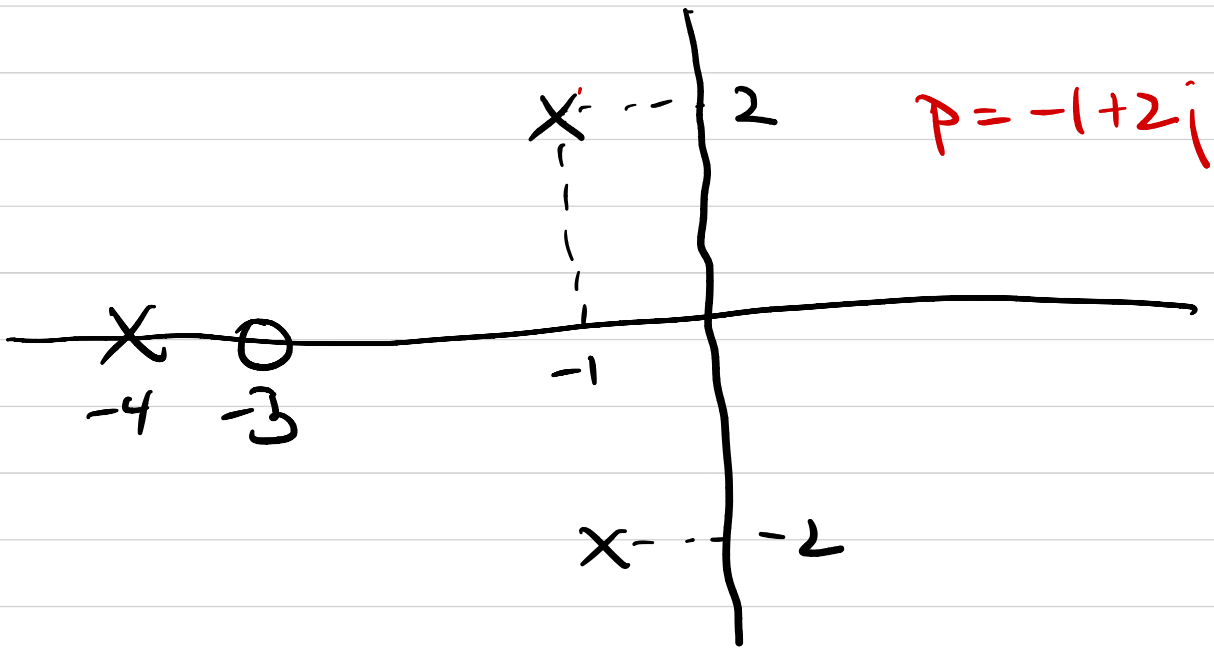
$$\measuredangle L(s)\Big|_{\varepsilon\text{-circle}} = \measuredangle L'(p) - \delta$$

$\delta = \measuredangle(s-p)$ on ε-circle.

$$\Rightarrow \quad \delta = \measuredangle L'(p) - (1+2\ell)180^\circ$$

choose ℓ so $\delta \in [-180^\circ, +180^\circ]$

Example



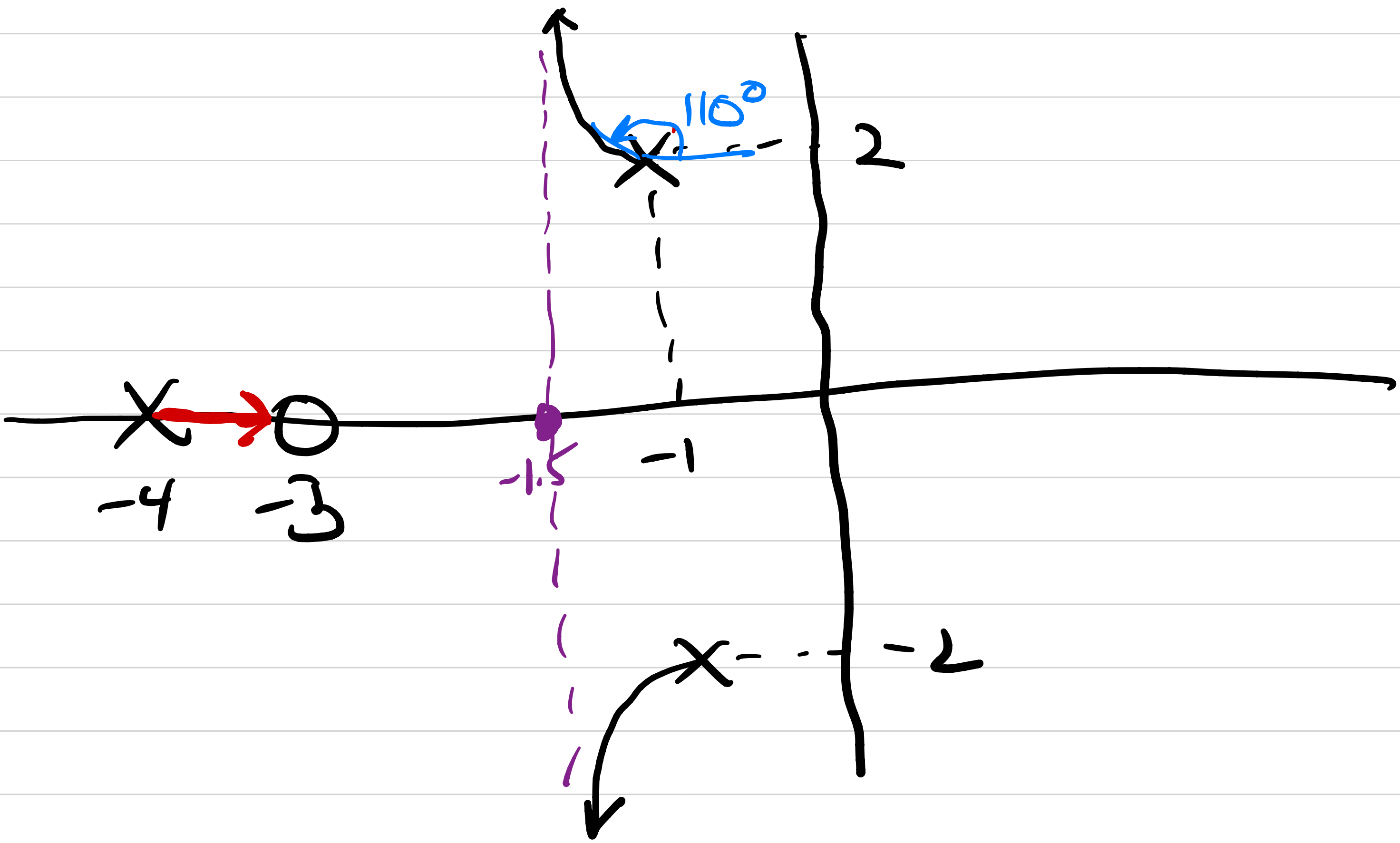
$$L'(s) = \frac{K(s+3)}{(s+4)(s-\bar{p})} \quad \text{here}$$

$$\angle L'(p) = \angle(2+2j) - \angle(3+2j) - \angle 4j$$

$$= 45^\circ - \tan^{-1}(2/3) - 90^\circ$$

$$\cong -78.7^\circ$$

$$\delta = -78.7^\circ + 180^\circ = \underline{\underline{101.3^\circ}}$$



=

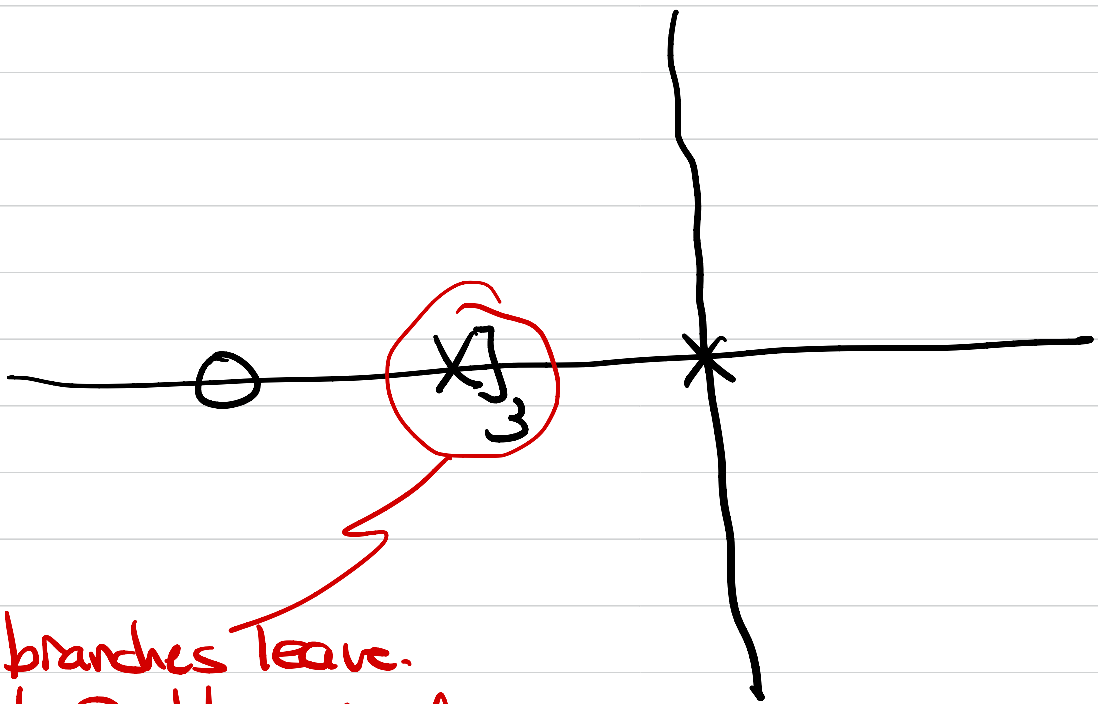
Similar calculations apply to arrival angle at complex zero.

Here $L'(s) = \left[\left(\frac{1}{s-z}\right) L(s)\right]$ (remove zero from $L(s)$)

and arrival angle δ satisfies

$$\delta = (1+2l)180° - \measuredangle L'(z)$$

Angle of departure/arrival multiple poles



3 branches leave.
One to Right on real axis, what about other 2?

Argument is similar. Suppose p repeated q times.

Let $L'(s) = [(s-p)^q L(s)]$ (remove repeated pole p from $L(s)$)

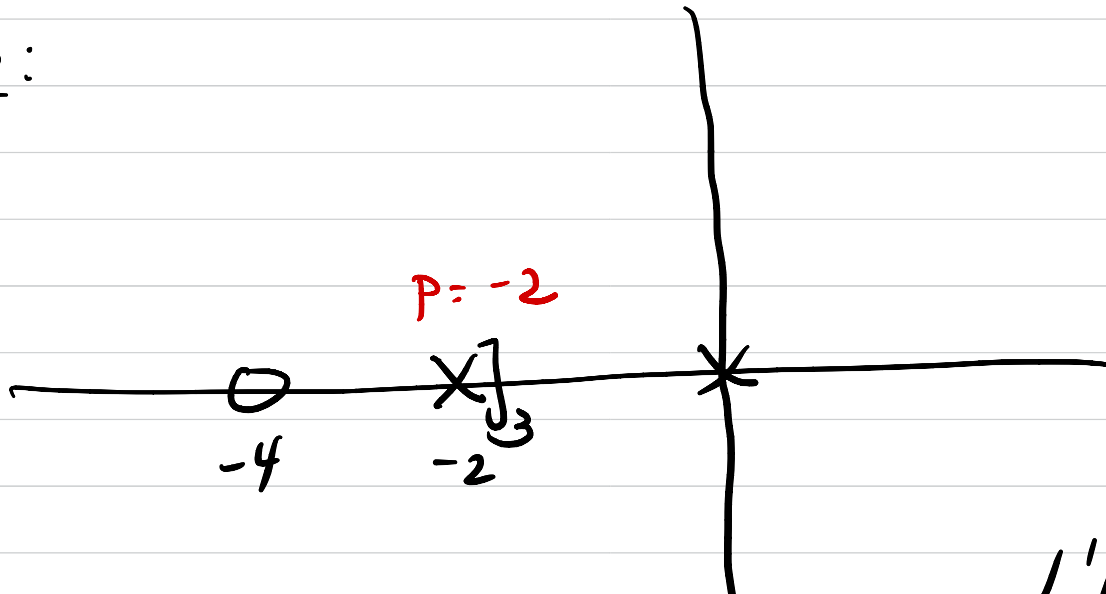
and consider tiny ε-circle around p.

Same calculation now gives

$$\delta = \frac{1}{q}\left[\angle L'(p) - (1+2\ell)180^\circ\right]$$

Defines q unique angles

Example:



$$L'(s) = \frac{K(s+4)}{s} \text{ here}$$

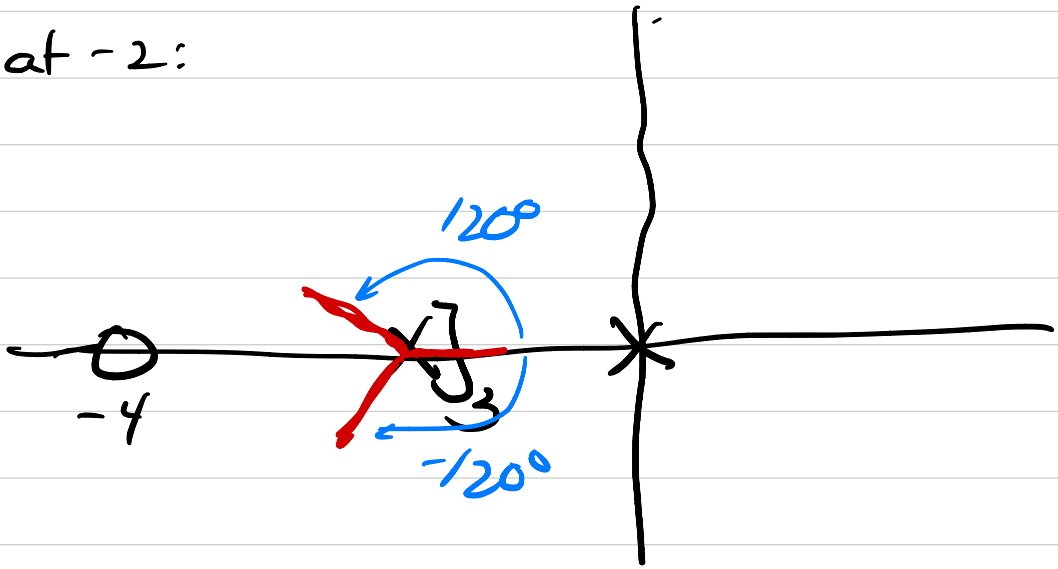
$$\measuredangle L'(p) = \measuredangle\left(\frac{2}{-2}\right) = -180^\circ$$

with $\ell = -1$, $\delta = \frac{1}{3}\left[-180^\circ - (-180^\circ)\right] = 0^\circ$

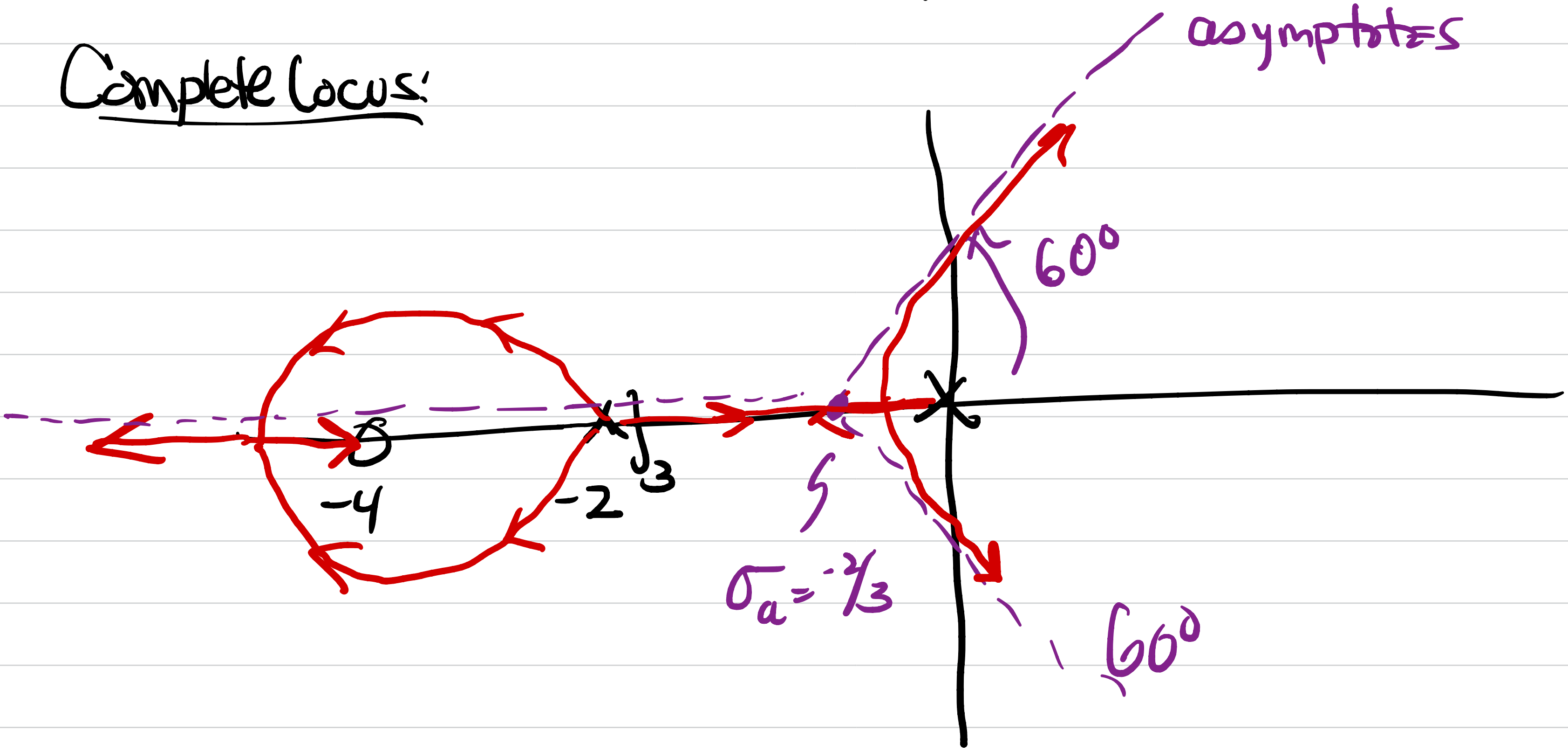
$\ell = 0$, $\delta = \frac{1}{3}\left[-180^\circ - 180^\circ\right] = -120^\circ$

$\ell = -2$, $\delta = \frac{1}{3}\left[-180^\circ - (-540^\circ)\right] = +120^\circ$

Branches at -2:



Complete Locus:



Again, similar considerations apply to calculating arrival angles for branches at a repeated zero:

Let $$\angle'(s) = \left[\frac{1}{(s-z)^q}\angle(s)\right]$$

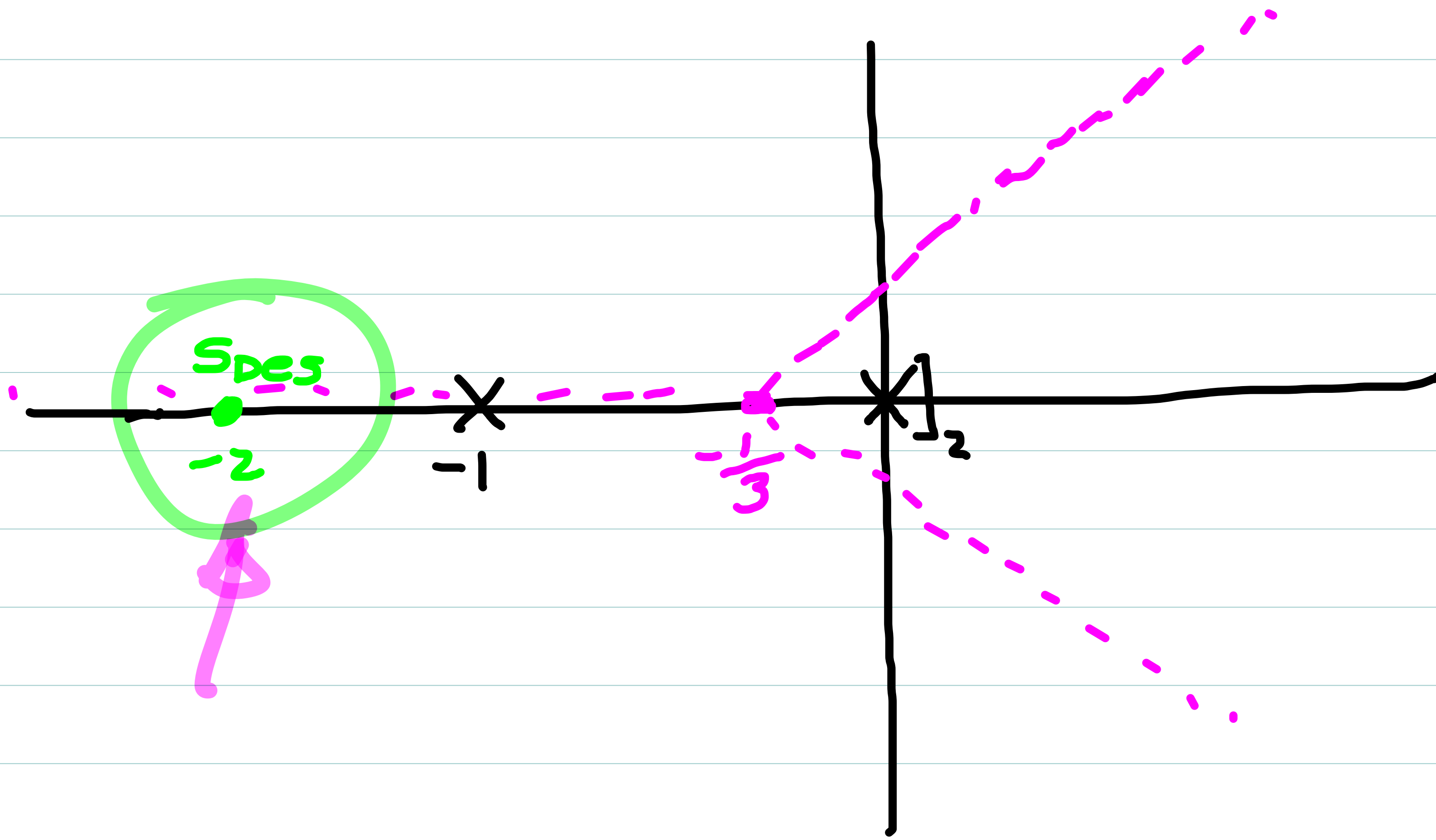
and then

$$\delta = \frac{1}{q}\left[(1+2l)180^\circ - \angle'(z)\right] \quad \Big| \quad q \ \underline{\text{unique}} \ \text{directions}$$

Example #3

This is where we originally started our investigation

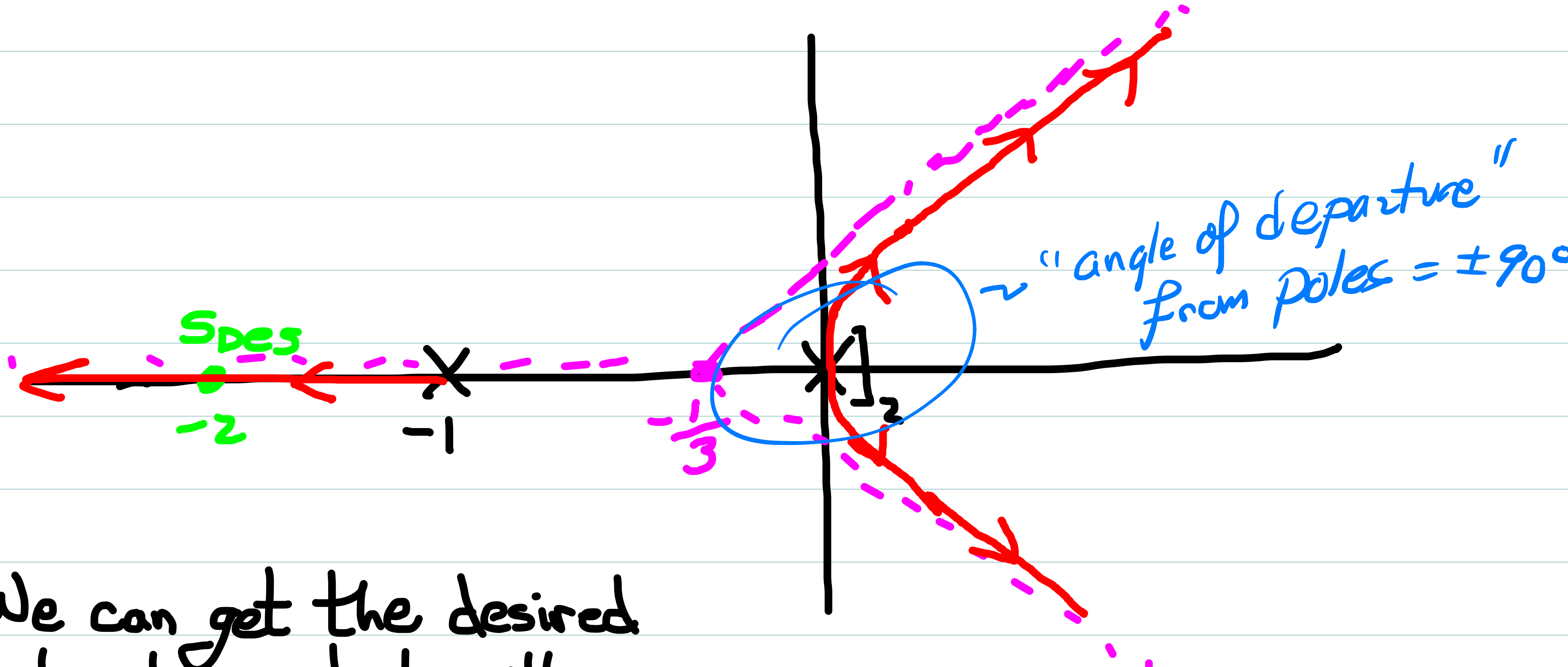
with $H(s)=K$, $L(s) = \dfrac{K}{s^2(s+1)}$



Example #3

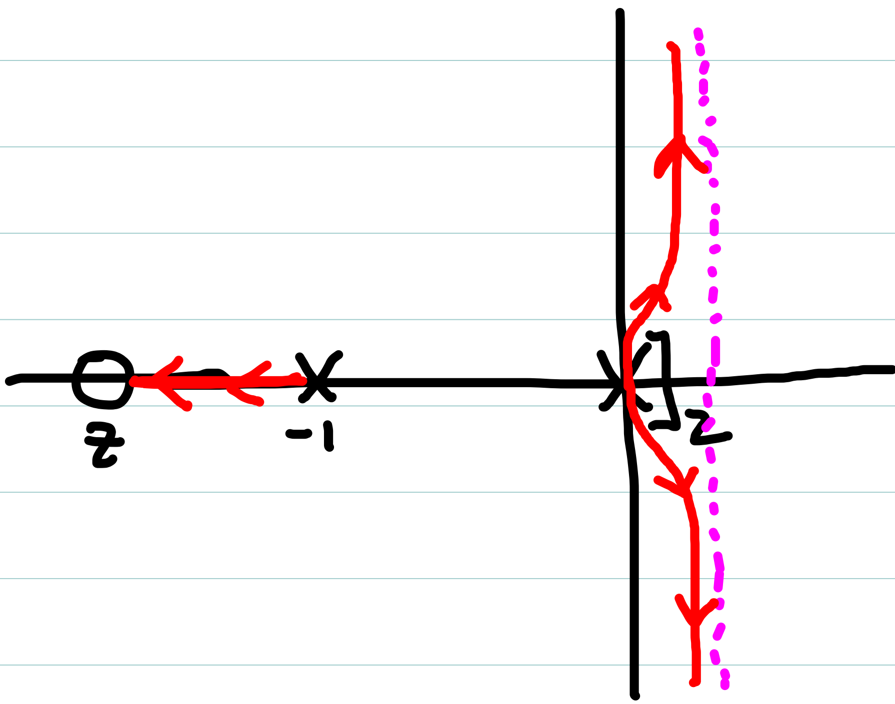
This is where we originally started our investigation

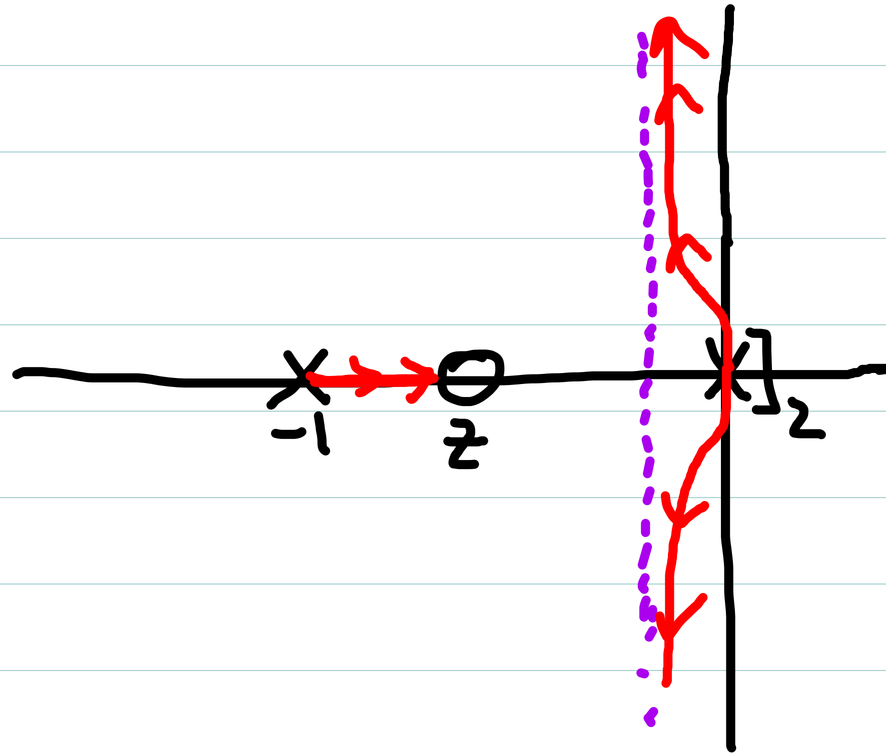
$$L(s) = \frac{K}{s^2(s+1)}$$



We can get the desired pole at -2, but will inevitably have poles of T(s) in RHP

With instead $H(s) = K(s-z)$ (PD compensator)





$z < -1 \Rightarrow \alpha_e = \pm 90^\circ$

$\sigma_a = \frac{-1}{2}(1+z) > 0$

$0 < z < -1$

$\Rightarrow \sigma_a < 0$

So, with $H(s) = K(s-z)$ we can stabilize the system as long as $|z| < 1$ (which would agree with a Nyquist/phase margin analysis)

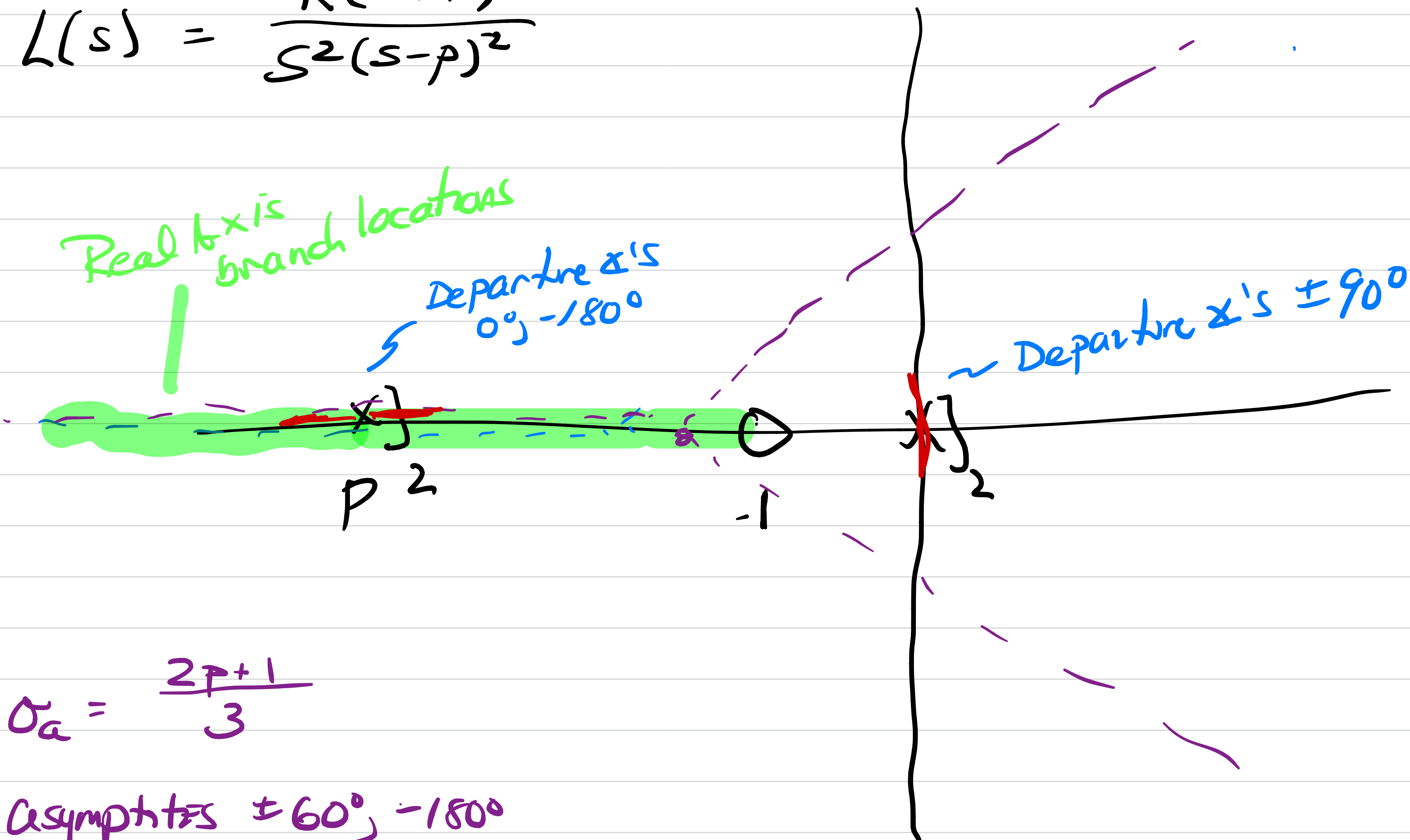
But we would have to accept a real pole > -1, and moreover this pole would not be dominant

An implementable compensator which could allow a real dominant CL pole near -2 would be

$$H(s) = K\left[\frac{(s+1)^2}{(s-p)^2}\right]$$

which has an interesting locus (next page)

$$L(s) = \frac{K(s+1)}{s^2(s-p)^2}$$

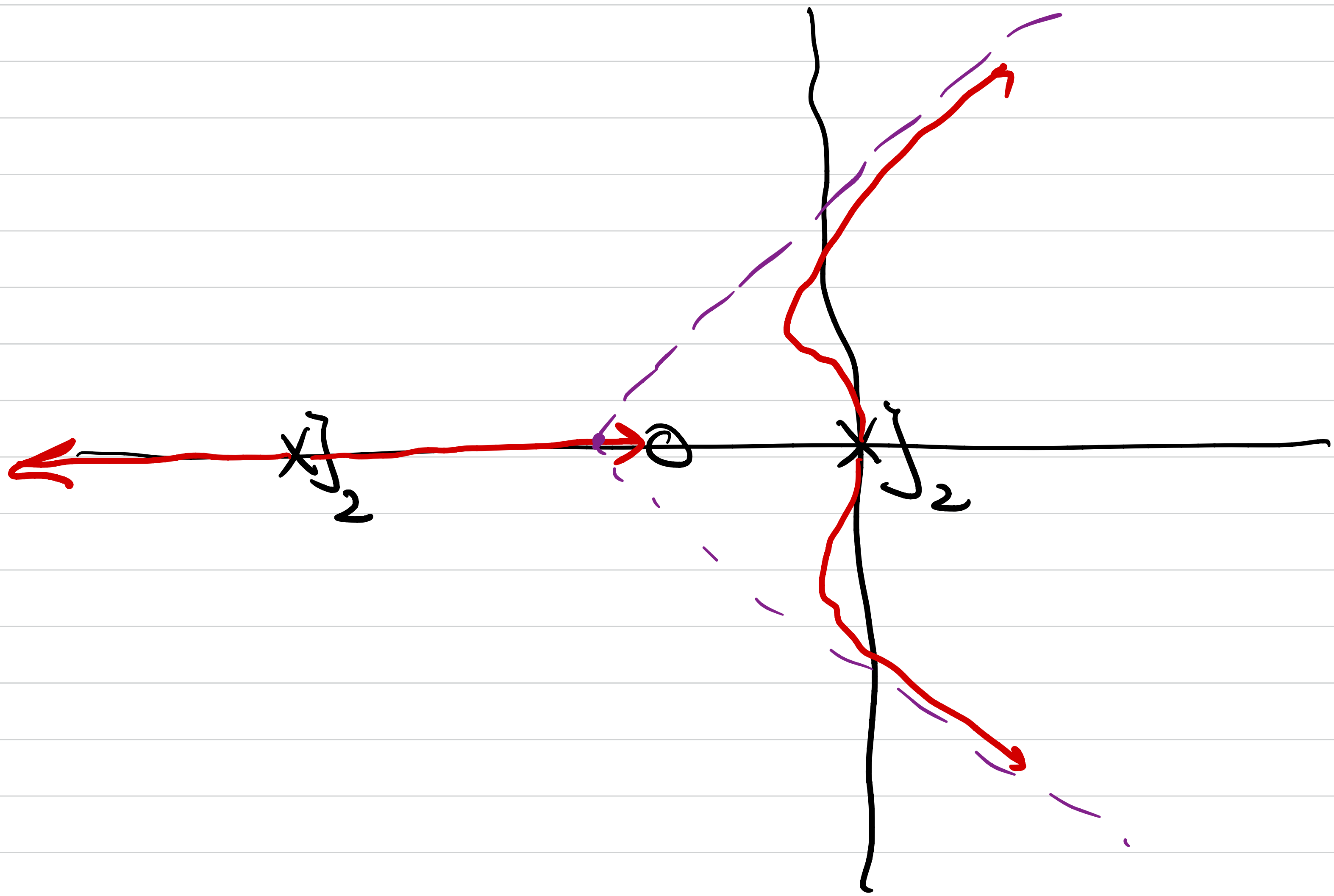


$$\sigma_a = \frac{2p+1}{3}$$

asymptotes $\pm 60°$, $-180°$

$(n - m = 3 \text{ here})$

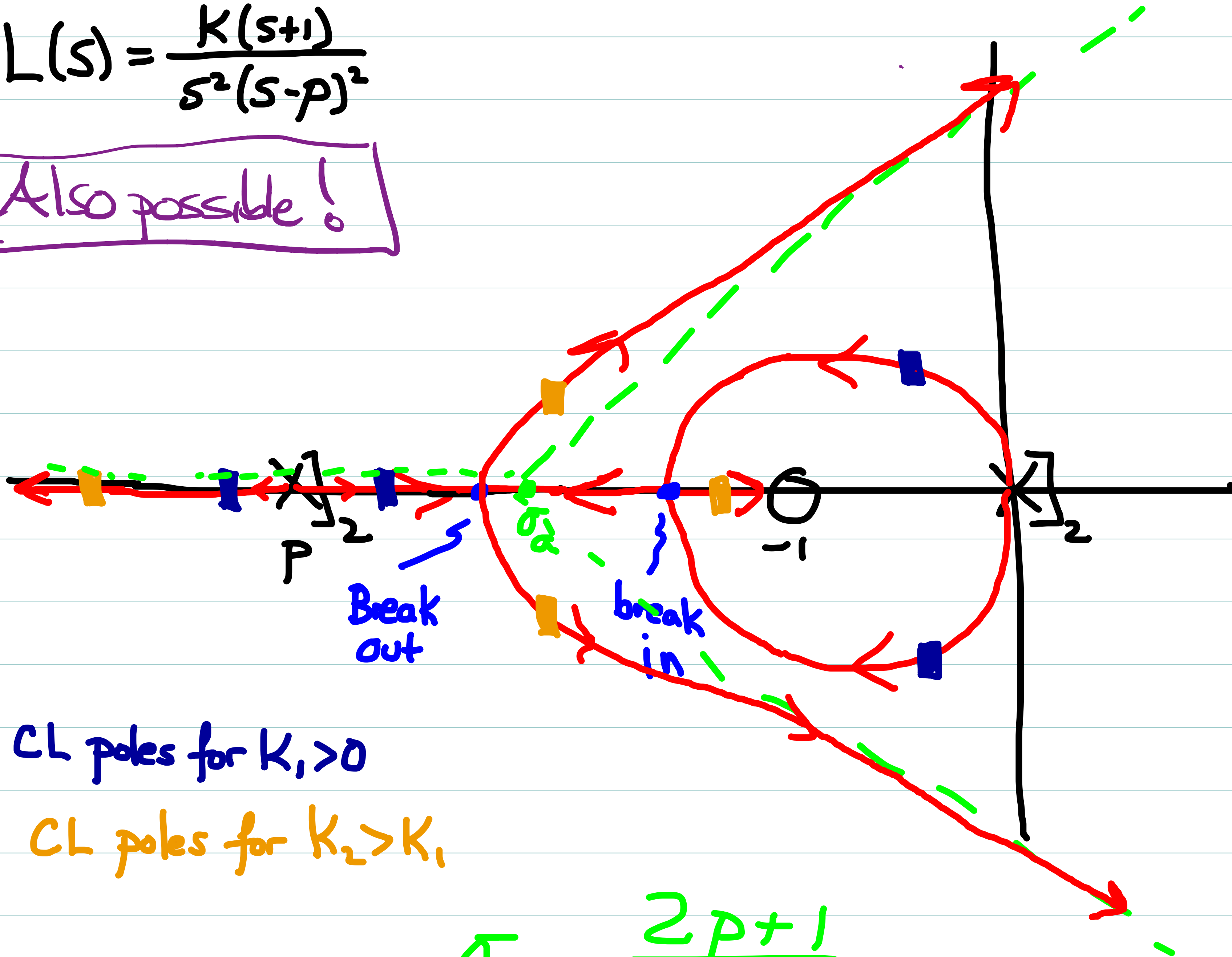
So, we could have this:



<u>Simplest</u> shape consistent w/rules, but not the only one!

$$L(s) = \frac{K(s+1)}{s^2(s-p)^2}$$

Also possible!



■ CL poles for $K_1 > 0$

■ CL poles for $K_2 > K_1$

$$\sigma_a = \frac{2p+1}{3}$$

- Which shape we get is highly dependent on exact location of poles in $H(s)$
- Simple rules are insufficient to identify which shape we get

(But note the break-in calculations would show the differences, however they require factoring a 3rd order polynomial here)

- 2nd shape would get close to our requirements (dominant CL pole at -2), 1st shape would not.

Comments on root locus method

⇒ Rules are not determinative; there may be many locus shapes consistent with calculations (although Matlab rlocus command will show you an exact plot).

⇒ Cannot adapt method to account for effects of time delay

⇒ Can adapt method only for very simple kinds of robustness analysis.

⇒ Bode/Nyquist methods preferred in professional practice.

⇒ But root locus does provide useful additional insights which are not available using freq. methods

⇒ Familiarity with both gives "best of both worlds"

Controller Implementation

Recall we have shown

$$U(s) = H(s)E(s) = \left[c_0 + \sum_i \frac{c_i}{s-a_i} \right] E(s)$$

where a_i are poles of $H(s)$, and $c_0, c_1, c_2 \ldots$ are PFE coefs, with $c_0 = \emptyset$ if $\rho(H) > \emptyset$ ($H(s)$ has more poles than zeros).

Let $X_i(s) = \left[\frac{1}{(s-a_i)} \right] E(s)$

then $u(t) = c_0 e(t) + \sum c_i x_i(t)$

where $e(t) = y_d(t) - y(t)$, and $x_i(t)$ are sol'ns of

$$\dot{x}_i(t) = a_i x_i(t) + e(t)$$

The discrete time stepping under which the computer and sensor/actuator electronics operate mean that $u(t)$ will be computed only at integer multiples of the sample interval, T_s.

i.e. at $t_K = KT_s$, for $K = \emptyset, 1, 2, \ldots$

Let $u_K = u(t_K) = u(KT_s)$, and $e_K = e(t_K) = e(KT_s)$

From above:

$$u_K = c_0 e_K + \sum c_i x_i(t_K)$$

We need to know how to evaluate $x_i(t_K)$, i.e. sol'n of

$$\dot{x}_i(t) = a_i x_i(t) + e(t) \quad \text{evaluated at } t = t_K$$

Focus on just a single one of these eq'ns, since they are identical except for coefs a_i:

$$\dot{x}(t) = a x(t) + \varepsilon(t)$$

(Let $\varepsilon(t) = e(t)$ here, to avoid confusion with e^{at})

<u>Assume</u> $\varepsilon(t)$ is a step of size ε_0, and $x(\phi) = x_0$

Then $X(s) = \dfrac{x_0}{s-a} + \dfrac{\varepsilon_0}{s(s-a)}$

$$\Rightarrow x(t) = e^{at} x_0 + \frac{1}{(-a)}\left(1 - e^{at}\right)\varepsilon_0$$

<u>Note</u>: $e(t)$ will <u>Not</u> generally be a step, even if $y_d(t)$ is!

But, the above is a useful intermediate result, as we will see Next.

Here

$$x(t) = e^{at}x_0 + \left(\frac{-1}{a}\right)(1 - e^{at})\varepsilon_0$$

So $x(T_s) = e^{aT_s}x_0 + \left(\frac{-1}{a}\right)(1 - e^{aT_s})\varepsilon_0$

Let $\boxed{\alpha = e^{aT_s}}$, $\boxed{\beta = (1-\alpha)/(-a)}$

Then $x(T_s) = \alpha x_0 + \beta\varepsilon_0$

Let $x(T_s) = x_1 \Rightarrow \boxed{x_1 = \alpha x_0 + \beta\varepsilon_0 = x(t_1)}$

Now, how does this help generally?

Sampling of output at discrete times $t_K = KT_s$ means that error $e(t)$ will have a staircase graph

i.e. $e(t)$ will be constant with level ε_K on the interval $t_K \leq t < t_{K+1}$.

Note that at t_0, $e(t)$ does look like a step.

So, it is true for first sample interval that

$$X(t_1) = x_1 = \alpha X_0 + \beta \varepsilon_0 \qquad \text{(as above)}$$

But what about subsequent time steps??

Exploit time invariance: when solving constant coeff DE's, the time called zero is <u>arbitrary</u>. All that matters is the initial cond'n and the time <u>elapsed</u> since initial time.

So, to get sol'n for next sample time t_2, we can "reset" the zero time to t_1, and use $X(t_1)$ as new initial cond'n.

Then, from new zero time $t = t_1$, error $e(t)$ looks like a step of height (initially) ε_1, so:

$$X(t_2) = X_2 = \alpha X_1 + \beta \varepsilon_1 \quad \text{by same logic as above}$$

We can repeat this trick for all subsequent t_K:

$$\boxed{x_{K+1} = \alpha x_K + \beta \varepsilon_K}$$

where $x_K = x(t_K) = x(KT_s)$

$$\boxed{\alpha = e^{aT_s}, \quad \beta = (1-\alpha)/(-a)}$$

$\varepsilon_K = e(t_K)$ (error at K^{th} sample time)

We have thus shown that:

$$\boxed{x(t_{K+1}) = \alpha x(t_K) + \beta e(t_K)}$$

is an <u>iterative</u> algorithm for generating the <u>exact</u> sol'n for $x(t)$ at each of the sample times t_K, given the staircase structure of $e(t)$.

So generally:

$$u(t_K) = u_K = C_0 e(t_K) + \sum c_i x_i(t_K)$$

Where Each $x_i(t_K)$ is computed iteratively using

$$x_i(t_{K+1}) = \alpha_i x_i(t_K) + \beta_i e(t_K)$$

where $\alpha_i = \exp[a_i T_s]$, $\beta_i = \left[\frac{1-\alpha_i}{(-a_i)}\right]$

and T_s is the sample interval.

Real-time implementation

$$u(t_K) = u_K = c_0 e(t_K) + \sum c_i x_i(t_K) \quad , \quad t_K \text{ is } K^{th} \text{ update time}$$

Where Each $x_i(t_K)$ is computed iteratively using

$$x_i(t_{K+1}) = \alpha_i x_i(t_K) + \beta_i e(t_K)$$

and $\Rightarrow \alpha_i = \exp[a_i T_s]$, $\beta_i = \left[\frac{1-\alpha_i}{(-a_i)}\right]$

$\Rightarrow T_s$ is the sample interval (inverse of sample rate)

$\Rightarrow a_i$ are poles of $H(s)$

$\Rightarrow c_0, c_1, c_2 \ldots$ are PFE coefs of $H(s)$

$\Rightarrow e(t_K) = y_d(t_K) - y(t_K)$

Matlab code

Simple case: $H(s) = K \Rightarrow u(t) = Ke(t)$

```
function u=control(yd,y)

% define K (number!)
K=...

% compute u
e = yd-y;
u = K*e;

end
```

H(s) with 1 pole

$$H(s) = K\frac{(s-z_c)}{(s-p_c)} = c_0 + \frac{c_1}{(s-p_c)}$$

```
function u=control(yd,y)

% define c0, c1, alpha, beta (as numbers!)
c0=...
c1=...
alpha=...
beta=...

% compute u
e = yd-y;
u = c0*e+c1*x;

% update x
x = alpha*x+beta*e;

end
```

alpha = exp[p_c*T] here, and

Beta = (1-alpha)/(-p_c)

$H(s)$ with 1 pole

$$H(s) = K\frac{(s-z_c)}{(s-p_c)} = c_0 + \frac{c_1}{(s-p_c)}$$

```
function u=control(yd,y)

% define c0, c1, alpha, beta
c0=...
c1=...
alpha=...
beta=...

% compute u
e = yd-y;
u = c0*e+c1*x;

% update x
x = alpha*x+beta*e;

end
```

Unfortunately, won't work as written!

We need the function to "remember" the values of x between calls.

Remember: Matlab functions (like C/C++ functions) have their own, private workspace (storage) for their variables, which is separate from the main workspace (main function).

Local variables in functions are <u>cleared</u> after the function runs.

Can prevent this clearing by declaring the variable to be "persistent" in Matlab ("static" in C/C++).

$$H(s) = K\frac{(s-z_c)}{(s-p_c)} = c_0 + \frac{c_1}{(s-p_c)}$$

```
function u=control(yd,y)

persistent x

% define c0, c1, alpha, beta
c0=...
c1=...
alpha=...
beta=...

% compute u
e = yd-y;
u = c0*e+c1*x;

% update x
x = alpha*x+beta*e;

end
```

$$H(s) = K\frac{(s-z_c)}{(s-p_c)} = c_0 + \frac{c_1}{(s-p_c)}$$

```
function u=control(yd,y)

persistent x

% define c0, c1, alpha, beta
c0=...
c1=...
alpha=...
beta=...

% compute u
e = yd-y;
u = c0*e+c1*x;

% update x
x = alpha*x+beta*e;

end
```

Still won't work!

x needs to be <u>initialized</u>, but only the 1^{st} time the function is called.

Matlab initializes a persistent variable as an empty array the first time the function is run

We can test for this, and set initial value of x to our pleasing: "isempty" function for test

Note: simplest to initialize x to zero, unless there is a compelling reason not to (very rare)

$$H(s) = K\frac{(s-z_c)}{(s-p_c)} = c_0 + \frac{c_1}{(s-p_c)}$$

```
function u=control(yd,y)

persistent x

if isempty(x)
   x=0;
end

% define c0, c1, alpha, beta
c0=...
c1=...
alpha=...
beta=...

% compute u
e = yd-y;
u = c0*e+c1*x;

% update x
x = alpha*x+beta*e;

end
```

"ZOH" zero order hold

initialize x first time

Works!

All our mathematical analysis ultimately boils down to 4 "magic numbers" that we plug into this standard template.

$$H(s) = 30\left[\frac{s+3}{s+9}\right] = 30 - \frac{180}{s+9}, \quad T_s = 0.1 \ (10\,\text{Hz})$$

```
function u=control(yd,y)

persistent x

if isempty(x)
   x=0;
end

% define c0, c1, alpha, beta
c0 = 30;
c1 = -180;
alpha = 0.4066;
beta = 0.0659;

% compute u
e = yd-y;
u = c0*e+c1*x;

% update x
x = alpha*x+beta*e;

end
```

Implementation of pole at origin

If $p_c = \phi$ (comp pole at origin), then clearly

$$\alpha = \exp[\phi T_s] = 1$$

in the implementation eq'n. However $\beta = \frac{(1-1)}{\phi}$ is indeterminate.

If we look more carefully at $\lim_{p_c \to \phi} \left[\frac{1 - \exp[p_c T_s]}{-p_c} \right]$

This yields the correct value $\beta = T_s$ for this case.

Thus for $\quad \dot{x}(t) = e(t)$

we have $\quad x(t_{k+1}) = x(t_k) + T_s\, e(t_k)$

i.e. $$x_{k+1} = x_k + T_s e_k$$

Real-time implementation

$$u(t_K) = u_K = C_0 e(t_K) + \sum C_i x_i(t_K)$$, t_K is K^{th} update time

$t_K = K T_s$

Where each $x_i(t_K)$ is computed iteratively using

$$x_i(t_{K+1}) = \alpha_i x_i(t_K) + \beta_i e(t_K)$$

and $\Rightarrow$ $\alpha_i = \exp[a_i T_s]$, $\beta_i = \left[\frac{1-\alpha_i}{(-a_i)}\right]$

$\Rightarrow$ T_s is the sample interval (inverse of sample rate)
(secs) (Hz)

$\Rightarrow$ a_i are poles of $H(s)$

$\Rightarrow$ $C_0, C_1, C_2 \ldots$ are PFE coefs of $H(s)$

$\Rightarrow$ $e(t_K) = y_d(t_K) - y(t_K)$

$$H(s) = K\frac{(s-z_c)}{(s-p_c)} = c_0 + \frac{c_1}{(s-p_c)}$$

```
function u=control(yd,y)

persistent x

if isempty(x)
   x=0;
end

% define c0, c1, alpha, beta
c0=...
c1=...
alpha=...
beta=...

% compute u
e = yd-y;
u = c0*e+c1*x;

% update x
x = alpha*x+beta*e;

end
```

"ZOH" zero order hold

initialize x first time

Works!

All our mathematical analysis ultimately boils down to 4 "magic numbers" that we plug into this standard template.

$$H(s) = 30\left[\frac{s+3}{s+9}\right] = 30 - \frac{180}{s+9} \quad , T_s = 0.1 \ (10\,\text{Hz})$$

```
function u=control(yd,y)

persistent x

if isempty(x)
   x=0;
end

% define c0, c1, alpha, beta
c0 = 30;
c1 = -180;
alpha = 0.4066;
beta = 0.0659;

% compute u
e = yd-y;
u = c0*e+c1*x;

% update x
x = alpha*x+beta*e;

end
```

precomputed from $\alpha = \exp(\underset{-9}{p_c} * \underset{0.1}{T_s})$

$\beta = \underset{-\frac{1}{9}}{\frac{1}{p_c}}(\alpha - 1)$

Implementation of pole at origin

If $p_c = \phi$ (comp pole at origin), then clearly

$$\alpha = \exp[\phi T_s] = 1$$

in the implementation eq'n. However $\beta = \frac{(1-1)}{\phi}$ is indeterminate.

If we look more carefully at $\lim_{p_c \to \phi} \left[\frac{1 - \exp[p_c T_s]}{-p_c} \right]$

This yields the correct value $\beta = T_s$ for this case.

Thus for $\dot{x}(t) = e(t)$

we have $x(t_{k+1}) = x(t_k) + T_s e(t_k)$

i.e. $x_{k+1} = x_k + T_s e_k$

$$H(s) = K\frac{(s-z_{c1})(s-z_{c2})}{(s-p_{c1})(s-p_{c2})} = C_0 + \frac{C_1}{s-p_{c1}} + \frac{C_2}{s-p_{c2}}$$

```
function u=control(yd,y)

persistent x1 x2

if isempty(x1)
   x1=0;
   x2=0;
end

% define constants
% ...
%


% compute u
e = yd-y;
u = c0*e+c1*x1+c2*x2;

% update x
x1 = alpha1*x1+beta1*e;
x2 = alpha2*x2+beta2*e;

end
```

An implementation with 2 poles in $H(s)$, hence 2 Diff. eq'ns which need to be solved ($x_1(t_k), x_2(t_k)$)

$$H(s) = K\frac{(s-z_{c1})(s-z_{c2})}{(s-p_{c1})(s-p_{c2})} = c_0 + \frac{c_1}{s-p_{c1}} + \frac{c_2}{s-p_{c2}}$$

```
function u=control(yd,y)

persistent x

if isempty(x)
   x=[0;0];
end

% define constants
% ...
%

% compute u
e = yd-y;
u = c0*e+c1*x(1)+c2*x(2);

% update x
x(1) = alpha1*x(1)+beta1*e;
x(2) = alpha2*x(2)+beta2*e;

end
```

Alternate implementation using Matlab arrays

Extention to 3 or more poles in H(s) straightforward following same general pattern.

a(1)
a(2)

b(1)
b(2)

(Can put α, β coefs into arrays ako)

Note that, with two or more poles in $H(s)$, we could write the implementation equations in Matlab in the general form:

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix}$$

$$u = \underline{C} * \underline{x} + D * e$$

Matlab ".*" array multi.

$$\underline{x} = \underline{a} .* \underline{x} + \underline{b} * e$$

Where $\underline{x}$ is vector with all the different x_i variables

$$\underline{a} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_M \end{bmatrix} \qquad \underline{b} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_M \end{bmatrix}$$

M = # poles in $H(s)$

$$\underline{C} = [c_1 \; c_2 \cdots c_M] \qquad D = c_0$$

Even more generally:

$$u = \underline{c} * \underline{x} + D * e$$

$$\underline{x} = A * \underline{x} + \underline{b} * e$$

"State space" representation of (discretized) controller dynamics

where

$$A = \text{diag}\{\alpha_1, \alpha_2, \ldots, \alpha_M\}$$

Special structure for A matrix

$$= \begin{bmatrix} \alpha_1 & & & \emptyset \\ & \alpha_2 & & \\ & & \ddots & \\ \emptyset & & & \alpha_M \end{bmatrix}$$

$\underline{M} \times \underline{M}$ matrix

$(\underline{M} = \#\text{ poles in } H(s))$

```
function u = control(yd,y)
% Stub to illustrate most general form
% of discretized implementation equations

% Note: more efficient to define numerical
% components once, when we initialize x
  persistent x A B C D
  if isempty(x)
    A = []; % square matrix
    B = []; % column vector
    C = []; % row vector
    D = []; % scalar
        x = zeros(size(A,1),1);
        % Note: one state (xi variable) for each row of A
        %  (equivalently, for each pole in H(s))
  end

% Do the actual calculations
  e = yd - y;
  u = C*x + D*e;
  x = A*x + B*e;

end
```

fill in NUMBERS for A, B, C, D

x = 0 initially

Matrix-vector version of controller calculations = Different control strategies correspond to Different numerical values for matrices A, B, C, D

"State Space" Models

The diagonal state space form of the discretized controller equations is mirrored in the form of the continuous implementation eqns:

i.e. $$u(t) = c_0 e(t) + \sum_{i=1}^{M} c_i x_i(t)$$

$$\dot{x}_i(t) = a_i x_i(t) + e(t) \qquad i = 1, \ldots, M$$

$$\Rightarrow \qquad u(t) = C\underline{x}(t) + De(t)$$
$$\dot{\underline{x}}(t) = A\underline{x}(t) + Be(t)$$

where $C = [c_1 \; c_2 \; \ldots \; c_M]$ $\qquad D = c_0$

$$A = \text{diag}\{a_1, a_2, \ldots, a_m\} \quad \text{and}$$

$$B = [1 \;\; 1 \;\; 1 \; \cdots \; 1]^T \quad \text{(column vector)}$$

- This diagonal form for the state-space equations for $H(s)$ (and its discretization) is known as the "modal form".
- However, it is not the only possible state-space form we could use for $H(s)$ and its discretization.
- We have already seen that there are many (infinitely!) different state-space models for a given transfer fn.
- One other form we have examined is the "companion form", where the values in A, B, C, D come from coefs of numerator + denom. polys of $H(s)$, instead of pole locations and residues.

Companion form (review)

$$H(s) = \frac{b_{n-1}s^{n-1} + \cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0}$$

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & 1 \\ -a_0 & -a_1 & \cdots & \cdots & \cdots & -a_{n-1} \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$$C = [\, b_0 \;\; b_1 \;\; \cdots \;\; b_{n-1} \,]$$

$=$

Note that $A_1 = A^T,\ B_1 = C^T,\ C_1 = B^T$

defines an alternate, equivalent, companion form.

General State Space Models

- Modal and companion forms are two common state-space models for a transfer function – again, there are an infinite number of other possible models
- Discretization is easy for modal form, so why would we want to consider alternate forms?

  ⇒ When $H(s)$ has repeated or complex poles, the diagonal modal form is not always possible

  ⇒ Thus we need to consider discretization of more general state-space models for $H(s)$.

Suppose

$$\dot{\underline{x}} = A_H \underline{x} + B_H \underline{e}$$

$$\underline{u} = C_H \underline{x} + D_H \underline{e}$$

Arbitrary state-space model for $H(s)$.

is any state-space model for $H(s)$. Then the corresponding discrete equivalent is

$$\underline{u}_K = C_d \underline{x}_K + D_d \underline{e}_K$$

$$\underline{x}_{K+1} = A_d \underline{x}_K + B_d \underline{e}_K$$

For ZOH discretization with sample interval T_s.

$$A_d = e^{A_H T_s} \qquad C_d = C_H$$

$$B_d = A_H^{-1}(A_d - I)B_H \qquad D_d = D_H$$

(compare with $\alpha = e^{aT_s}$, $\beta = \left(\frac{1}{a}\right)(\alpha - 1)$ in scalar case)

$A_d = e^{A_H T_s}$ $\qquad$ $B_d = A_H^{-1}[A_d - I]B_H$

recall

e^{At} is the "matrix exponential" function

Matlab function "expm"

$$e^{A_H t} = \mathcal{L}^{-1}\{(sI - A_H)^{-1}\}$$

which is an $n \times n$ matrix whose entries are scalar exponential functions determined by the eigenvalues of A_H (which are the same as the poles of $H(s)$)

$\Rightarrow$

<u>Note</u>: if A_H is singular, then A_H^{-1} DNE and B_d must be calculated using a limiting process similar to scalar case (we'll use Matlab for this).

Example I

(1)

$$H(s) = \frac{5(s+1)^2}{(s+5)^2} = 5 - \frac{40s+120}{s^2+10s+25}$$

Can take

a) $A_H = \begin{bmatrix} 0 & 1 \\ -25 & -10 \end{bmatrix}$ $\quad B_H = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $\quad C_H = [-120 \;\; -40]$

companion form

$D_H = 5$

for $T_s = 0.1$

$$A_d = \text{expm}[.1 A_H] = \begin{bmatrix} 0.91 & 0.061 \\ -1.52 & 0.30 \end{bmatrix}$$

$$B_d = A_H^{-1}[A_d - I]B_H = \begin{bmatrix} .0036 \\ .061 \end{bmatrix}$$

$$C_D = [-120 \;\; -40] \qquad D_D = 5$$

Example I, cont

(1)

$$H(s) = \frac{5(s+1)^2}{(s+5)^2} = C_H(sI - A_H)^{-1}B_H$$

b) Suppose we use instead

$$A_H = \begin{bmatrix} -5 & 1 \\ 0 & -5 \end{bmatrix} \quad B_H = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad C_H = \begin{bmatrix} 80 & -40 \end{bmatrix}$$

$$D_H = 5$$

[This is a "block modal" form for A_H]

Then here

$$A_d = \begin{bmatrix} 0.61 & 0.061 \\ 0 & 0.61 \end{bmatrix} \quad B_d = \begin{bmatrix} .0036 \\ .079 \end{bmatrix}$$

$$C_D = \begin{bmatrix} 80 & -40 \end{bmatrix} \quad D_D = 5$$

And this will yield exactly same result as a) above.

Example II:

2.) $H(s) = \dfrac{5(s+1)^2}{s^2+10s+100} = 5 - \dfrac{40s+495}{s^2+10s+100}$

a.) Can take

$$A_H = \begin{bmatrix} 0 & 1 \\ -100 & -10 \end{bmatrix} \quad B_H = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

companion form

$$C_H = \begin{bmatrix} -495 & -40 \end{bmatrix} \quad D_H = 5$$

then

$$A_d = \begin{bmatrix} 0.66 & 0.053 \\ -5.33 & 0.13 \end{bmatrix} \quad B_d = \begin{bmatrix} .0034 \\ .0534 \end{bmatrix}$$

$C_D = C_H$, $D_D = D_H$

Example II, cont

2) $H(s) = \dfrac{5(s+1)^2}{s^2 + 10s + 100}$ ~ poles at $-5 \pm \frac{10\sqrt{3}}{2} j$

b) alternately write

$$A_H = \begin{bmatrix} -5 & \frac{10\sqrt{3}}{2} \\ -\frac{10\sqrt{3}}{2} & -5 \end{bmatrix} \quad B_H = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Complex block modal form

$$C_H = \begin{bmatrix} -34 & -40 \end{bmatrix} \quad D_H = 5$$

and then

$$A_d = \begin{bmatrix} .393 & .462 \\ -.462 & .393 \end{bmatrix} \quad B_d = \begin{bmatrix} .0295 \\ .0704 \end{bmatrix}$$

$$C_d = C_H \qquad D_D = D_H$$

and again, this would be completely equivalent to the model in a) above

Which form to choose?

- The ultimate question is which form is going to yield the best accuracy given finite precision of computer calculation
- Matlab's "c2d" function does attempt to generate matrices with this goal in mind
- We will consider uses of this function, and other techniques to improve the accuracy of the discretization, in the next lecture.

```
function u = control(yd,y)
% Stub to illustrate most general form
% of discretized implementation equations

% Note: more efficient to define numerical
% components once, when we initialize x
  persistent x Ad Bd Cd Dd
  if isempty(x)
    Ad = []; % square matrix
    Bd = []; % column vector
    Cd = []; % row vector
    Dd = []; % scalar
    x = zeros(size(Ad,1),1);
    % Note: one state (xi variable) for each row of A
    %  (equivalently, for each pole in H(s))
  end

% Do the actual calculations
  e = yd - y;
  u = Cd*x + Dd*e;
  x = Ad*x + Bd*e;

end
```

Fill with #s for your H(s)

Standard template

Implementation of pole at origin (ZOH)

If $p_c = \emptyset$ (comp pole at origin), then clearly

$$\alpha = \exp[\emptyset T_s] = 1$$

in the implementation eq'n. However $\beta = \frac{(1-1)}{\emptyset}$ is indeterminate.

If we look more carefully at $\lim_{p_c \to \emptyset} \left[\frac{1 - \exp[p_c T_s]}{-p_c} \right]$

This yields the correct value $\beta = T_s$ for this case.

Thus for $\dot{x}(t) = e(t)$

we have $x(t_{k+1}) = x(t_k) + T_s\, e(t_k)$

i.e. $x_{k+1} = x_k + T_s e_k$

A closer look

$$\dot{x}(t) = e(t) \Rightarrow x_{k+1} = x_k + T_s e_k$$

equiv $\Rightarrow$ $$x(t+dt) = x(t) + dt\, e(t)$$

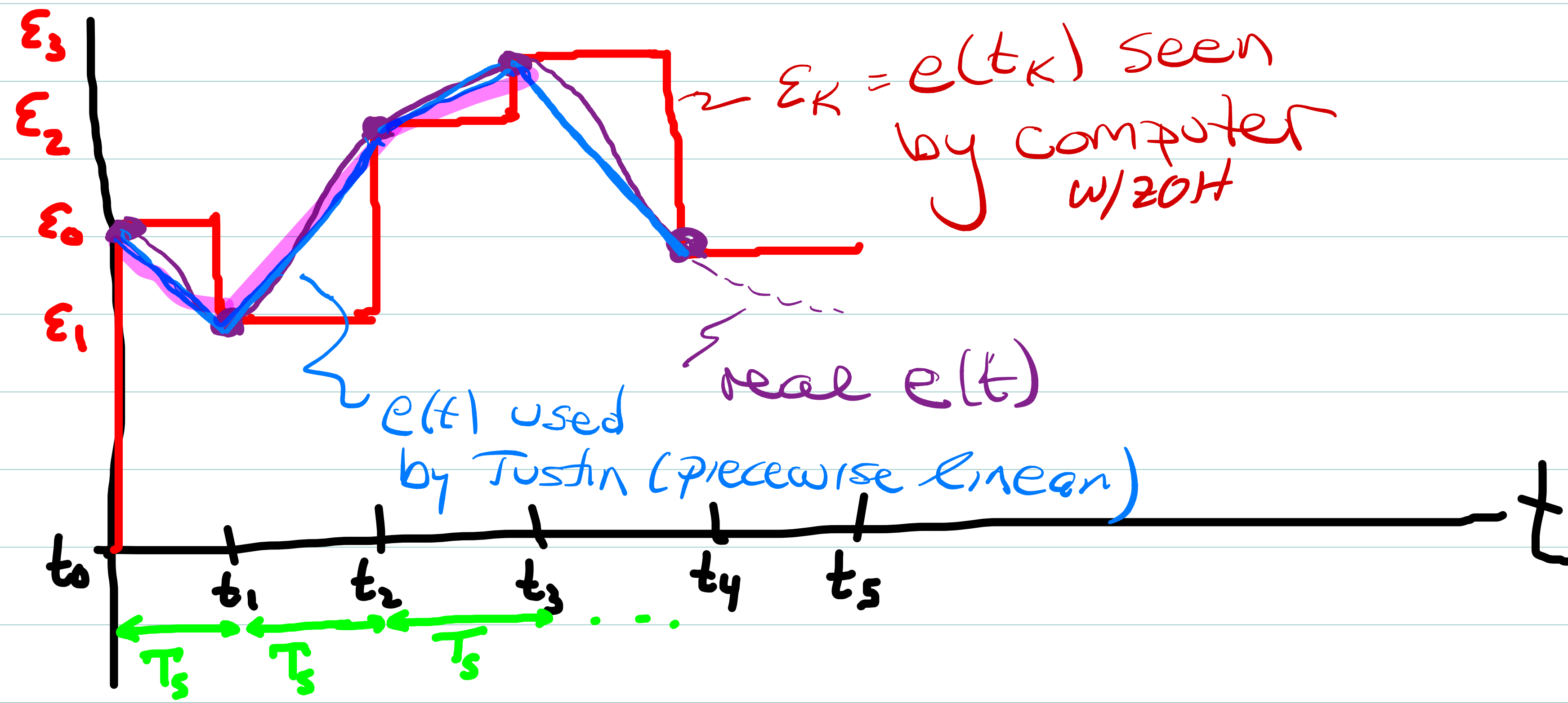
So our ZOH discretization strategy is equivalent to a simple (and not terribly accurate) Euler method for numerically integrating

Better idea:

$$x(t+dt) = x(t) + \frac{dt}{2}\left[e(t) + e(t+dt)\right]$$

i.e. a trapezoidal numerical approximation

<u>Sampling</u> of output at discrete times $t_K = KT_s$ means that error $e(t)$ will have a <u>staircase</u> graph



i.e. $e(t)$ will be constant with level $\mathcal{E}_K$ on the interval $t_K \leq t < t_{K+1}$.

Note that at t_0, $e(t)$ <u>does</u> look like a step.

Equivalent discrete equations (trap. integrate)

$$x(t+dt) = x(t) + \frac{dt}{2}\left[e(t) + e(t+dt)\right]$$

$$\Rightarrow \qquad x_{k+1} = x_k + \frac{T_s}{2}\left[e_k + e_{k+1}\right]$$

Which seems to require knowledge of future (e_{k+1})

But:

Let $\qquad z_k = x_k - \frac{T_s}{2} e_k$

Then $\qquad z_{k+1} = x_{k+1} - \frac{T_s}{2} e_{k+1}$

$$= x_k + \frac{T_s}{2} e_k + \cancel{\frac{T_s}{2} e_{k+1}} - \cancel{\frac{T_s}{2} e_{k+1}}$$

$$\Rightarrow z_{k+1} = z_k + T_s e_k$$

Trapezoidal ("Tustin") Discretization

So $\dot{x}(t) = e(t)$ can more accurately be discretized with the pair of equations

$$\boxed{\begin{aligned} z_{k+1} &= z_k + T_s e_k \\ x_k &= z_k + \frac{T_s}{2} e_k \end{aligned}}$$

Extension to general 1st order DEs is known as

"Tustin's method"

Can be ~~Generally~~ more accurate than simple ZOH.

⇒ most commonly used in practice

Straightforward to calculate, but algebraically tedious

Matlab's "c2d" function is very helpful to get the $[A_d, B_d, C_d, D_d]$ for either ZOH (default) or Tustin discretization

$$[A_d, B_d, C_d, D_d] = \text{ssdata}(\text{c2d}(H, T_s, \text{option}))$$

Omit "option" for ZOH, or use 'tustin' to specify that method.

Example: $$H(s) = \frac{.25(s+3)^3}{s(s/15+1)^2}$$

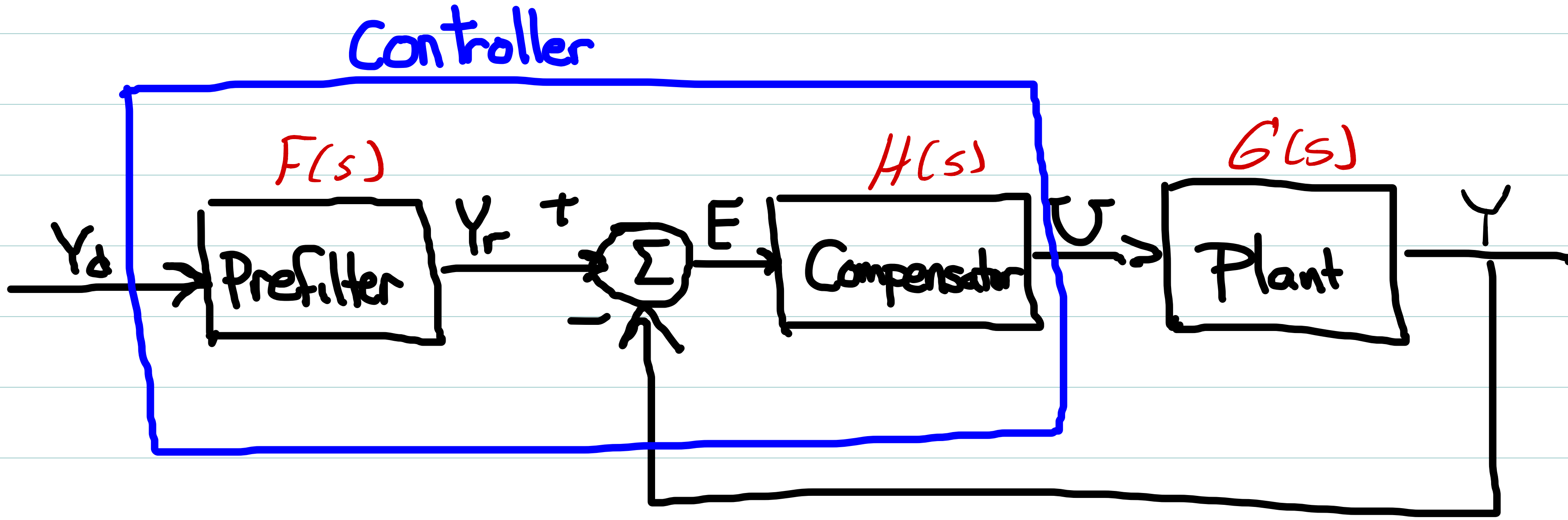
$$[A_b, B_b, C_d, D_d] = \text{ssdata}(\text{c2d}(H, .05, \text{'tustin'}));$$

<see m-file>

```
function u = control(yd,y)
% Specific illustration of implementation
% equation using the results of discretizing
% the example H(s) in makesscomp.m

  persistent x Ad Bd Cd Dd
  if isempty(x)
    Ad = [1.9091   -1.1157    0.4132;
          1.0000         0         0;
          0          0.5000        0];
    Bd = [8; 0; 0];
    Cd = [-3.1061    5.1075   -3.9777];
    Dd = 36.9609;
    x = zeros(size(Ad,1),1);
  end
  % Note: the default display of numerical results in Matlab↙
is
  % 4 decimal digits -- less than provided by a C/C++ "float"↙
type.
  % This may not provide sufficient accuracy in practice.
  % Recall that Matlab actually does all of its calculations
  % in double precision (15 decimal digits), and you can see↙
all
  % of them (to copy into control.m) using "format long".

  e = yd - y;
  u = Cd*x + Dd*e;
  x = Ad*x + Bd*e;

end
```

For Tustin disc. of $H(s) = \dfrac{.25(s+3)^3}{s(s/15+1)^2}$

"Prefilter" Designs

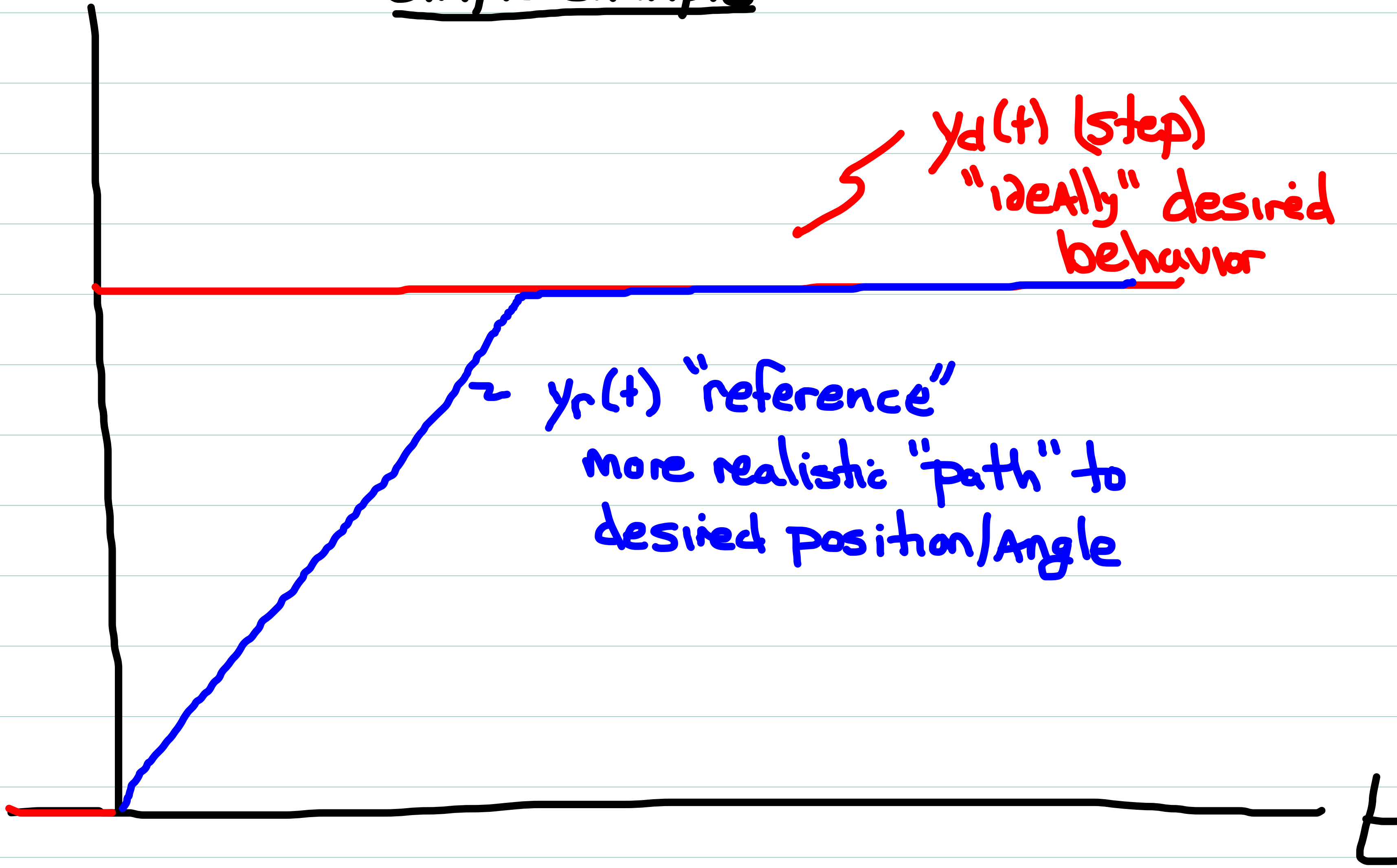


⇒ Prefilter is an extra degree of freedom in controller design

⇒ "Smooths" or "shapes" $y_d(t)$ into a "more reasonable" trajectory $y_r(t)$ which is easier for feedback loop to track

⇒ Can minimize some undesireable features of transient response, especially overshoot.

Simple Example



Reference trajectory goes to same value as $\theta_d(t)$, but in a smoother, less sudden, fashion

A useful framework for studying prefilter is to assume its action can be modeled by a transfer function $F(s)$:

$$Y_r(s) = F(s)\,Y_d(s)$$

For example, if $F(s) = \dfrac{1}{\tau s+1}$, $\tau > 0$ then

When using a prefilter we have:

$$Y(s) = T(s)Y_r(s) = \boxed{T(s)F(s)Y_d(s)}$$

where $T(s) = \dfrac{G(s)H(s)}{1+G(s)H(s)}$ as usual.

Recall that $H(s)$ typically has LHP zeros

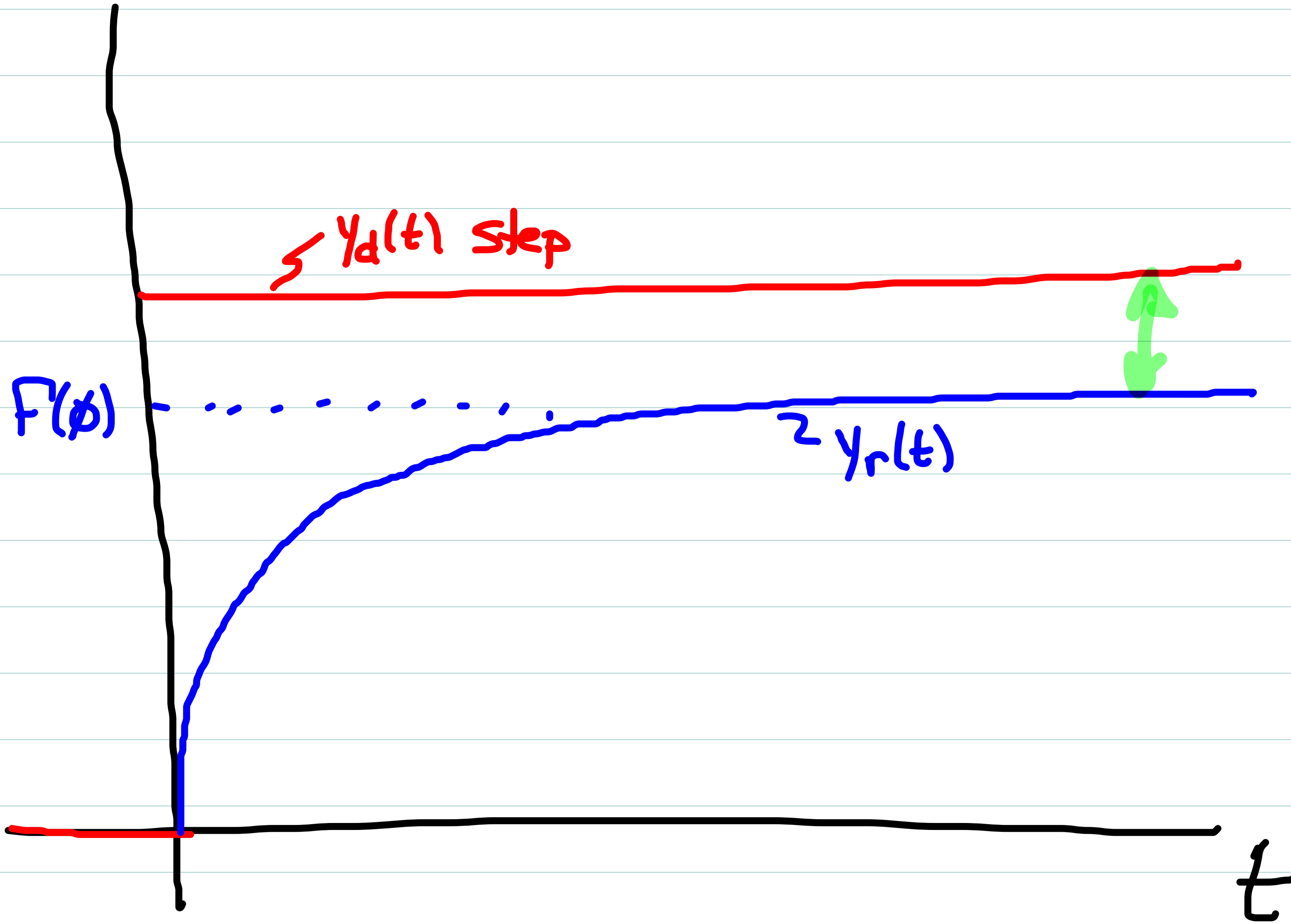
$\Rightarrow$ These zeros are also zeros of $T(s)$

$\Rightarrow$ They can substantially increase the overshoot

Use new degree of freedom $F(s)$ to <u>cancel</u> some or all zeros in $T(s)$, especially zeros used in compensator

$\Rightarrow$ $F(s)$ could have <u>poles</u> where $H(s)$ has (LHP) <u>zeros</u>!

Add'l constraint: need $F(\emptyset) = 1$ (Bode gain of 1)



If $F(\emptyset) \neq 1$, $y_r(t)$ will not converge to actual desired behavior

When using a prefilter:

$$Y(s) = \left[T(s)F(s)\right] Y_d(s)$$ Use to predict transients

$$E(s) = \left[1 - T(s)F(s)\right] Y_d(s)$$ Use to predict bandwidth

$$U(s) = \left[R(s)F(s)\right] Y_d(s)$$ Use to predict control usage

Generally a prefilter designed as above will:

$\Rightarrow$ greatly improve overshoot

$\Rightarrow$ slightly worsen tracking bandwidth

$\Rightarrow$ moderately reduce peak control efforts.

Generally advantageous (but increases complexity of implementation)

However, when using a prefilter:

=> still use $L(s)$ to design for stability (Nyquist / phase margin)

=> still use $S_i(s)$ to predict disturbance rejection

=> still use $T_o(s)$ to predict robustness and noise rejection (Δ-test)

$\boxed{\text{Prefilter does Not affect "internal" properties of feedback loop.}}$

=> $F(s)$ designed after designing a good compensator $H(s)$. All the usual design rules for $H(s)$ are unaffected by use of a prefilter.

=> Prefilter just adds a way to further "clean up" response of system to sharp changes in $y_d(t)$

⇒ Diff'l eq'ns corresponding to $F(s)$ can be implemented on computer in exactly same manner as for $H(s)$.

⇒ Do a PFE on $F(s)$, and use the resulting equations to generate $y_r(t)$ from $y_d(t)$

$$Y_r(s) = F(s)\,Y_d(s)$$

$$= \left[\frac{c_1}{s-f_1} + \frac{c_2}{s-f_2} + \cdots\right] Y_d(s)$$

⇒ Generate equivalent $x_k(t)$ diff eq'n driven by $y_d(t)$, and do a ZOH discretization just like for $H(s)$ equations
or Tustin

⇒ Then replace $y_d(t)$ with $y_r(t)$ in controller implementation i.e. use $e(t) = y_r(t) - y(t)$ in calculations for $u(t)$.

⇒ If plant has nonzero IC, good idea to initialize prefilter with $y_r(\emptyset) = y(\emptyset)$ in implementation.

Code modification w/prefilter:

$y_r = C_r * x_r + D_r * y_d$ add

$e = y_r - y$; change

$u = C_d * x + D_d * e$

$x = A_d * x + B_d * e$

same

$x_r = A_r * x_r + B_r * y_d$ add

$\Rightarrow$

$[A_r, B_r, C_r, D_r]$ obtained from $F(s)$ exactly like $[A_d, B_d, C_d, D_d]$ obtained from $H(s)$, using c2d w/same sample rate

"Pole placement" problem

Our objective in control would be easier (but not as easy as you might think!) if we could freely set all the CL poles to specified desired values.

Suppose $\lambda_1, \dots, \lambda_n$ are the desired CL poles. Then it must be the case that

$$1 + L(s) = 0$$

$$= \underline{(s-\lambda_1)(s-\lambda_2)\cdots(s-\lambda_n)}$$

$$= \underline{s^n + \gamma_{n-1} s^{n-1} + \cdots + \gamma_1 s + \gamma_0} = D_T(s)$$

i.e. Char Eq'n $1 + L(s) = 0$ must expand into above polynomial where γ_k coefs are determined by desired CL poles.

Root locus is "ouput pole placement"

Root locus examines this problem when $L(s) = \frac{K\,N(s)}{D(s)}$

where K is adjustable parameter, $N(s)$ $D(s)$ fixed, so

$$1 + L(s) = 0 \Rightarrow D(s) + K N(s) = 0$$

Let $D(s) = s^n + \alpha_{n-1}s^{n-1} + \dots + \alpha_1 s + \alpha_0$

Assume initially $N(s) = 1 \Rightarrow L(s) = \frac{K}{D(s)}$

Then

$$1 + L(s) = 0 \Rightarrow \boxed{s^n + \alpha_{n-1}s^{n-1} + \dots + \alpha_1 s + (\alpha_0 + K)}$$

if we want $\boxed{s^n + \gamma_{n-1}s^{n-1} + \dots + \gamma_1 s + \gamma_0}$

Clearly we can only "match" desired polynomial in constant term, i.e. choose K so that $\gamma_0 = \alpha_0 + K$

Not enough DOF to affect other coefs $\Rightarrow$ can't solve PP.

If, more generally, $N(s) = s^{n-1} + \beta_{n-2}s^{n-2} + \cdots + \beta_1 s + \beta_0$

then $1 + L(s) = 0$

$$= s^n + (\alpha_{n-1} + K\beta_{n-1})s^{n-1} + \cdots + (\alpha_1 s + K\beta_1)s + (\alpha_0 + K\beta_0)$$

Can affect all coefs, but not independently.

Still only 1 DOF.

Probably cannot simultaneously satisfy

$$\alpha_0 + K\beta_0 = \gamma_0$$
$$\alpha_1 + K\beta_1 = \gamma_1$$
$$\vdots$$
$$\alpha_{n-1} + K\beta_{n-1} = \gamma_{n-1}$$

except in special cases $\Rightarrow$ Root locus shows these feasible sol'ns (possible CL poles) !

Controllers with more DOF have more flexibility in "placing" CL poles.

Suppose $H(s) = K_D s + K_P$, (PD comp) $G(s) = \frac{1}{s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s + \alpha_0}$

Then

$$1 + L(s) = 0$$

$$= s^n + \alpha_{n-1}s^{n-1} + \cdots + (\alpha_1 + K_D)s + (\alpha_0 + K_P)$$

Can now independently affect 2 coefs in char poly, i.e. match

$$\gamma_0 = \alpha_0 + K_P$$

$$\gamma_1 = \alpha_1 + K_D$$

$\Rightarrow$ Larger set of CL poles feasible.

Can we generalize this idea?

i.e. Keep introducing more DOF into H(s) to allow us to achieve an arbitrary set of CL poles

⇒ While still keeping H(s) "proper"

(i.e. $\rho(H) \geq 0$ ⇒ at least as many poles as zeros)

Consider first a simple example, similar to HW problem.

Terminology:

"Proper": no more zeros than poles

"strictly proper": at least 1 more pole than zeros

Suppose $G(s) = \dfrac{1}{(s-1)^2}$

and we want all CL poles at -1 (i.e. $D_T(s) = (s+1)^k$)

Try $H(s) = \dfrac{K(s-z)}{s-p} = \dfrac{b_1 s + b_0}{s + a_0}$ (lead comp?)

Poles of $T(s)$ satisfy

$$1 + G(s)H(s) = 0$$

$$\Rightarrow 1 + \frac{b_1 s + b_0}{(s-1)^2(s+a_0)} = 0$$

3rd order poly

or $\underbrace{(s-1)^2(s+a_0) + b_1 s + b_0} = 0$

$\Rightarrow$ want this to factor as $(s+1)^3$

$$(s^2-2s+1)(s+a_0)+b_1 s+b_0$$

$$= s^3+(a_0-2)s^2+(1-2a_0+b_1)s+a_0+b_0$$

$$\stackrel{?}{=} s^3+\gamma_2 s^2+\gamma_1 s+\gamma_0 \quad \text{(Desired CL poly)}$$

$$\Rightarrow \begin{cases} \gamma_0 = a_0+b_0 & \text{Determines } b_0 = \gamma_0 - a_0 \\ \gamma_1 = 1-2a_0+b_1 & \text{Determines } b_1 = \gamma_1 - 1 + 2a_0 \\ \gamma_2 = a_0-2 & \text{Determines } a_0 = \gamma_2 + 2 \end{cases}$$

Desired CL poles

for $\lambda_1=\lambda_2=\lambda_3=-1 \Rightarrow (s+1)^3 = s^3+\underbrace{3}_{\gamma_2}s^2+\underbrace{3}_{\gamma_1}s+\underbrace{1}_{\gamma_0}$

$\gamma_2 = 3 \Rightarrow a_0 = 5$

$\gamma_1 = 3 \Rightarrow b_1 = 12$

$\gamma_2 = 1 \Rightarrow b_0 = -4$

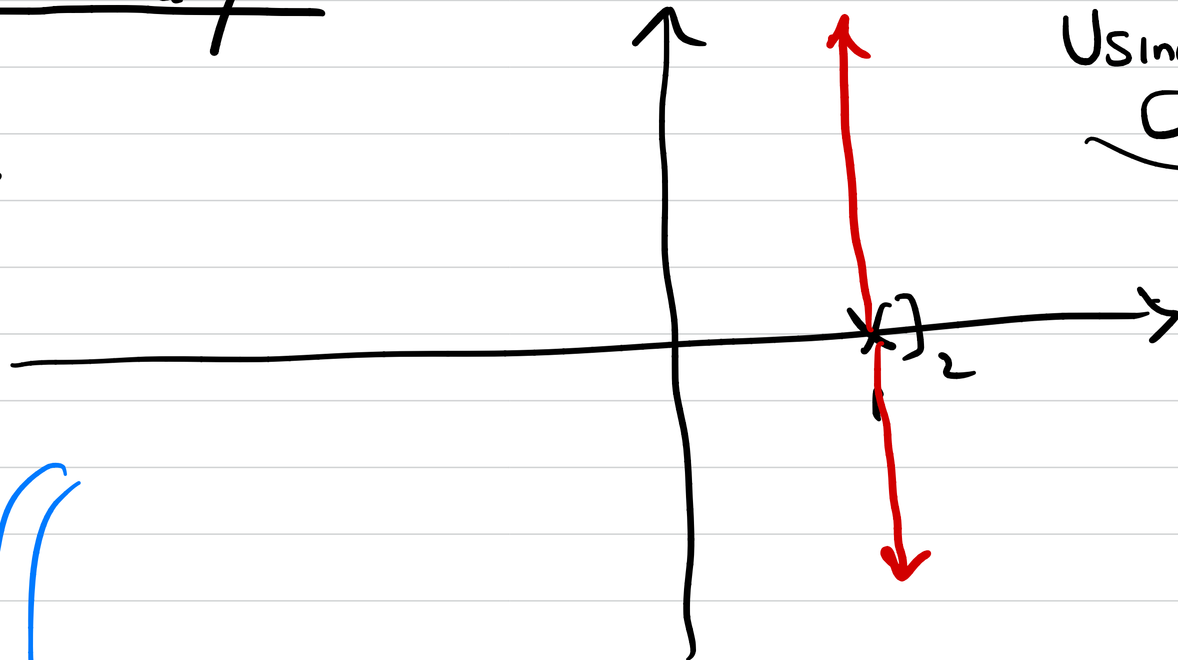
$$H(s) = \frac{12s-4}{s+5}$$

$$H(s) = \frac{K(s-1/3)}{s+5} \qquad K=12$$

Root locus analysis

$G(s) = \frac{1}{(s-1)^2}$

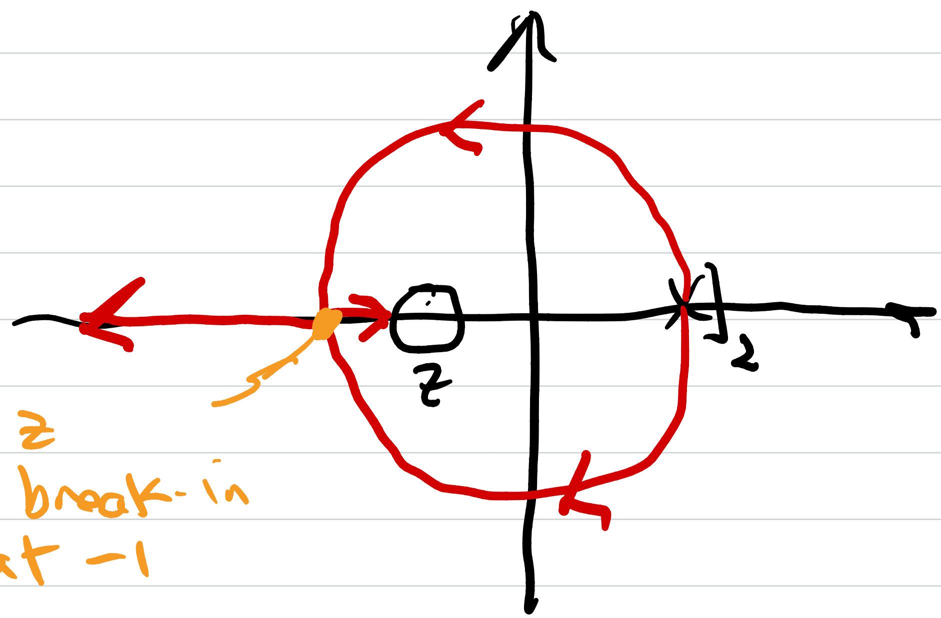
Using $H(s) = K$



Using $H(s) = K_D s + K_p$

$= K(s-z)$

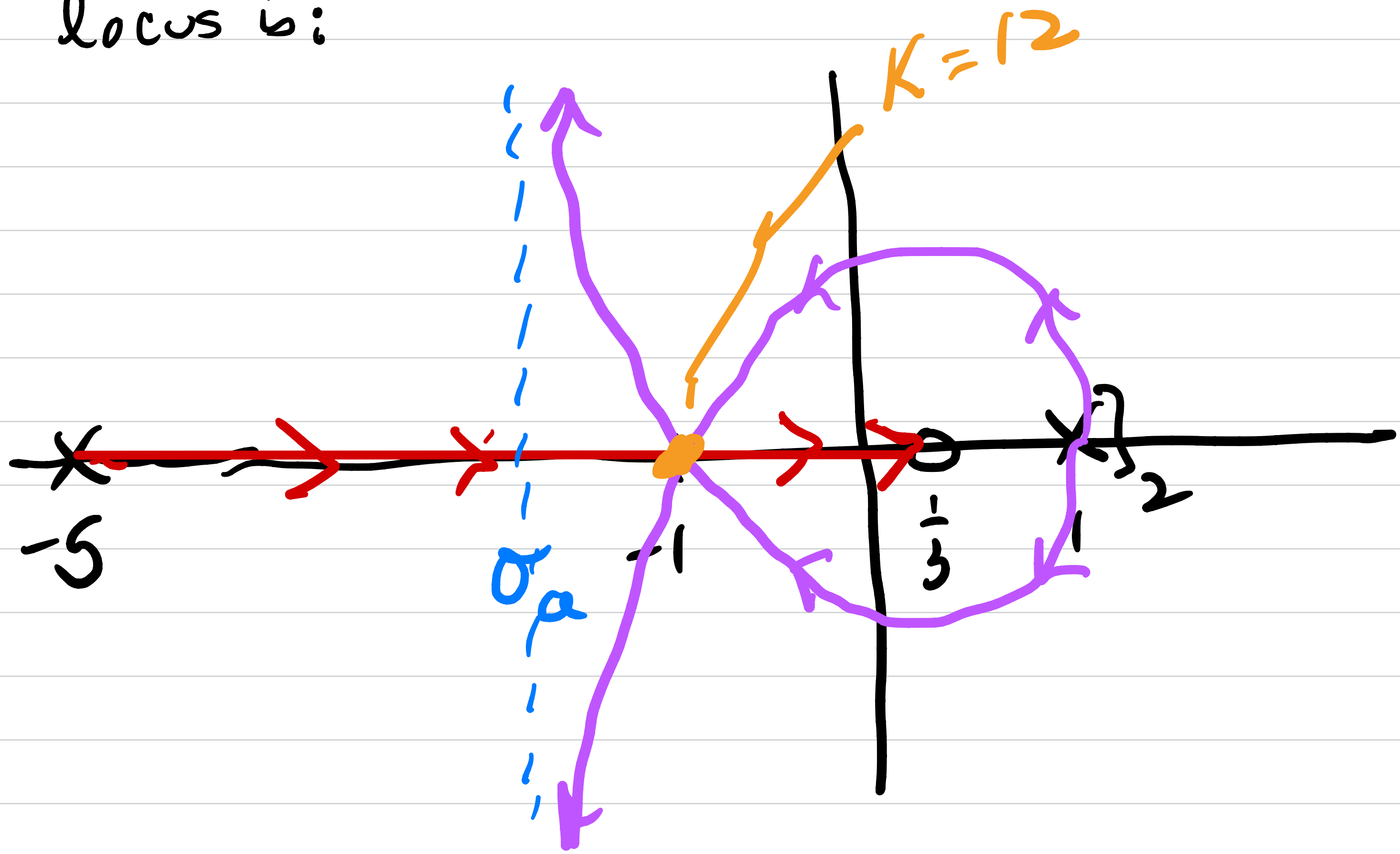
(PD – not proper/implementable)



choose z to get break-in at -1

But Using $H(s) = \dfrac{K(s-1/3)}{s+5}$ | Our Design from above

locus is:



$\gamma_e = \pm 90^\circ$

$\sigma_a = \frac{-10}{6}$

Root locus interpretation:

We have added a zero and a pole in just the Right locations to "pull" the RHP branches left in a loop that meets the New branch coming from the added pole at exactly -1

$\Rightarrow$ With enough experience in root locus we might have guessed this could happen but matching DOFs in $H(s)$ polys to desired char poly for $T(s)$ made this straightforward!

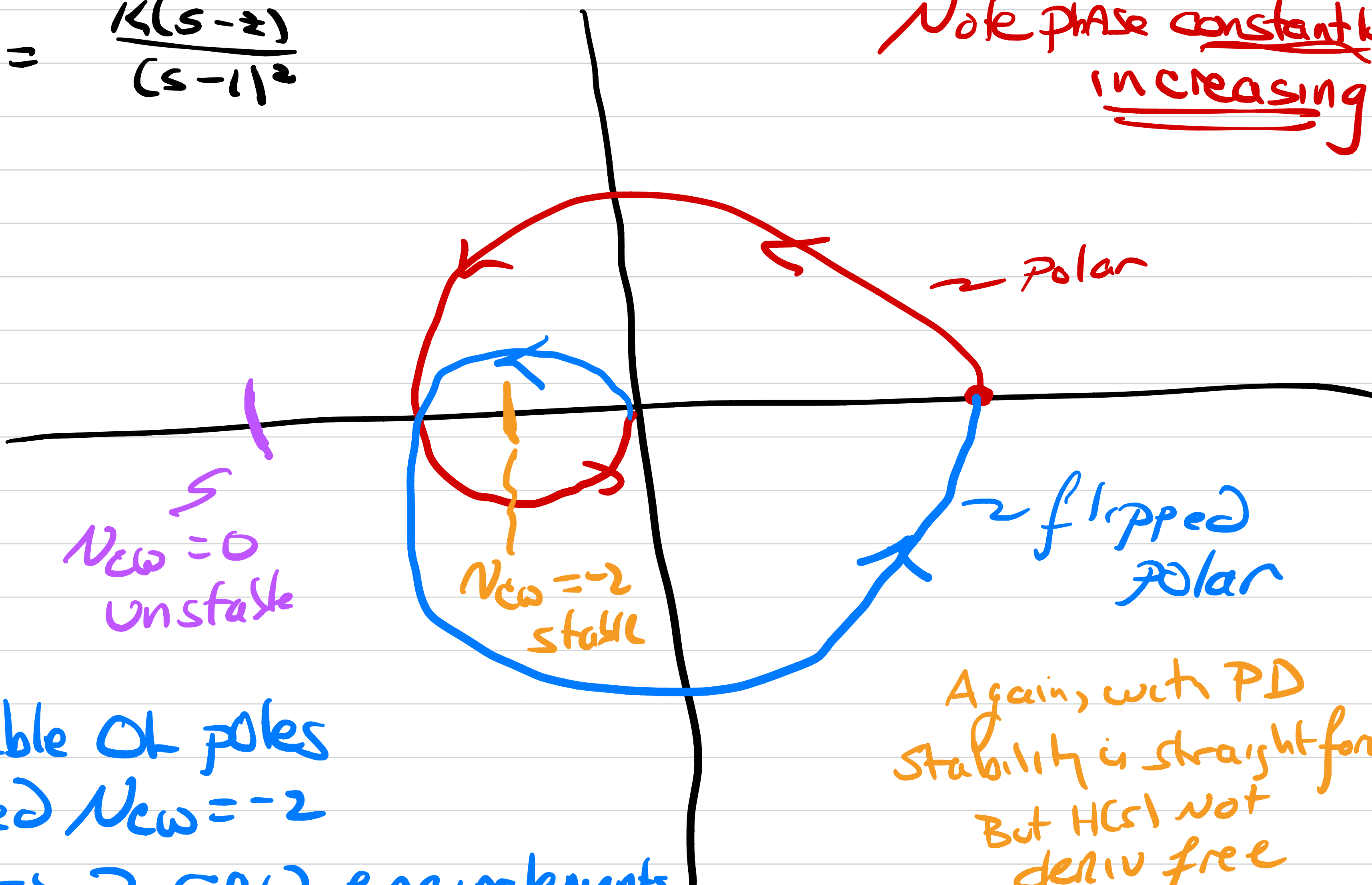
Nyquist interpretation:

$G(s) = \frac{1}{(s-1)^2}$

w/ $H(s) = K_D s + K_P = K(s-z)$ $(z<0)$ (PD comp)

$L(s) = \frac{K(s-z)}{(s-1)^2}$

Note phase constantly increasing

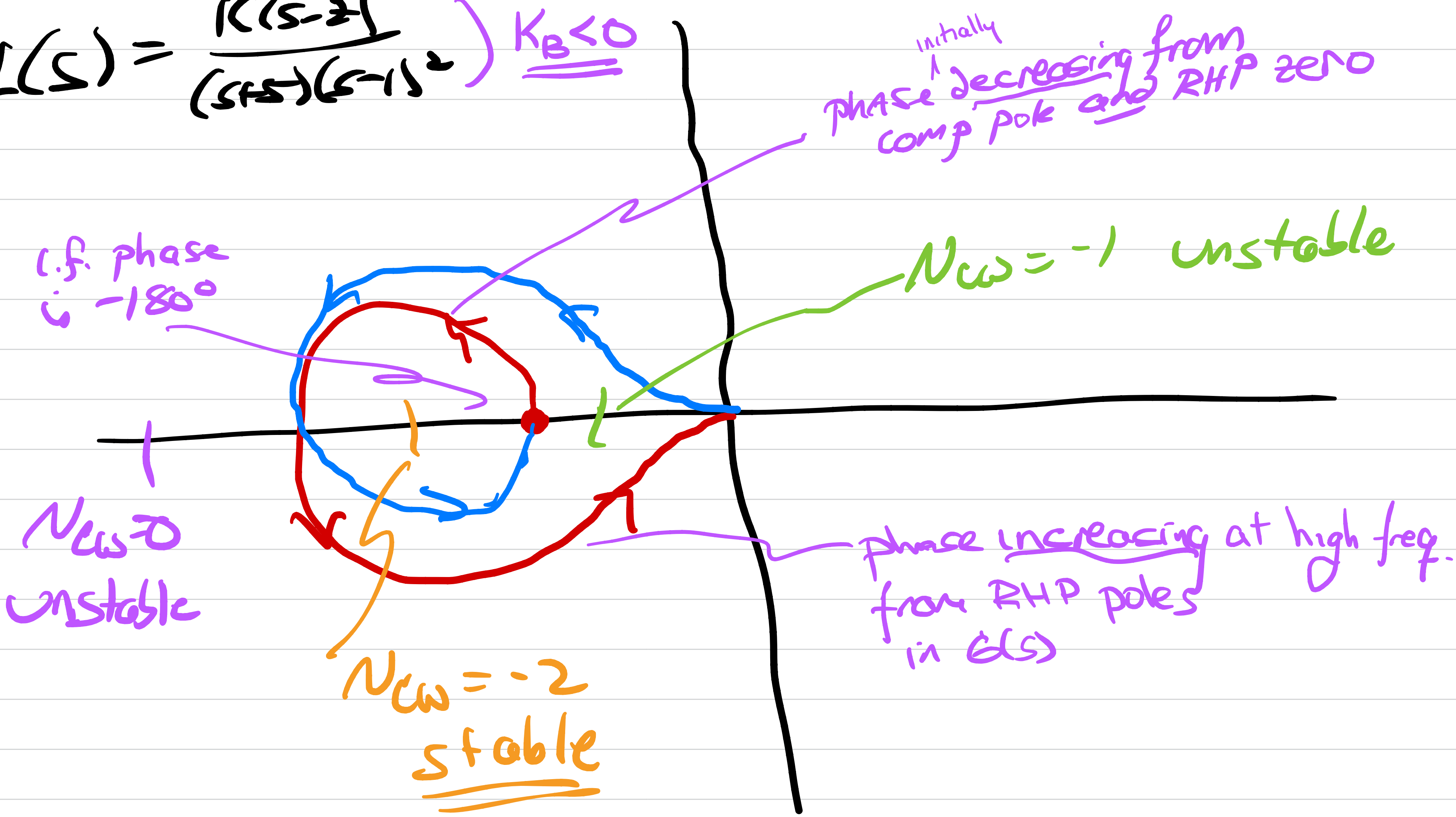


2 unstable OL poles
$\Rightarrow$ need $N_{cw} = -2$
$\Rightarrow$ 2 CCW encirclements

Again, with PD stability is straightforward
But H(s) not deriv free

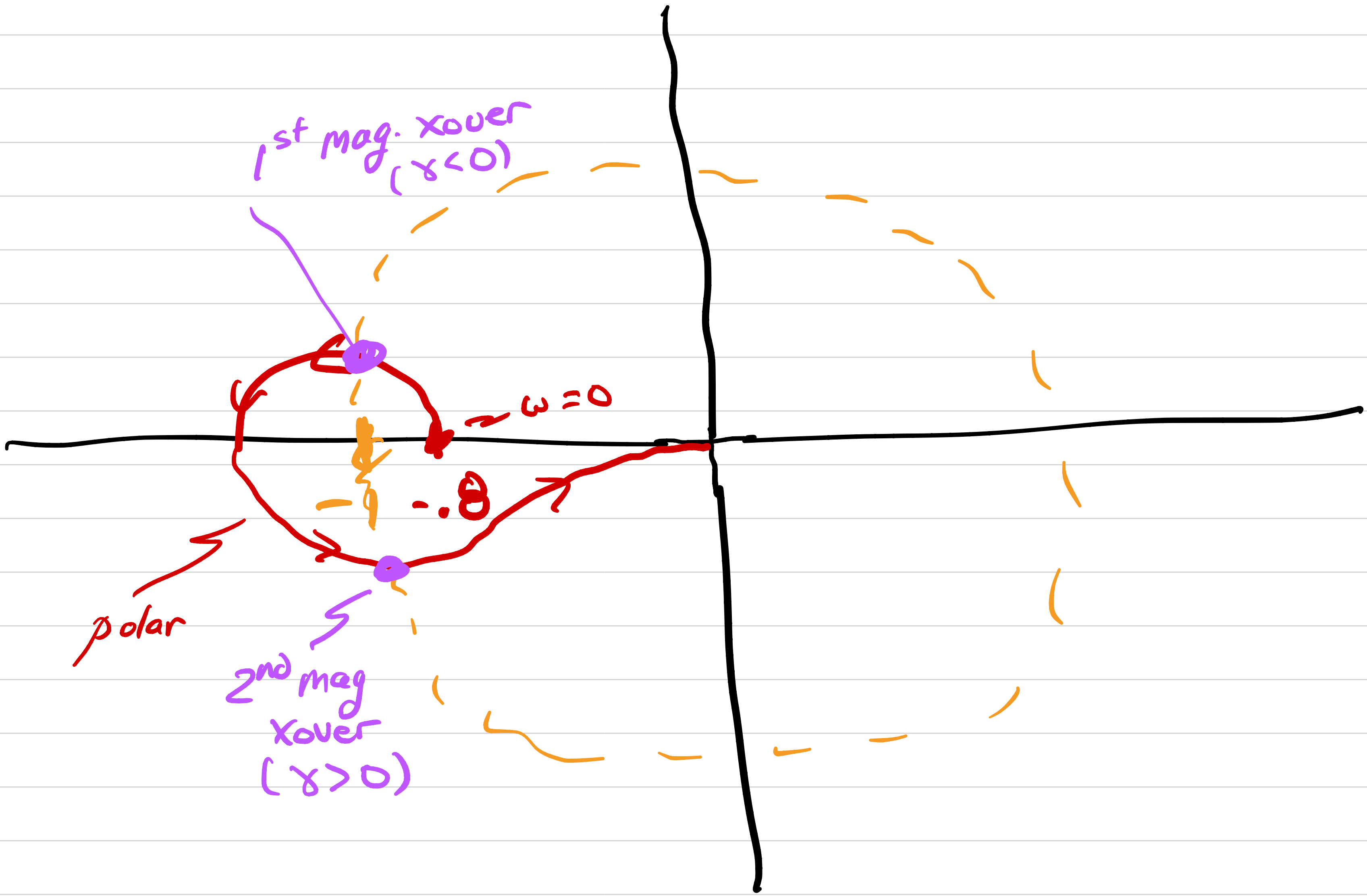
$$\text{w/ } H(s) = K\frac{(s-z)}{(s+5)} \quad (\underline{z > 0}, \text{ RHP zero!})$$

$$L(s) = \frac{K(s-z)}{(s+5)(s-1)^2}$$ $\underline{K_B < 0}$



Note: H(s) (counterintuitively) contributes negative phase at every freq here!

Specifically w/ $H(s) = \dfrac{12s-4}{5s+5}$ $\left[G(s) = \dfrac{1}{(s-1)^2} \right]$



2 mag xovers = 2 phase margins
one positive, one negative

The above "game" – matching poly coefficients in $H(s)$ to desired coefs in char poly for $T(s)$ → yields design insights that are difficult (or even counter-intuitive) from either root locus or Bode/Nyquist perspective.

⇒ A useful additional tool, if it can be systematically generalized!

Generically solve

Solve

$$\underline{D_G(s)}\,\overline{D_H(s)} + \underline{N_G(s)}\,\overline{N_H(s)} = \underline{D_T(s)}$$

specified

Generically solve

$$\underbrace{D_G(s)}_{\text{specified}}\overline{D_H}(s) + \underbrace{N_G(s)}_{\text{specified}}\overline{N_H}(s) = \underbrace{D_T(s)}_{\text{specified}}$$

Solve

Where

$$G(s) = \frac{N_G(s)}{D_G(s)} \quad \text{(fixed)}$$

$$H(s) = \frac{N_H(s)}{D_H(s)} \quad \text{free to choose}$$

$$T(s) = \frac{N_T(s)}{D_T(s)}$$

set by desired CL poles.

"Polynomial design/matching"

When can we solve this problem generally?

Suppose $G(s)$ is n^{th} order (n poles), strictly proper and either

- $H(s)$ is proper, $(n-1)$th order

or

- $H(s)$ is strictly proper, n^{th} order

then the matching problem can be solved from

$$\boxed{\underline{M}\,\underline{c} = \underline{d}}$$

$\underline{d}$ — vector of coefs from desired char poly

$\underline{c}$ — vector of coefs of num + Den of $H(s)$ — $2n-1$ or $2n$ total

$\underline{M}$ — Square Matrix of coefs of num, den in $G(s)$ — $(2n-1)$ or $(2n)$ dimension

Example:

$$G(s) = \frac{\beta_1 s + \beta_0}{s^2 + \alpha_1 s + \alpha_0}$$

Dim $n=2$

$$H(s) = \frac{b_1 s + b_0}{(s + a_0)}$$

$H(s)$ proper

3 DOF $= 2n-1$

$$1 + h = 0 = (s^2 + \alpha_1 s + \alpha_0)(s + a_0) + (\beta_1 s + \beta_0)(b_1 s + b_0)$$

$$= s^3 + \overbrace{(a_0 + \alpha_1 + \beta_1 b_1)}^{\gamma_2} s^2 + \overbrace{[\alpha_1 a_0 + \alpha_0 + \beta_1 b_0 + \beta_0 b_1]}^{\gamma_1} s + \underbrace{[\alpha_0 a_0 + \beta_0 b_0]}_{\gamma_0}$$

(3 COEFS)

$$\gamma_0 = \alpha_0 a_0 + \beta_0 b_0$$

$$\gamma_1 = a_1 \alpha_0 + \beta_1 b_0 + \beta_0 b_1 + \alpha_0$$

$$\gamma_2 = a_0 + \alpha_1 + \beta_1 b_1$$

$$\underbrace{\begin{bmatrix} \alpha_0 & \beta_0 & 0 \\ \alpha_1 & \beta_1 & \beta_0 \\ 1 & 0 & \beta_1 \end{bmatrix}}_{M\ (2n-1)\times(2n-1)} \underbrace{\begin{bmatrix} a_0 \\ b_0 \\ b_1 \end{bmatrix}}_{\underline{c}\ 2n-1} = \underbrace{\begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 - \alpha_1 \end{bmatrix}}_{\underline{d}\ 2n-1}$$

Example #2)

$$G(s) = \frac{\beta_1 s + \beta_0}{s^2 + \alpha_1 s + \alpha_0}$$

$$H(s) = \frac{b_1 s + b_0}{s^2 + a_1 s + a_0}$$

$$1 + L = s^4 + \gamma_3 s^3 + \gamma_2 s^2 + \gamma_1 s + \gamma_0$$

(4 coefs)

$\rightarrow$ 4 DOF $= 2n$ | strictly proper

$$1 + GH = 0$$

$$= (s^2 + \alpha_1 s + \alpha_0)(s^2 + a_1 s + a_0) + \underbrace{(\beta_1 s + \beta_0)(b_1 s + b_0)}_{\beta_1 b_1 s^2 + (\beta_1 b_0 + \beta_0 b_1)s + \beta_0 b_0}$$

$$s^4 + (a_1 + \alpha_1)s^3 + (a_0 + \alpha_0 + \alpha_1 a_1)s^2 + (\alpha_1 a_0 + \alpha_0 a_1)s + \alpha_0 a_0$$

$$\gamma_0 = \alpha_0 a_0 + \beta_0 b_0$$

$$\gamma_1 = (\alpha_1 a_0 + \alpha_0 a_1 + \beta_1 b_0 + \beta_0 b_1)$$

$$\gamma_2 = \beta_1 b_1 + a_0 + \alpha_0 + \alpha_1 a_1$$

$$\gamma_3 = a_1 + \alpha_1$$

$$\underset{2n \times 2n}{\underbrace{\begin{bmatrix} \alpha_0 & \beta_0 & 0 & 0 \\ \alpha_1 & \beta_1 & \alpha_0 & \beta_0 \\ 1 & 0 & \alpha_1 & \beta_1 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{M}} \underset{2n}{\underbrace{\begin{bmatrix} a_0 \\ b_0 \\ a_1 \\ b_1 \end{bmatrix}}_{c}} = \underset{2n}{\underbrace{\begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 - \alpha_0 \\ \gamma_3 - \alpha_1 \end{bmatrix}}_{d}}$$

A different way to think about this approach

Note: Not strictly proper PD

Recall if $H(s) = K_D s + K_P$, $G(s) = \dfrac{1}{s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s + \alpha_0}$

(PD comp)

Then

$$1 + L(s) = 0$$

$$= s^n + \alpha_{n-1}s^{n-1} + \cdots + (\alpha_1 + K_D)s + (\alpha_0 + K_P)$$

Can now independently affect 2 coefs in char poly,

i.e. match

$$\gamma_0 = \alpha_0 + K_P$$

$$\gamma_1 = \alpha_1 + K_D$$

$\Rightarrow$ Larger set of CL poles feasible.

An (extreme + impractical) way to add sufficient DOFs is

$$H(s) = K_{n-1} s^{n-1} + \cdots + K_1 s + K_0 \quad \Big\} \; p = -(n-1)$$

all zeros, no poles!

$$\Rightarrow \quad 1 + L(s) = 0$$

$$= s^n + (\alpha_{n-1} + K_{n-1}) s^{n-1} + \cdots + (\alpha_1 + K_1) s + (\alpha_0 + K_0)$$

We can <u>independently</u> change <u>all</u> coefs in char poly $\Rightarrow$ can "place" <u>any</u> desired set of CL poles.

"impractical" because above implies:

$$u(t) = K_0 e(t) + \underbrace{K_1 \dot{e}(t) + K_2 \ddot{e}(t) + \cdots + K_{n-1} e^{(n-1)}(t)}_{\text{measured???}}$$

Measurement of all these derivs is probably impractical (if not impossible)

But, equivalent measurements may be feasible.

Example: consider model of A/C pitch angle (θ) dynamics

Typical TF has 4 poles $\Rightarrow$ would need to measure

$\theta, \dot{\theta}, \ddot{\theta}, \dddot{\theta}$

probably not measurable.

However,

A state-space model which yields the same TF has states θ, α (angle of attack), $\dot{\theta}$, and v (air speed)

These 4 physical variables typically can be measured.

So, if we shift focus from controllers using output + its derivs to controllers using state meas, we may be able to increase effective DOFs in design + solve pole placement