Topic 11 Notes

Jeremy Orloff

11 Argument Principle

11.1 Introduction

The argument principle (or principle of the argument) is a consequence of the residue theorem. It connects the winding number of a curve with the number of zeros and poles inside the curve. This is useful for applications (mathematical and otherwise) where we want to know the location of zeros and poles.

11.2 Principle of the argument

Setup.
γ a simple closed curve, oriented in a counterclockwise direction.
$f(z)$ analytic on and inside γ, except for (possibly) some finite poles inside (not on) γ and some zeros inside (not on) γ.

Let $p_1, \ldots, p_m$ be the poles of f inside γ.
Let $z_1, \ldots, z_n$ be the zeros of f inside γ.
Write $\text{mult}(z_k)$ = the multiplicity of the zero at z_k. Likewise write $\text{mult}(p_k)$ = the order of the pole at p_k.

We start with a theorem that will lead to the argument principle.

Theorem 11.1. With the above setup

$$\int_\gamma \frac{f'(z)}{f(z)}\, dz = 2\pi i \left(\sum \text{mult}(z_k) - \sum \text{mult}(p_k) \right).$$

Proof. To prove this theorem we need to understand the poles and residues of $f'(z)/f(z)$. With this in mind, suppose $f(z)$ has a zero of order m at z_0. The Taylor series for $f(z)$ near z_0 is

$$f(z) = (z - z_0)^m g(z)$$

where $g(z)$ is analytic and never 0 on a small neighborhood of z_0. This implies

$$\begin{aligned} \frac{f'(z)}{f(z)} &= \frac{m(z-z_0)^{m-1}g(z) + (z-z_0)^m g'(z)}{(z-z_0)^m g(z)} \\ &= \frac{m}{z-z_0} + \frac{g'(z)}{g(z)} \end{aligned}$$

Since $g(z)$ is never 0, $g'(z)/g(z)$ is analytic near z_0. This implies that z_0 is a simple pole of $f'(z)/f(z)$ and

$$\text{Res}\left(\frac{f'(z)}{f(z)}, z_0 \right) = m = \text{mult}(z_0).$$

Likewise, if z_0 is a pole of order m then the Laurent series for $f(z)$ near z_0 is

$$f(z) = (z - z_0)^{-m} g(z)$$

where $g(z)$ is analytic and never 0 on a small neighborhood of z_0. Thus,

$$\begin{aligned}\frac{f'(z)}{f(z)} &= -\frac{m(z-z_0)^{-m-1}g(z) + (z-z_0)^{-m}g'(z)}{(z-z_0)^{-m}g(z)} \\ &= -\frac{m}{z-z_0} + \frac{g'(z)}{g(z)}\end{aligned}$$

Again we have that z_0 is a simple pole of $f'(z)/f(z)$ and

$$\text{Res}\left(\frac{f'(z)}{f(z)}, z_0\right) = -m = -\text{mult}(z_0).$$

The theorem now follows immediately from the Residue Theorem:

$$\begin{aligned}\int_\gamma \frac{f'(z)}{f(z)}\,dz &= 2\pi i \text{sum of the residues} \\ &= 2\pi i\left(\sum \text{mult}(z_k) - \sum \text{mult}(p_k)\right).\end{aligned}$$

Definition. We write $Z_{f,\gamma}$ for the sum of multiplicities of the zeros of f inside γ. Likewise for $P_{f,\gamma}$. So the Theorem 11.1 says,

$$\int_\gamma \frac{f'}{f}\,dz = 2\pi i\left(Z_{f,\gamma} - P_{f,\gamma}\right). \tag{1}$$

Definition. Winding number. We have an intuition for what this means. We define it formally via Cauchy's formula. If γ is a closed curve then its winding number (or index) about z_0 is defined as

$$\text{Ind}(\gamma, z_0) = \frac{1}{2\pi i}\int_\gamma \frac{1}{z - z_0}\,dz.$$

(In class I'll draw some pictures. You should draw a few now.)

11.2.1 Mapping curves: $f \circ \gamma$

One of the key notions in this topic is mapping one curve to another. That is, if $z = \gamma(t)$ is a curve and $w = f(z)$ is a function, then $w = f \circ \gamma(t) = f(\gamma(t))$ is another curve. We say f maps γ to $f \circ \gamma$. We have done this frequently in the past, but it is important enough to us now, so that we will stop here and give a few examples. This is a key concept in the argument principle and you should make sure you are very comfortable with it.

Example 11.2. Let $\gamma(t) = e^{it}$ with $0 \le t \le 2\pi$ (the unit circle). Let $f(z) = z^2$. Describe the curve $f \circ \gamma$.

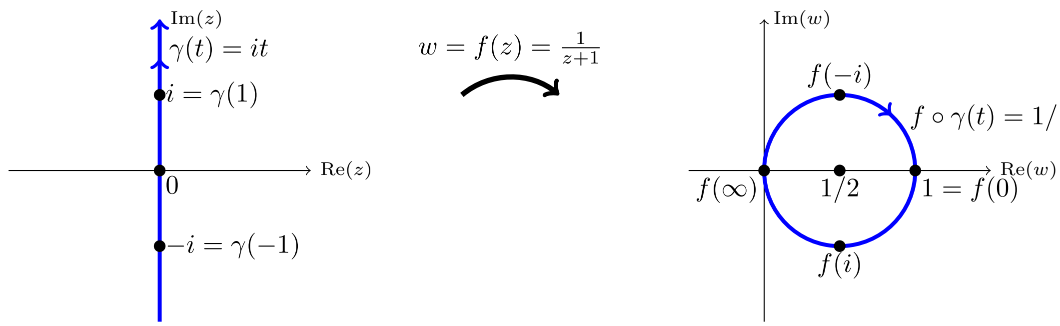
Solution: Clearly $f \circ \gamma(t) = e^{2it}$ traverses the unit circle twice as t goes from 0 to 2π.

Example 11.3. Let $\gamma(t) = it$ with $-\infty < t < \infty$ (the y-axis). Let $f(z) = 1/(z+1)$. Describe the curve $f \circ \gamma(t)$.

Solution: $f(z)$ is a fractional linear transformation and maps the line given by γ to the circle through the origin centered at $1/2$. By checking at a few points:

$$f(-i) = \frac{1}{-i+1} = \frac{1+i}{2}, \quad f(0) = 1, \quad f(i) = \frac{1}{i+1} = \frac{1-i}{2}, \quad f(\infty) = 0.$$

We see that the circle is traversed in a clockwise manner as t goes from $-\infty$ to ∞.



The curve $z = \gamma(t) = it$ is mapped to $w = f\circ\gamma(t)) = 1/(it + 1)$.

11.2.2 Argument principle

You will also see this called the principle of the argument.

Theorem 11.4. Argument principle. For f and γ with the same setup as above

$$\int_\gamma \frac{f'(z)}{f(z)}\,dz = 2\pi i\,\mathrm{Ind}(f\circ\gamma, 0) = 2\pi i(Z_{f,\gamma} - P_{f,\gamma}) \tag{2}$$

Proof. Theorem 11.1 showed that

$$\int_\gamma \frac{f'(z)}{f(z)}\,dz = 2\pi i(Z_{f,\gamma} - P_{f,\gamma})$$

So we need to show is that the integral also equals the winding number given. This is simply the change of variables $w = f(z)$. With this change of variables the countour $z = \gamma(t)$ becomes $w = f\circ\gamma(t)$ and $dw = f'(z)\,dz$ so

$$\int_\gamma \frac{f'(z)}{f(z)}\,dz = \int_{f\circ\gamma} \frac{dw}{w} = 2\pi i\,\mathrm{Ind}(f\circ\gamma, 0)$$

The last equality in the above equation comes from the definition of winding number.

Note that by assumption γ does not go through any zeros of f, so $w = f(\gamma(t))$ is never zero and $1/w$ in the integral is not a problem.

Here is an easy corollary to the argument principle that will be useful to us later.

Corollary. Assuming that $f\circ\gamma$ does not go through -1, i.e. there are no zeros of $1 + f(z)$ on γ then

$$\int_\gamma \frac{f'}{f+1} = 2\pi i\,\mathrm{Ind}(f\circ\gamma, -1) = 2\pi i(Z_{1+f,\gamma} - P_{f,\gamma}). \tag{3}$$

Proof. Applying the argument principle in Equation 2 to the function $1 + f(z)$, we get

$$\int_\gamma \frac{(1+f)'(z)}{1+f(z)}\,dz = 2\pi i\,\mathrm{Ind}(1+f\circ\gamma,0) = 2\pi i(Z_{1+f,\gamma} - P_{1+f,\gamma})$$

Now, we can compare each of the terms in this equation to those in Equation 3:

$$\begin{aligned}
\int_\gamma \frac{(1+f)'(z)}{1+f(z)}\,dz &= \int_\gamma \frac{f'(z)}{1+f(z)}\,dz && \text{(because } (1+f)' = f'\text{)}\\
\mathrm{Ind}(1+f\circ\gamma,0) &= \mathrm{Ind}(f\circ\gamma,-1) && (1+f \text{ winds around } 0 \Leftrightarrow f \text{ winds around -1})\\
Z_{1+f,\gamma} &= Z_{1+f,\gamma} && \text{(same in both equations)}\\
P_{1+f,\gamma} &= P_{f,\gamma} && \text{(poles of } f = \text{poles of } 1+f\text{)}
\end{aligned}$$

Example 11.5. Let $f(z) = z^2+z$ Find the winding number of $f\circ\gamma$ around 0 for each of the following curves.

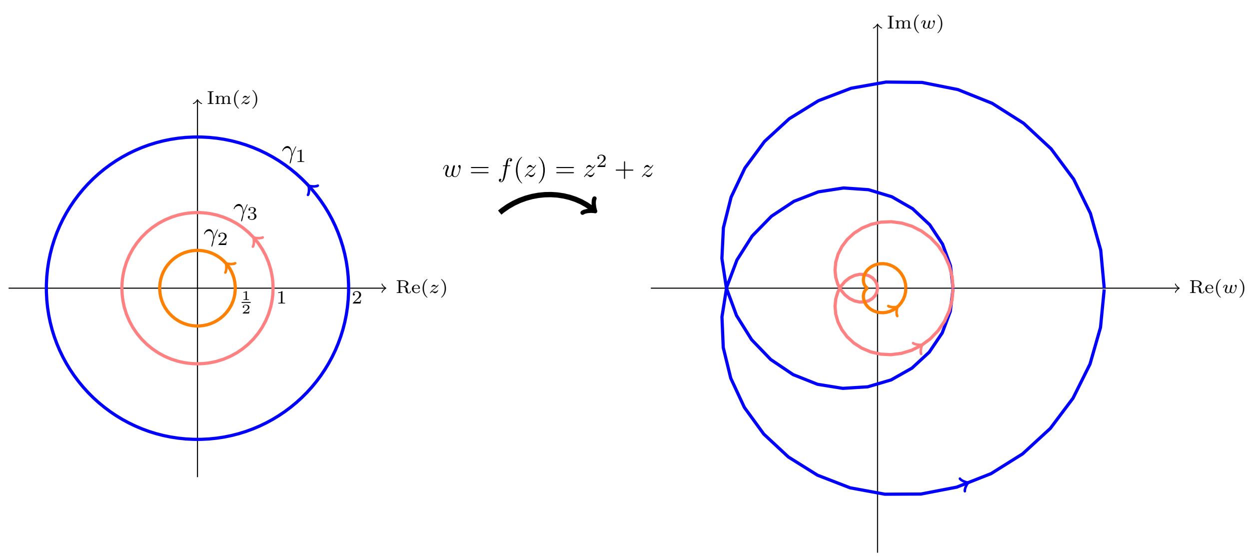
1. γ_1 = circle of radius 2.
2. γ_2 = circle of radius 1/2.
3. γ_3 = circle of radius 1.

answers. $f(z)$ has zeros at 0, -1. It has no poles.

So, f has no poles and two zeros inside γ_1. The argument principle says $\mathrm{Ind}(f\circ\gamma_1, 0) = Z_{f,\gamma_1} - P_{f,\gamma} = 2$

Likewise f has no poles and one zero inside γ_2, so $\mathrm{Ind}(f\circ\gamma_2, 0) = 1 - 0 = 1$

For γ_3 a zero of f is on the curve, i.e. $f(-1) = 0$, so the argument principle doesn't apply. The image of γ_3 is shown in the figure below – it goes through 0.



The image of 3 different circles under $f(z) = z^2 + z$.

11.2.3 Rouché's theorem.

Theorem 11.6. Rouchés theorem. Make the following assumptions:
γ is a simple closed curve

f, h are analytic functions on and inside γ, except for some finite poles.
There are no poles of f and h on γ.
$|h| < |f|$ everywhere on γ.

Then

$$\mathrm{Ind}(f\circ\gamma, 0) = \mathrm{Ind}((f+h)\circ\gamma, 0).$$

That is,

$$Z_{f,\gamma} - P_{f,\gamma} = Z_{f+h,\gamma} - P_{f+h,\gamma} \tag{4}$$

Proof. In class we gave a heuristic proof involving a person walking a dog around $f\circ\gamma$ on a leash of length $h\circ\gamma$. Here is the analytic proof.

The argument principle requires the function to have no zeros or poles on γ. So we first show that this is true of f, $f+h$, $(f+h)/f$. The argument is goes as follows.

Zeros: The fact that $0 \le |h| < |f|$ on γ implies f has no zeros on γ. It also implies $f+h$ has no zeros on γ, since the value of h is never big enough to cancel that of f. Since f and $f+h$ have no zeros, neither does $(f+h)/f$.

Poles: By assumption f and h have no poles on γ, so $f+h$ has no poles there. Since f has no zeros on γ, $(f+h)/f$ has no poles there.

Now we can apply the argument principle to f and $f+h$

$$\frac{1}{2\pi i}\int_\gamma \frac{f'}{f}\,dz = \mathrm{Ind}(f\circ\gamma, 0) = Z_{f,\gamma} - P_{f,\gamma}. \tag{5}$$

$$\frac{1}{2\pi i}\int_\gamma \frac{(f+h)'}{f+h}\,dz = \mathrm{Ind}((f+h)\circ\gamma, 0) = Z_{f+h,\gamma} - P_{f+h,\gamma}. \tag{6}$$

Next, by assumption $\left|\frac{h}{f}\right| < 1$, so $\left(\frac{h}{f}\right)\circ\gamma$ is inside the unit circle. . This means that $1+\frac{h}{f} = \frac{f+h}{f}$ maps γ to the inside of the unit disk centered at 1. (You should draw a figure for this.) This implies that

$$\mathrm{Ind}\left(\left(\frac{f+h}{f}\right)\circ\gamma, 0\right) = 0.$$

Let $g = \frac{f+h}{f}$. The above says $\mathrm{Ind}(g\circ\gamma, 0) = 0$. So, $\int_\gamma \frac{g'}{g}\,dz = 0$. (We showed above that g has no zeros or poles on γ.)

Now, it's easy to compute that $\frac{g'}{g} = \frac{(f+h)'}{f+h} - \frac{f'}{f}$. So, using

$$\mathrm{Ind}(g\circ\gamma, 0) = \int_\gamma \frac{g'}{g}\,dz = \int_\gamma \frac{(f+h)'}{f+h}\,dz - \int_\gamma \frac{f'}{f}\,dz = 0 \Rightarrow \mathrm{Ind}((f+h)\circ\gamma, 0) = \mathrm{Ind}(f\circ\gamma, 0).$$

Now equations 5 and 6 tell us $Z_{f,\gamma} - P_{f,\gamma} = Z_{f+h,\gamma} - P_{f+h,\gamma}$, i.e. we have proved Rouchés theorem.

Corollary. Under the same hypotheses, If h and f are analytic (no poles) then

$$Z_{f,\gamma} = Z_{f+h,\gamma}.$$

Proof. Since the functions are analytic $P_{f,\gamma}$ and $P_{f+h,\gamma}$ are both 0. So Equation 4 shows $Z_f = Z_{f+h}$. QED.

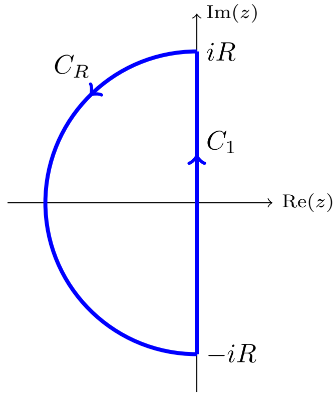
We think of h as a small perturbation of f.

Example 11.7. Show all 5 zeros of $z^5 + 3z + 1$ are inside the curve $C_2 : |z| = 2$.

Solution: Let $f(z) = z^5$ and $h(z) = 3z + 1$. Clearly all 5 roots of f (really one root with multiplicity 5) are inside C_2. Also clearly, $|h| < 7 < 32 = |f|$ on C_2. The corollary to Rouchés theorem says all 5 roots of $f + h = z^5 + 3z + 1$ must also be inside the curve.

Example 11.8. Show $z + 3 + 2e^z$ has one root in the left half-plane.

Solution: Let $f(z) = z + 3$, $h(z) = 2e^z$. Consider the contour from $-iR$ to iR along the y-axis and then the left semicircle of radius R back to $-iR$. That is, the contour $C_1 + C_R$ shown below.



To apply the corollary to Rouchés theorem we need to check that (for R large) $|h| < |f|$ on $C_1 + C_R$. On C_1, $z = iy$, so

$$|f(z)| = |3 + iy| \geq 3, \quad |h(z)| = 2|e^{iy}| = 2.$$

So $|h| < |f|$ on C_1.

On C_R, $z = x + iy$ with $x < 0$ and $|z| = R$. So,

$$|f(z)| > R - 3 \text{ for } R \text{ large}, \quad |h(z)| = 2|e^{x+iy}| = 2e^x < 2 \text{ (since } x < 0).$$

So $|h| < |f|$ on C_R.

The only zero of f is at $z = -3$, which lies inside the contour.

Therefore, by the Corollary to Rouchés theorem, $f + h$ has the same number of roots as f inside the contour, that is 1. Now let R go to infinity and we see that $f + h$ has only one root in the entire half-plane.

Theorem. Fundamental theorem of algebra.

Rouchés theorem can be used to prove the fundamental theorem of algebra as follows.

Proof. Let

$$P(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_0$$

be an nth order polynomial. Let $f(z) = z^n$ and $h = P - f$. Choose an R such that $R > \max(1, n|a_{n-1}|, \ldots, n|a_0|)$. Then on $|z| = R$ we have

$$|h| \leq |a_{n-1}|R^{n-1} + |a_{n-2}|R^{n-2} + \ldots + |a_0| \leq \frac{R}{n}R^{n-1} + \frac{R}{n}R^{n-2} + \ldots + \frac{R}{n} < R^n.$$

On $|z| = R$ we have $|f(z)| = R^n$, so we have shown $|h| < |f|$ on the curve. Thus, the corollary to Rouchés theorem says $f + h$ and f have the same number of zeros inside $|z| = R$. Since we know f has exactly n zeros inside the curve the same is true for the polynomial $f + h$. Now let R go to infinity, we've shown that $f + h$ has exactly n zeros in the entire plane.

Note. The proof gives a simple bound on the size of the zeros: they are all have magnitude less than or equal to $\max(1, n|a_{n-1}|, \ldots, n|a_0|)$.

11.3 Nyquist criterion for stability

The Nyquist criterion is a graphical technique for telling whether an unstable linear time invariant system can be stabilized using a negative feedback loop. We will look a little more closely at such systems when we study the Laplace transform in the next topic. For this topic we will content ourselves with a statement of the problem with only the tiniest bit of physical context.

Note. You have already encountered linear time invariant systems in 18.03 (or its equivalent) when you solved constant coefficient linear differential equations.

11.3.1 System functions

A linear time invariant system has a system function which is a function of a complex variable. Typically, the complex variable is denoted by s and a capital letter is used for the system function.

Let $G(s)$ be such a system function. We will make a standard assumption that $G(s)$ is meromorphic with a finite number of (finite) poles. This assumption holds in many interesting cases. For example, quite often $G(s)$ is a rational function $Q(s)/P(s)$ (Q and P are polynomials).

We will be concerned with the stability of the system.

Definition. The system with system function $G(s)$ is called stable if all the poles of G are in the left half-plane. That is, if all the poles of G have negative real part.

The system is called unstable if any poles are in the right half-plane, i.e. have positive real part.

For the edge case where no poles have positive real part, but some are pure imaginary we will call the system marginally stable. This case can be analyzed using our techniques. For our purposes it would require and an indented contour along the imaginary axis. If we have time we will do the analysis.

Example 11.9. Is the system with system function $G(s) = \dfrac{s}{(s+2)(s^2+4s+5)}$ stable?

Solution: The poles are -2, $-2 \pm i$. Since they are all in the left half-plane, the system is stable.

Example 11.10. Is the system with system function $G(s) = \dfrac{s}{(s^2-4)(s^2+4s+5)}$ stable?

Solution: The poles are ± 2, $-2 \pm i$. Since one pole is in the right half-plane, the system is unstable.

Example 11.11. Is the system with system function $G(s) = \dfrac{s}{(s+2)(s^2+4)}$ stable?

Solution: The poles are -2, $\pm 2i$. There are no poles in the right half-plane. Since there are poles on the imaginary axis, the system is marginally stable.

Terminology. So far, we have been careful to say 'the system with system function $G(s)$'. From now on we will allow ourselves to be a little more casual and say 'the system $G(s)$'. It is perfectly clear and rolls off the tongue a little easier!

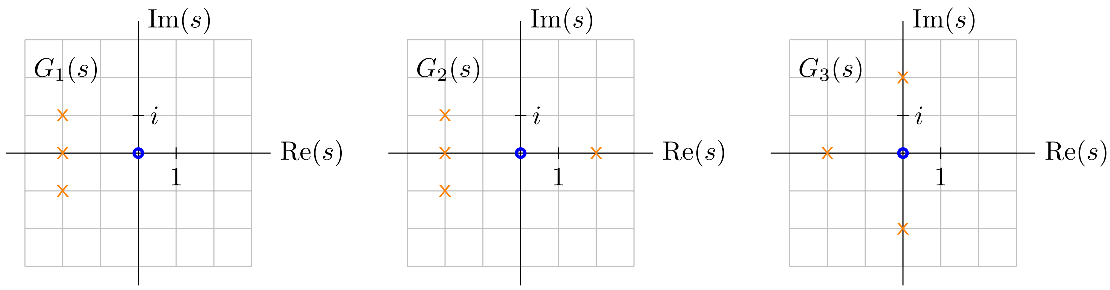
11.3.2 Pole-zero diagrams

We can visualize $G(s)$ using a pole-zero diagram. This is a diagram in the s-plane where we put a small cross at each pole and a small circle at each zero.

Example 11.12. Give zero-pole diagrams for each of the systems

$$G_1(s) = \frac{s}{(s+2)(s^2+4s+5)}, \quad G_2(s) = \frac{s}{(s^2-4)(s^2+4s+5)}, \quad G_3(s) = \frac{s}{(s+2)(s^2+4)}$$

Solution: These are the same systems as in the examples just above. We first note that they all have a single zero at the origin. So we put a circle at the origin and a cross at each pole.



Pole-zero diagrams for the three systems.

11.3.3 A bit about stability

This is just to give you a little physical orientation. Given our definition of stability above, we could, in principle, discuss stability without the slightest idea what it means for physical systems.

The poles of $G(s)$ correspond to what are called modes of the system. A simple pole at s_1 corresponds to a mode $y_1(t) = e^{s_1 t}$. The system is stable if the modes all decay to 0, i.e. if the poles are all in the left half-plane.

Physically the modes tell us the behavior of the system when the input signal is 0, but there are initial conditions. A pole with positive real part would correspond to a mode that goes to infinity as t grows. It is certainly reasonable to call a system that does this in response to a zero signal (often called 'no input') unstable.

To connect this to 18.03: if the system is modeled by a differential equation, the modes correspond to the homogeneous solutions $y(t) = e^{st}$, where s is a root of the characteristic equation. In 18.03 we called the system stable if every homogeneous solution decayed to 0. That is, if the unforced system always settled down to equilibrium.

11.3.4 Closed loop systems

If the system with system function $G(s)$ is unstable it can sometimes be stabilized by what is called a negative feedback loop. The new system is called a closed loop system. Its system function is given by Black's formula

$$G_{CL}(s) = \frac{G(s)}{1 + kG(s)}, \tag{7}$$

where k is called the feedback factor. We will just accept this formula. Any class or book on control theory will derive it for you.

In this context $G(s)$ is called the open loop system function.

Since G_{CL} is a system function, we can ask if the system is stable.

Theorem. The poles of the closed loop system function $G_{CL}(s)$ given in Equation 7 are the zeros of $1 + kG(s)$.

Proof. Looking at Equation 7, there are two possible sources of poles for G_{CL}.

1. The zeros of the denominator $1 + kG$. The theorem recognizes these.

2. The poles of G. Since G is in both the numerator and denominator of G_{CL} it should be clear that the poles cancel. We can show this formally using Laurent series. If G has a pole of order n at s_0 then

$$G(s) = \frac{1}{(s - s_0)^n}\left(b_n + b_{n-1}(s - s_0) + \ldots a_0(s - s_0)^n + a_1(s - s_0)^{n+1} + \ldots\right),$$

where $b_n \neq 0$. So,

$$\begin{aligned} G_{CL}(s) &= \frac{\frac{1}{(s-s_0)^n}\left(b_n + b_{n-1}(s - s_0) + \ldots a_0(s - s_0)^n + \ldots\right)}{1 + \frac{k}{(s-s_0)^n}\left(b_n + b_{n-1}(s - s_0) + \ldots a_0(s - s_0)^n + \ldots\right)} \\ &= \frac{\left(b_n + b_{n-1}(s - s_0) + \ldots a_0(s - s_0)^n + \ldots\right)}{(s - s_0)^n + k\left(b_n + b_{n-1}(s - s_0) + \ldots a_0(s - s_0)^n + \ldots\right)}, \end{aligned}$$

which is clearly analytic at s_0. (At s_0 it equals $b_n/(kb_n) = 1/k$.)

Example 11.13. Set the feedback factor $k = 1$. Assume a is real, for what values of a is the open loop system $G(s) = \dfrac{1}{s + a}$ stable? For what values of a is the corresponding closed loop system $G_{CL}(s)$ stable?

(There is no particular reason that a needs to be real in this example. But in physical systems, complex poles will tend to come in conjugate pairs.)

Solution: $G(s)$ has one pole at $s = -a$. Thus, it is stable when the pole is in the left half-plane, i.e. for $a > 0$.

The closed loop system function is

$$G_{CL}(s) = \frac{1/(s + a)}{1 + 1/(s + a)} = \frac{1}{s + a + 1}.$$

This has a pole at $s = -a - 1$, so it's stable if $a > -1$. The feedback loop has stabilized the unstable open loop systems with $-1 < a \leq 0$. (Actually, for $a = 0$ the open loop is marginally stable, but it is fully stabilized by the closed loop.)

Note. The algebra involved in canceling the $s + a$ term in the denominators is exactly the cancellation that makes the poles of G removable singularities in G_{CL}.

Example 11.14. Suppose $G(s) = \dfrac{s + 1}{s - 1}$. Is the open loop system stable? Is the closed loop system stable when $k = 2$.

Solution: $G(s)$ has a pole in the right half-plane, so the open loop system is not stable. The closed loop system function is

$$G_{CL}(s) = \frac{G}{1+kG} = \frac{(s+1)/(s-1)}{1+2(s+1)/(s-1)} = \frac{s+1}{3s+1}.$$

The only pole is at $s = -1/3$, so the closed loop system is stable. This is a case where feedback stabilized an unstable system.

Example 11.15. $G(s) = \dfrac{s-1}{s+1}$. Is the open loop system stable? Is the closed loop system stable when $k = 2$.

Solution: The only pole of $G(s)$ is in the left half-plane, so the open loop system is stable. The closed loop system function is

$$G_{CL}(s) = \frac{G}{1+kG} = \frac{(s-1)/(s+1)}{1+2(s-1)/(s+1)} = \frac{s-1}{3s-1}.$$

This has one pole at $s = 1/3$, so the closed loop system is unstable. This is a case where feedback destabilized a stable system. It can happen!

11.3.5 Nyquist plot

For the Nyquist plot and criterion the curve γ will always be the imaginary s-axis. That is

$$s = \gamma(\omega) = i\omega, \quad \text{where} \quad -\infty < \omega < \infty.$$

For a system $G(s)$ and a feedback factor k, the Nyquist plot is the plot of the curve

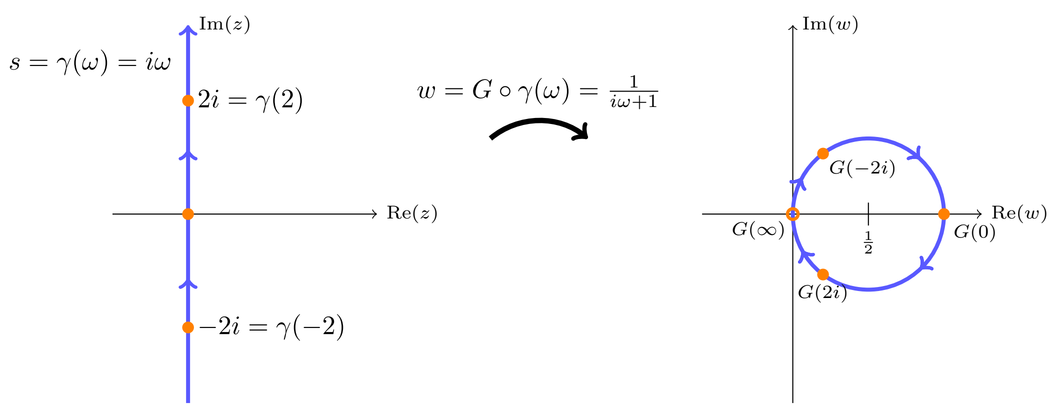
$$w = kG\circ\gamma(\omega) = kG(i\omega).$$

That is, the Nyquist plot is the image of the imaginary axis under the map $w = kG(s)$.

Note. In $\gamma(\omega)$ the variable is a greek omega and in $w = G\circ\gamma$ we have a double-u.

Example 11.16. Let $G(s) = \dfrac{1}{s+1}$. Draw the Nyquist plot with $k = 1$.

Solution: In this case $G(s)$ is a fractional linear transformation, so we know it maps the imaginary axis to a circle. It is easy to check it is the circle through the origin with center $w = 1/2$. You can also check that it is traversed clockwise.



Nyquist plot of $G(s) = 1/(s+1)$, with $k = 1$.

Example 11.17. Take $G(s)$ from the previous example. Describe the Nyquist plot with gain factor $k = 2$.

Solution: The Nyquist plot is the graph of $kG(i\omega)$. The factor $k = 2$ will scale the circle in the previous example by 2. That is, the Nyquist plot is the circle through the origin with center $w = 1$.

In general, the feedback factor will just scale the Nyquist plot.

11.3.6 Nyquist criterion

The Nyquist criterion gives a graphical method for checking the stability of the closed loop system.

Theorem 11.18. Nyquist criterion. Suppose that $G(s)$ has a finite number of zeros and poles in the right half-plane. Also suppose that $G(s)$ decays to 0 as s goes to infinity. Then the closed loop system with feedback factor k is stable if and only if the winding number of the Nyquist plot around $w = -1$ equals the number of poles of $G(s)$ in the right half-plane.

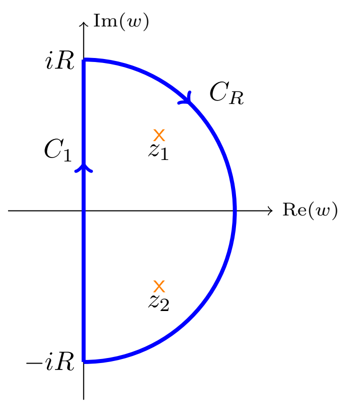
More briefly,

$$G_{CL}(s) \text{ is stable } \Leftrightarrow \text{Ind}(kG \circ \gamma, -1) = P_{G,\text{RHP}}$$

Here, γ is the imaginary s-axis and $P_{G,\text{RHP}}$ is the number of poles of **the original open loop system function** $G(s)$ in the right half-plane.

Proof. G_{CL} is stable exactly when all its poles are in the left half-plane. Now, recall that the poles of G_{CL} are exactly the zeros of $1 + kG$. So, stability of G_{CL} is exactly the condition that the number of zeros of $1 + kG$ in the right half-plane is 0.

Let's work with a familiar contour.



Let $\gamma_R = C_1 + C_R$. Note that γ_R is traversed in the *clockwise* direction. Choose R large enough that the (finite number) of poles and zeros of G in the right half-plane are all inside γ_R. Now we can apply Equation 3 in the corollary to the argument principle to $kG(s)$ and γ to get

$$-\text{Ind}(kG \circ \gamma_R, -1) = Z_{1+kG,\gamma_R} - P_{G,\gamma_R}$$

(The minus sign is because of the clockwise direction of the curve.) Thus, for all large R

$$\text{the system is stable} \;\Leftrightarrow\; Z_{1+kG,\gamma_R} = 0 \;\Leftrightarrow\; \text{Ind}(kG \circ \gamma_R, -1) = P_{G,\gamma_R}.$$

Finally, we can let R go to infinity. The assumption that $G(s)$ decays 0 to as s goes to ∞ implies that in the limit, the entire curve $kG \circ C_R$ becomes a single point at the origin. So in the limit $kG \circ \gamma_R$ becomes $kG \circ \gamma$. QED

11.3.7 Examples using the Nyquist Plot mathlet

The Nyquist criterion is a visual method which requires some way of producing the Nyquist plot. For this we will use one of the MIT Mathlets (slightly modified for our purposes).

Open the Nyquist Plot applet at
http://web.mit.edu/jorloff/www/jmoapplets/nyquist/nyquistCrit.html

Play with the applet, read the help.

Now refresh the browser to restore the applet to its original state. Check the *Formula* box. The formula is an easy way to read off the values of the poles and zeros of $G(s)$. In its original state, applet should have a zero at $s = 1$ and poles at $s = 0.33 \pm 1.75\,i$.

The left hand graph is the pole-zero diagram. The right hand graph is the Nyquist plot.

Example 11.19. To get a feel for the Nyquist plot. Look at the pole diagram and use the mouse to drag the yellow point up and down the imaginary axis. Its image under $kG(s)$ will trace out the Nyquis plot.

Notice that when the yellow dot is at either end of the axis its image on the Nyquist plot is close to 0.

Example 11.20. Refresh the page, to put the zero and poles back to their original state. There are two poles in the right half-plane, so the open loop system $G(s)$ is unstable. With $k = 1$, what is the winding number of the Nyquist plot around -1? Is the closed loop system stable?

Solution: The curve winds twice around -1 in the counterclockwise direction, so the winding number $\text{Ind}(kG \circ \gamma, -1) = 2$. Since the number of poles of G in the right half-plane is the same as this winding number, the closed loop system is stable.

Example 11.21. With the same poles and zeros, move the k slider and determine what range of k makes the closed loop system stable.

Solution: When k is small the Nyquist plot has winding number 0 around -1. For these values of k, G_{CL} is unstable. As k increases, somewhere between $k = 0.65$ and $k = 0.7$ the winding number jumps from 0 to 2 and the closed loop system becomes stable. This continues until k is between 3.10 and 3.20, at which point the winding number becomes 1 and G_{CL} becomes unstable.

Answer: The closed loop system is stable for k (roughly) between 0.7 and 3.10.

Example 11.22. In the previous problem could you determine analytically the range of k where $G_{CL}(s)$ is stable?

Solution: Yes! This is possible for small systems. It is more challenging for higher order systems, but there are methods that don't require computing the poles.

In this case, we have

$$G_{CL}(s) = \frac{G(s)}{1 + kG(s)} = \frac{\frac{s-1}{(s-0.33)^2+1.75^2}}{1 + \frac{k(s-1)}{(s-0.33)^2+1.75^2}} = \frac{s-1}{(s-0.33)^2 + 1.75^2 + k(s-1)}$$

So the poles are the roots of

$$(s - 0.33)^2 + 1.75^2 + k(s - 1) = s^2 + (k - 0.66)s + 0.33^2 + 1.75^2 - k$$

For a quadratic with positive coefficients the roots both have negative real part. This happens when

$$0.66 < k < 0.33^2 + 1.75^2 \approx 3.17.$$

Example 11.23. What happens when k goes to 0.

Solution: As k goes to 0, the Nyquist plot shrinks to a single point at the origin. In this case the winding number around -1 is 0 and the Nyquist criterion says the closed loop system is stable if and only if the open loop system is stable.

This should make sense, since with $k = 0$,

$$G_{CL} = \frac{G}{1 + kG} = G.$$

Example 11.24. Make a system with the following zeros and poles:

A pair of zeros at $0.6 \pm 0.75i$
A pair of poles at $-0.5 \pm 2.5i$.
A single pole at 0.25.

Is the corresponding closed loop system stable when $k = 6$?

Solution: The answer is no, G_{CL} is not stable. G has one pole in the right half plane. The mathlet shows the Nyquist plot winds once around $w = -1$ in the *clockwise* direction. So the winding number is -1, which does not equal the number of poles of G in the right half-plane.

If we set $k = 3$, the closed loop system is stable.

11.4 A bit on negative feedback

Given Equation 7, in 18.04 we can ask if there are any poles in the right half-plane without needing any underlying physical model. Still, it's nice to have some sense of where this fits into science and engineering.

In a negative feedback loop the output of the system is looped back and subtracted from the input.

Example 11.25. The heating system in my house is an example of a system stabilized by feedback. The thermostat is the feedback mechanism. When the temperature outside (input signal) goes down the heat turns on. Without the thermostat it would stay on and overheat my house. The thermostat turns the heat up or down depending on whether the inside temperature (the output signal) is too low or too high (negative feedback).

Example 11.26. Walking or balancing on one foot are examples negative feedback systems. If you feel yourself falling you compensate by shifting your weight or tensing your muscles to counteract the unwanted acceleration.

MIT OpenCourseWare
https://ocw.mit.edu

18.04 Complex Variables with Applications
Spring 2018

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.