Control System Design
Prof. G.R. Jayanth
Department of Instrumentation and Applied Physics
Indian Institute of Science, Bangalore

Lecture - 45
Fundamental properties of nonminimum phase systems (part 1/2)

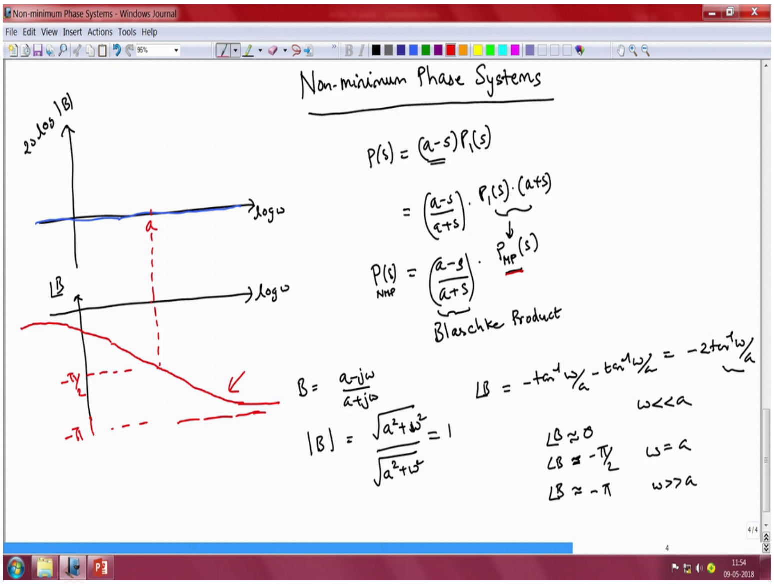
Hello, in the previous clip we initiated discussion on nonminimum phase systems and tried to look at how nonminimum phase systems could potentially provide post challenges to us as control engineers and we took some case studies, we took some from specific numerical examples through which we showed that in the presence of a time delay for instance or abilities as control engineers gets, drastically diminished and we often have to settle for closed loop systems with extremely low closed loop bandwidth, in order to make sure that the closed loop system is stable in the presence of the nonminimum phase term such as, a time delay.

Now, we shall in this clip we shall continue with this discussion. We shall first briefly look at, nonminimum phase system which have right half plane 0's and subsequently look at the Fundamental limitations on the gain crossover frequency and the phase crossover frequency that are imposed by the nonminimum phase terms. Again you will first do this for the case of the right of planes zeroes and repeat the same exercise for the case of time delay.

Subsequently, we will look at the Fundamental limitations that are imposed on the achievable loop gain at frequencies below the gain crossover frequency as a result of the fact that there exists such a thing as the maximum achievable gain crossover frequency.

So, let us first get started with Non-minimum Phase Systems which have right half planes 0's. As we discussed in the previous clip, the general transfer function for a plant P of s which has a single right half plane 0 at s is equal a is given by P of s is equal to a minus s times, P 1 of s, where P 1 of s is assumed to be a transfer function all of whose poles and 0's are on the left half of the complex plane.

(Refer Slide Time: 02:02)



So, this term is a Non-minimum Phase term and for the sake of greater clarity in understanding the effect of this term we shall rewrite this expression by multiplying the right hand side and dividing it by a term a plus s. So, that on the right hand side we would have a minus s by a plus s times P 1, times s P 1 of s times a plus s. So, I have done nothing but multiply and divide the right hand side by the same term namely a plus s and I shall call the all the product of P 1 and a plus s as the minimum phase part of the transfer function. So, I shall give it a special name as P MP of s.

So, P 1 times a plus s shall be called P minimum phase of s and that gets multiplied with a minus s by a plus s; to give me the actual transfer function of the plant P of s which once again I shall re-label as P NMP of s to highlight the fact that we are not talking of a nonminimum phase plant.

So, the the nonminimum phase plant can be written as a product of a term of this kind namely a minus s by a plus s times a minimum phase transfer function P MP of s. Now, this particular term has a special name it is called as a Blaschke Product, Blaschke Product. So, to understand the effect of this Blaschke Product on the overall magnitude and phase characteristics of the nonminimum phase plant. Let us first draw the bode plot of the Blaschke Product alone. So, to draw the bode plot of the Blaschke Product we said s is equal to j omega in which case the Blaschke Product is going to be given by a minus j omega by a plus, j omega. So, the magnitude of the Blaschke Product is going to be equal to square root of a square plus omega square divided by square root of a square, plus omega square once again and it is going to be equal to 1.

So, you note in this case a similarity in the magnitude of the Blaschke Product with that of the time delay. So, once again this particular term does not modify the magnitude characteristics of the minimum phase plant in any particular way it is gain is going to be one at all frequencies; however, if we compute the phase of the Blaschke Product we note that in the numerator we have a minus j omega.

So, the phase associated with that would be minus of tan inverse of omega by a and then we have a denominator of the kind a plus j omega and a phase of a plus j omega is going to be equal to tan inverse of omega by a and since it appears in the denominator it is going to be minus of this particular value or in other words of minus of tan inverse omega by a and hence the net phase of b a function of frequency is given by minus 2 times, tan inverse, omega by a.

So, let us now having derived the magnitude and the phase characteristics of the Blaschke Product; let us now, graph it is bode plot. So, the magnitude plot where the x axis is log omega and the y axis is 20, log, of magnitude of b, will be essentially a horizontal straight line that is coincident with the x axis of the plot because the gain of the Blaschke Product at all frequencies is going to be equal to 0 db or in linear terms it is going to be equal to 1.

Now, as far as the phase is concerned once again, the x axis will be log omega and a y axis will be the angle of the Blaschke Product. We note that, when omega is much less than a then the angle of the Blaschke Product is going to be approximately equal to 0 and when omega is exactly equal to a, we would have tan inverse of omega by a to be equal to tan inverse of 1, which is equal to pi by 4 and hence the angle of b is going to be equal to minus 2 times pi by 4 which is minus pi by 2.

So, angle of b is going to be equal to, minus pi by 2 at omega equal to a. Now, in the limit that omega tends to infinity or when omega is much greater than a we note that the term tan inverse of omega by a will tend to 10 tan inverse of infinity which is pi by 2 and hence the angle of b will tend to minus 2 times pi by 2 or in other words minus pi. So, by making note of the angle of b at these 3 particular limits of omega; we can now draw a rough phase characteristic for the Blaschke Product

So, if I were to mark out the point a, on the bode plot. We note that at the location omega equal to a the phase lag is going to b minus pi by 2 and for frequencies much less than a the phase lag is going to close to 0; the phase is going to decrease in this particular manner, is going to reach minus pi by 2 at omega equal to a and then, it is going to continue decreasing as omega tends to infinity and it is going to assume tactically approach the angle minus pi.

So, as omega tends to infinity the phase lag, will attend to minus pi, but it will never cross the line minus pi. So, this is the phase characteristic of the Blaschke Product and you can immediately see that is phase characteristic once again adds an extra phase lag to the overall, phase characteristics of the minimum phase part of the plant namely P MP of s without affecting it is magnitude characteristics.

So, in that respect therefore, it does something similar to what a time delay does to the overall system. Even in case of the time delay we have no modification of the magnitude characteristics of the minimum phase part of the plant it is just that the phase lag gets increased by an amount omega times capital T where, capital T is the time delay and in this case where we have a nonminimum phase 0; we note that once again the magnitude characteristic remains, unaffected.

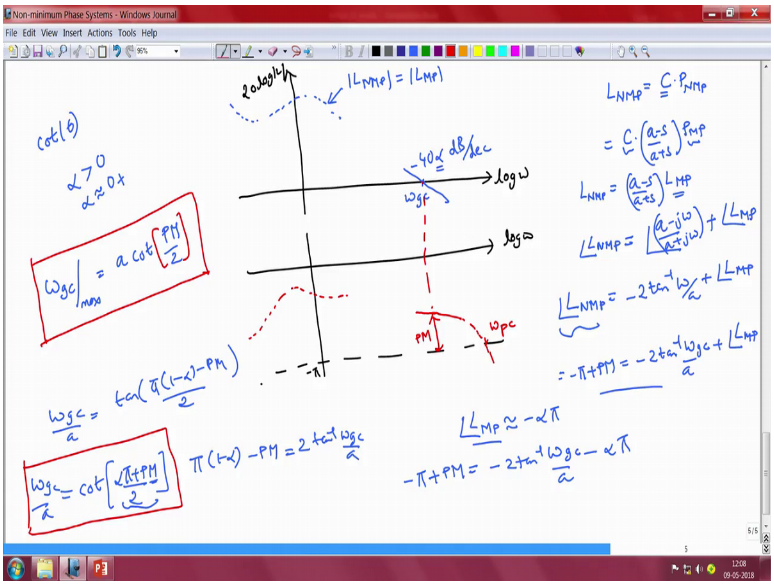
So, the magnitude characteristics of P minimum phase is going to be equal to the magnitude characteristics of P nonminimum phase; however, the phase characteristics of P Nonminimum phase is going to have an extra phase lag compared to that of the phase characteristics of P minimum phase by the amount given by minus 2 tan inverse omega by a; and once again we see that since the phase lag goes from 0 to minus pi. This term the Blaschke Product forces a plant to undergo phase crossover.

So, if we have a first order plant for instance if, P minimum phase is a first order system, then ordinarily this system would not have a phase crossover frequency at all because, the phase goes from 0 to minus pi by 2; however, if we were to cascade such a first order, minimum phase plant with the Blaschke Product there will be a particular frequency at which the phase will cross over. And hence we will have to start to worry about the stability of the closed loop system even, when we are using fairly simple controllers such as a proportional controller.

So, the problems that a, a nonminimum phase term a nonminimum phase 0 causes for stability has very close parallels with the kind of problems that we encountered in the case of the time delay. So, I shall not take extra numerical examples to underscore the problems that this particular term would cause because, the magnitude and the phase characteristics of this term have certain parallels with the magnitude and phase characteristics of the time delay and once and hence one can easily predict and anticipate the kind of problems that this particular kind of phase characteristic; where the phase lag goes from 0 to pi will cause to a closed loop system it the problems temps essentially from the fact that this kind of a phase lag forces the open loop system to have a phase crossover frequency that is much less than the natural phase crossover frequency of the minimum phase part of the plant it itself.

Now, both the nonminimum phase 0 as well as time delay as we discussed imposed certain fundamental limitations on our abilities as control engineers. So, let us now see how one can derive these fundamental limits that are imposed by these terms. And I want to underscore at the outset that the gain phase relationship that was discovered by hw bode plays a central role in coming up with the approximate expression for the Fundamental limitations that we would be talking about in connection with these nonminimum phase terms.

(Refer Slide Time: 12:08)



So, what I shall do next is to draw the bode plot of a general open loop system which has a nonminimum phase 0 unit. So, the magnitude plot will have 20 log, magnitude of L plotted versus log of omega; and let me assume, that the magnitude plot does something at low frequencies. So, depending on the poles and zeros of the open loop system the magnitude might reduce or increase and so on and so forth. So, I shall not draw it is exact characteristics; the exact characteristics are of course, dependent on the structure of the plant as well as the control of that we are using, but no matter what it does there needs to be a particular frequency at which the gain will cross over because of plants gain will reduced and the controllers gain will also reduced beyond a certain frequency.

There will be a certain frequency omega gc; the gain cross over frequency at which point the, the magnitude crosses over from some value above 0 db to some value less than 0 db. Now, I shall also draw the phase characteristics associated with the corresponding magnitude characteristics. So,, I have drawn the line theta equal to minus pi as the dashed line here.

So,, at very low frequencies the phase does something and then, I shall once again not drawn the exact phase characteristics at very low frequencies because it once again depends on the exact structure of the plant and the controller. Now, near the frequency omega gc let us say, there is a certain phase margin. So, at the frequency omega gc the phase lag is such that it gives us a certain phase margin PM; and then the phase continues to decrease and finally, crosses over at some frequency omega pc. So, let us say, this is the phase characteristic and this is the general magnitude characteristic. Now, in the vicinity of the frequency omega gc, let us assume that the slope of the magnitude characteristic is minus forty alpha decibels, per decade. So, let us say that the poles and zeros of the plant and the controller are such that, we get a slope of minus 40 alpha decibels per decade near the gain cross over frequency.

Now, the first point I wish to make in connection with the bode plot the general bode plot of a nonminimum phase loop gain that I have drawn here is that if, we were to define L NMP, as the nonminimum phase loop gain which is going to be equal to the controller, times the plant, P NMP. Then we can write L NMP has, being equal to C times a minus s by a plus s, times P minimum phase; now, assuming that our controller is also a minimum phase controller, then, we can lump the terms C and P minimum phase and write that as a minus s, by a plus s, times L minimum phase.

So, in other words L nonminimum phase can be related to L minimum phase which is defined as C times P minimum phase by the expression that has been shown here; it is equal to the Blaschke Product times the transfer function L minimum phase. Now, I want to first point out that the magnitude characteristics of L nonminimum phase; which is what has been plotted at the top there is go to be exactly identical to the magnitude characteristics of L minimum phase that is because the Blaschke Product has a gain of one at all frequencies.

owever, the phase lag of nonminimum L nonminimum phase that is the angle of L NMP is going to be equal to the angle of the Blaschke Product which is a minus, j omega angle of a minus j omega by a plus j omega, plus the angle of L minimum phase. Hence, the angle of L nonminimum phase is going to be equal to minus 2 times, tan inverse of omega by a as we discovered in our previous slide plus the angle of L minimum phase.

Now, let us apply this particular equation at the gain crossover frequency. Now, at the gain crossover frequency we note that our phase of the nonminimum phase loop gain is going to be given by minus pi, plus the phase margin. So, suppose I have specified a certain phase margin, then the angle of L, L nonminimum phase at the gain crossover frequency by definition is going to be equal to minus pi plus PM; and that is going to be equal to minus 2 times, tan inverse of omega gc by a, plus the angle of L minimum phase.

Now, we can make another important simplification to the equation that I have returned here; we note that if, our magnitude characteristic is rolling off at minus 40 alpha decibels per decade. So, the magnitude characteristic I want to remind you corresponds to both that of the nonminimum phase loop gain as well as the minimum phase loop gain they both have the same magnitude characteristic. So, this minus 40 alpha decibels per decade also happens to be the magnitude characteristic of the minimum phase loop gain L MP. Know if, you have a minimum phase loop gain L MP whose magnitude is reducing at the rate of minus 40 alpha decibels per decade then bodes gain phase relationship which is applicable to minimum phase loop gains tells us, that the angle associated with this roll off which is the angle of L MP is going to be approximately equal to minus alpha pi.

So, this we get from bodes gain phase relationship which tells us, that the phase characteristics at any particular frequency is determined approximately almost entirely by a slope of the magnitude characteristic in the vicinity of that particular frequency. So, since the slope in the vicinity of omega gc is minus 40 alpha decibels per decade the angle of L minimum phase is approximately going to be equal to minus alpha pi. Now, if we substitute this into this equation we would have that the phase of my, nonminimum phase loop gain which is minus pi plus PM, is going to be equal to minus 2, tan, inverse, omega gc by a minus alpha pi.

Now, if we rearrange this equation we would have that pi, times 1 minus alpha minus PM is going to be equal to 2 times tan inverse, omega gc by a, or in other words, omega gc, by a, is going to be equal to tan of pi, times 1 minus alpha minus, PM, whole divided by 2. Now, by simplifying this further we would note that, the expression on the right hand side is going to be equal to cot of alpha pi, plus PM, divided by 2.

So, our gain crossover frequency is related to the location of our nonminimum phase 0 a and the specified phase margin PM and the role of near the gain crossover frequency alpha, according to the equation that we have given here. Now, this is an important equation, because it allows us to work out workout the best possible or the highest possible gain crossover frequency. Why is that so? That is so, because we note that, for any variable theta, the function cot theta, is a decreasing function of theta. So, cot theta assumes it is maximum value or theta equal to 0, at which point it is equal to infinity and as theta tends to pi by 2 cot theta will tend to 0 monotonically. Hence for the gain crossover frequency to be a maximum, we should have that term within the square brackets namely alpha pi plus PM by 2 to be as small as possible. So, the smaller the term alpha pi plus PM by 2 is the larger will be the term on the left hand side namely, omega gc by a. Since a is a constant that will automatically mean that omega gc will get maximized. Now, if you focus on the term within the square bracket namely alpha pi plus pm by 2; we note that the phase margin is something that you might have specified we want in the interest of stability our open loop system to have a certain phase margin.

So, we cannot play around with the term PM, all that is left with for us to play around with is the term alpha. So, the smaller the value of alpha the higher will be the gain crossover frequency and we desire a fairly high gain crossover frequency. So,, so, that our closed loop bandwidth can be adequately high and with that high closed loop bandwidth you will be able to reject a large range of disturbances or track a large range of references that is the purpose behind our objective to maximize omega gc.

So, the smaller the value of alpha the larger will be the value of omega gc. And the question now is, what is the smallest permissible value of alpha? To answer this question we need to focus on the magnitude characteristic near the gain crossover frequency. Since, by definition we note that we have the magnitude characteristic crossing over at omega gc our alpha here has to necessarily be some positive value if, alpha is a negative value then it means that the slope near the gain crossover frequency is positive which means that the gain will not reduced as function of frequency and hence will not even cross over. Therefore, we need alpha to be greater than 0; and the smallest value that alpha can assume is some number which is approximately equal to 0, but actually slightly greater than zero.

So, this is the minimum value that we can permit alpha to assume because if, alpha assumes values less than 0 then we are in real trouble; because our gain characteristics will not even reduce from sum value above 0 db to sum value below 0 db. In order for gain crossover to happen we need to slope near the gain crossover frequency to be negative which implies therefore, that our alpha has to be necessarily greater than 0 and the smallest value that alpha can assume therefore, is a value that is slightly larger than 0. So, if, that is the case if alpha is allowed to assume a value added slightly larger than 0 that gives us an estimate of the maximum possible gain crossover frequency and that is given by omega gc max and that is from this equation from the equation within the red box if you go to set alpha to be approximately equal to 0. So, that we can ignore the term alpha pi in relation to the term pm we would get omega gc max to be equal to a times cotangent of PM by 2. So, what is analysis reveals therefore, is that there is a maximum gain crossover frequency possible for a control system, which has a nonminimum phase 0 unit and that is given by this particular expression a cotangent hyperbolic of PM by 2.

And I want to underscore that, this is the first time, in this set of lectures that you are come across this possibility that there is such a thing as an upper limit to the gain crossover frequency. In all the design that we have undertaken so far both in the case of 2 degree of freedom control design as well as in 1 degree of freedom control design we have assume, minimum phase plants and for all such plants there was no such thing as an upper limit to the theoretical upper limit to the achievable gain crossover frequency.

It could be as large as 1 would wish to be and you can still achieve whatever phase margin we wanted to have for the closed loop system it is just that we had to invest in a more expensive controller that had a much wider bandwidth. But here, what this analysis reveals is that no matter how much money you are willing to pour? In designing of the controller there is nothing you can do to improve the bandwidth of the closed loop system. It is theoretically limited to the amount given by a cot PM by 2.

Now, in this analysis we assumed that our alpha was approximately equal to 0 and that essentially implies that the gain margin for the open loop system in this particular limit which gives us the maximum permissible, gain crossover frequency is actually equal to 0 db. So, we can extend this analysis in order to estimate the best possible gain crossover frequency when we also have specified a certain, gain margin. So, let us undertake this analysis next.