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

Lecture - 48
Consequences of actuator bandwidth limitations while controlling unstable systems

Hello, in the previous clip, we started our discussion on control of unstable systems. And in this clip we shall continue this discussion. So, we saw that unstable systems could possess Bode plots whose phase characteristics make it difficult for us to decide, what the correct conditions are for the closed loop system to be stable. So, subsequently we found that Nyquist plot is the one plot which helps us determine whether the closed loop system is stable or not. So, we have to revise first here rules for determining whether a closed loop system is stable or unstable in the Nyquist plot. And then we examine the Nyquist plot for the kind of open loop gain that we considered unstable loop gain.

And then found out for what case the closed loop system is stable. And subsequently we also discovered that they were fundamental limits to the gain crossover frequency of an unstable system. So, just as in the case of a non-minimum phase system, there was an upper limit to the achievable gain crossover frequency. In the case of an unstable system there is a lower limit to the gain crossover frequency. In other words, the gain crossover frequency has to at least be a certain value in the interest of stability in order to achieve a certain phase margin specification.

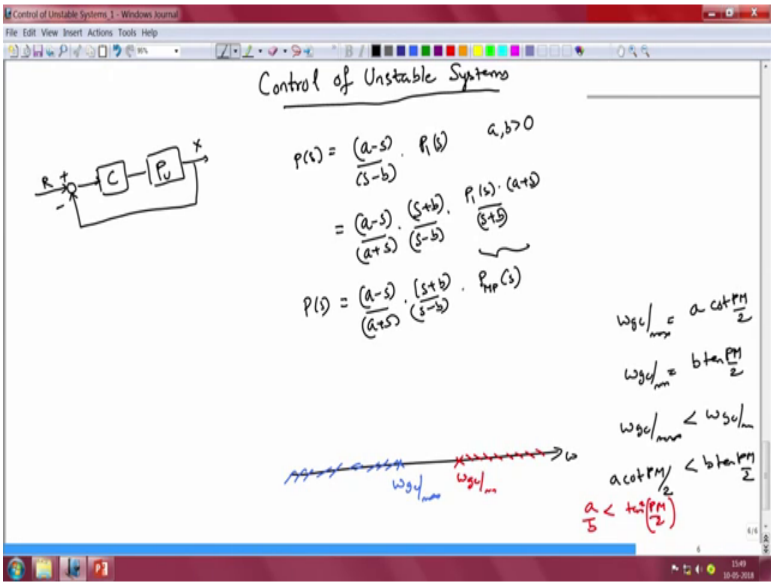
If you want a certain gain margin also in addition to the specified phase margin, then the minimum gain crossover frequency necessary will be even larger. So, as a rule of thumb we noticed that the minimum gain crossover frequency necessary for the closed loop system to be stable is generally several multiples of the location or the numerical value of the unstable pole of the open loop system. So, the fact that there has to be a minimum bandwidth necessary for an unstable system has some unfortunate consequences.

The first is the additional sensitivity of such systems to measurement noise. So, as we have discussed in the past, we want to minimise the bandwidth of the closed loop system, so that it is able to achieve the performance specifications. But beyond that frequency range within which we were expecting performance for the closed loop system, we want the loop gain of the closed loop system to be reduced as fast as possible, but without of course, a jeopardizing on the stability of the closed loop system. But what we discover in case of unstable open loop systems is that the closed loop system has to have a certain minimum bandwidth, no matter what the frequency range of this disturbance or reference we want to track or reject.

So, for instance, if even if we are interested in rejection of disturbances or tracking of references a frequencies very close to DC or at very low frequencies, we still need to maintain the closed loop bandwidth to be more than omega g c min, so which was given by a time p m by 2 or in other words if you multiples of the location of the unstable pole, so that is one unfortunate consequence.

The second even more tricky situation arises. When we are stuck with a plant which has both unstable dynamics or in other words a right half plane pole as well as a non-minimum phase term such as a time delay or a right half plane zero. So, in this case the right half plane zero imposes an upper limit on the achievable gain crossover frequency, while the unstable pole imposes a lower limit on the achievable gain crossover frequency.

(Refer Slide Time: 04:06)



For instance, if we consider a plant of the form P of s is equal to a minus s by s minus b times P 1 of s, where the terms a and b are both greater than 0. In other words, this is a plant which has a right half plane 0 as well as a right half plane pole unstable dynamics as well as non-minimum phase behaviour. And it is assume that the transfer function P 1 of s is a minimum phase transfer function all of its poles and 0s on the left half of the complex plane. Then this can be written as P of s is equal to a minus s by a plus s times s plus b by s minus b times P 1 of s times a plus s divided by s plus b.

So, all I have done in this case is to multiply and divide the right hand side by the same expressions namely a plus s and s plus b and this is the expression that we get. So, we note that the first two terms correspond to the familiar Blaschke products. And then we have the other term namely P 1 of s times a plus s by s plus b. So, now, we can call this part of the transfer function namely P 1 of s times a plus s by s plus b as P minimum phase P and P of s. So, we would therefore, have that the plant P of s would be given by a minus s by a plus s times s plus b by s minus b times P minimum phase of s.

Now, the Blaschke product a minus s by a plus s as we discussed imposes an upper limit on the achievable gain crossover frequency. In other words if this is the frequency axis, let me call as omega, we would have omega g c max to be equal to a cot PM by 2. So, this is the maximum achievable gain crossover frequency for a specified phase margin PM. If we want a certain gain margin, is going to be even smaller than this particular value. Now, since the system has unstable dynamics, the unstable dynamics demands a certain minimum gain crossover frequency and that is given by omega g c min equal to b time P M by 2.

Hence, the non-minimum phase 0 imposes an upper limit on the gain crossover frequency. The unstable poled imposes a lower limit on the gain crossover frequency. And it may or may not always be possible for us to satisfy both the specifications. For instance, if omega g c max, is less than omega g c min or in other words a cot PM by 2 is less than b time PM by 2. What that means essentially on the frequency axis is that there is a certain frequency, which we can call as omega g c max which is the maximum allowed frequency for us to have a certain phase margin, and with a only achieve gain crossover frequencies of magnitude less than that now, because you have unstable dynamics thus a minimum gain crossover frequency necessary and that we call omega g c min.

And we have we can only achieve can crossover frequencies a value greater than or equal to omega g c means in order for our closed loop system to be stable. Now, if these two permissible regions of our the gain crossover frequency do not intersect. What it means is that, it is not possible for us to have a control system. No matter what kind of controller we design that will give us a specified phase margin PM. So, in other words if a by b is less than tan square PM by 2, then we are in big trouble. It is not possible for us to design a closed loop system with the specified phase margin PM.

So, in other words even if we have unstable dynamics and non-minimum phase 0, the location of the non-minimum phase 0 a should be much greater than the location of the unstable dynamics b in order as to be able to design closed loop systems with respectable amounts of responsible extends of stability. Now, you might wonder under what circumstances would we have plants which exhibit both unstable dynamics as well as posses non-minimum phase behaviour. It turns out that such plants are more common than we imagine.

In fact, the humble bicycle is one which is both unstable as well as exhibits non-minimum phase dynamics. In fact, the steering dynamics of a bicycle shows non-minimum phase behaviour, and all of us have are aware that a bicycle is an unstable system. And we have learnt it the hard way, when we were trying to learn riding a bicycle. So, a bicycle is of course an unstable system, and it also exhibits non-minimum phase behaviour as far as its steering is concerned.

Now, one can design bicycles in such a manner that the location of the unstable pole, and the non-minimum phase 0 are such that it is not possible for us to have adequately high phase margins for the closed loop system that we might want to come up with to control the bicycle. In such a case the bicycle would become an un rideable bicycle one example of an un rideable bicycle is a bicycle which is rear steered.

(Refer Slide Time: 10:25)

So, I have shown here the photograph of a rear steer bicycle. So, what you see is that is bicycle for all intention purposes looks like a regular bicycle with the exception that the steering handle bar, which is shown here is connected to the rear wheel through this chain and sprocket arrangement here. Now, you can show that if you have a rear steer bicycle, then the location of the pole and 0 of this bicycle are such that are so close to one another that we cannot get appreciable phase margins for the closed loop system. And hence our closed loop system can never be in practice, can never be stabilized.

So, as a general rule of thumb if the ratio between a and b in other words if the ratio of the location of the non-minimum phase 0, and the unstable pole is less than 5, then such systems are so difficult to stabilize that it is not worth investing effort in designing controllers to stabilize such systems. Even if one what to try and attempt to stabilize systems, it is very likely that we would have very small phase margins. And therefore there is a very real threat of instability of such systems. So, as a rule of thumb, therefore if we have an a system that has both unstable dynamics as well as shows non-minimum phase behaviour, then the ratio of the non-minimum phase 0 a to the unstable pole b should be significantly greater than 5.

Now, when we talked of the stability of this bicycle, we notice that it was difficult to stabilize it, because of the particular locations of the non-minimum phase 0, and the unstable dynamics of the bicycle. This can be remedied by reducing the location of the unstable poles. So, if unstable pole is brought closer to the origin or in other words the unstable pole b is made smaller in magnitude, that it is possible to achieve this ratio of a by b to be large enough for us to stabilize the bicycle.

(Refer Slide Time: 12:41)

And are precisely what has been done in this new bicycle design which once again is a rear steel bicycle as you can see here. The handlebars are connected to the rear wheels of the bicycle, but this time the centre of gravity of the bicycle has been raised to such an extent that the location of the unstable pole is closer to the origin than for the bicycle that I showed in the previous slide. And hence for this bicycle a ratio of the non-minimum phase 0 to the unstable pole can be made large enough that we can stabilize it.

And therefore we can ride this bicycle this is also why it is possible for clowns in circus for instance to ride unicycles. If you have noticed the kind of unicycles that clowns in circuses right they are usually cycles with very tall handles, and the person who is riding it is sitting at the very top of this unicycle, why is that so that is so, because if this person is seated that far up on the unicycle, then the location of the unstable pole is going to be closer to the origin or in other words unstable dynamics is going to be quite slow, it is going to be quite take quite some time for the bicycle to flip over. And, this allows the person enough time to quickly correct for any tendency on part of the unicycle to flip over.

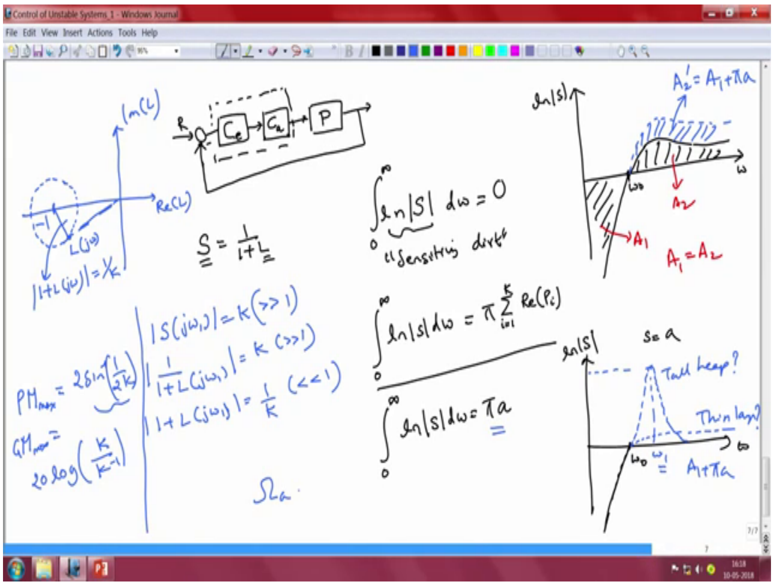
So, when you are therefore trying to control unstable systems the bandwidth or the speed of response of the actuator, when you are trying to control the system has to be adequately large, and if that is not so, then it may not even be possible to control the unstable system. So, as another example similar to the one that I just gave where, you had a unicycle with a very tall seat which helps the rider to balance the unicycle better. One can also think of balancing a stick on ones hand as part of your childhood experience you might be aware that it is easy to balance a tall stick on your hand, but significantly more difficult and sometimes even impossible to balance shorter sticks.

So, if I have a stick that is as short as this pen or even shorter, for example a match stick, it is practically impossible for us to stabilize this stabilize the stick. Why is that so, fundamentally it is so because larger sticks or longer sticks have very slow unstable dynamics, and that allows us we are the actuators in this case our hand is actuator our hands motion can be fast enough to compensate for any tendency on part of a long stick to get destabilized from its nominal position. Whereas, if you have very short stick its dynamics is very fast, it can flip over very quickly, and our hands are not fast enough to prevent any tendency on part of a short stick from flipping over insufficiently first time in order for it to state put at its nominal position.

So, this discussion brings us to the centrality of the bandwidth of actuators in the control of unstable systems. And sometimes if the bandwidth of the actuators is not chosen to be adequately large, it can have disastrous consequences. And on a few occasions in the recent past the importance of the bandwidth of actuation, when one is tend to control an unstable system has been often looked has been over looked of an even by professional controls designers. So, it is well worth visiting this topic, and understanding the limits of achievable performance in terms of stability and understanding the limits of achievable stability metrics such as the phase margin and the gain margin.

If you are given the location of the unstable pole, and if you are specify the bandwidth of the actuator, that you have to control this open loop unstable system. So, this analysis relies heavily on Bode sensitivity integrals. And Bode sensitivity integrals allow us to derive very simple relationships that can be used to predict the best possible phase margin and a best possible gain margin, if you know if you are given the bandwidth of your actuator and if you are told the location of your unstable pole for the plant. So, let us now discuss the limits on stability imposed by unstable dynamics in the presence of actuator bandwidth limitations. And we shall do this as I said by using Bode sensitivity integrals.

(Refer Slide Time: 17:29)



So, if we revisit the control one degree of freedom control system that we have been looking that so far, we have a controller and in all cases we assume that is controller was an electronic system that realised or implemented the control laws that we might choose to design, but in practice the controller can itself we look that as a cascade of 2 systems. One is the electronic subsystem of the controller which I shall call as C e, C subscript e which might result in a computer or which might be an electronic circuit that implements the control law that me chose to achieve.

And the output of this is fit to an electromechanical element which converts some voltage input, so I have a displacement or a velocity change for some such or it affects the physical variables that we are trying to control. And this element this electromechanical element is of course called as the actuator. So, we can call this as a C a. So, our controller, therefore is a cascade of the electronic subsystem, and the actuators of system. And the output of the actuator of course is fit to a plant.

The plant is moved by the actuator, and the output of the plant is measured by the sensor and is fed back in the interest of feedback control of the plant. So, the references provided here so that two blocks C e and C a put together actually form our controller. Now, in most cases the electronic subsystem of the controller is very fast is fast enough, that it is not going to limit the closed loop band width of our feedback control systems. This is especially true when we are controlling fairly big mechanical systems or processes industries, whose rate of change whose dynamics is very slow compare to dynamics of electronic elements.

The actuator however has limited bandwidth, because it is an electromechanical device, it could be a motor, it could be a valve, it could be an element of the kind. And the band width limit on the actuator intern imposes limits on the performance of the closed loop system. And this is particularly of concern for unstable systems, because there is possibility of a real accident to happen.

So, to understand this, let us now use Bode sensitivity integrals. We had visited Bode sensitivity integrals, if you clips back, in connection with the fundamental properties of the loop gain. And we had derived it in the case of minimum phase loop gains. And there we saw that the sensitivity function s satisfies this particular equation ln of mod of s b omega from integrated from 0 to infinity is going to be equal to 0. So, this was the expression that we had for Bode sensitivity integrals in the case of minimum phase loop gains or in other words, where sensitivity function S is equal to 1 by 1 plus L. And we assume that the poles and 0s of a l are all on the left half of the complex plane.

Now, we also talked of this integral being equal to 0 as quote unquote the law of conservation of sensitivity dirt. We called this ln of magnitude of s as sensitivity dirt, because generally as feedback control engineers we want sensitivity to be as small as possible, because if sensitivity is very small, if capital s is very small, it means that the loop gain is very large. And if loop gain is very large, then we are happy as control engineers, because that will allow us to reject disturbances, achieve robustness to plant parameter variations, achieve good small errors to tracking references and so on and so forth.

So, we want sensitivity to be very small, and hence we called it sensitivity dirt, because we would like dirt around us to also be small, but then what is equation told us, is that you cannot make the sensitivity dirt go away everywhere along the frequency axis. So, for instance, if the x axis is omega, and the y axis is ln of magnitude of S, we note that in the interest of performance you might choose to reduce the sensitivity over some frequency range, let me call that frequency as omega naught. So, below omega naught we might want to have a low sensitivity now what that means, is that if you have reduce sensitivity in this particular frequency range, then the exact area that we have reduced here has to be added on to the rest of the frequency range, so that we would have higher sensitivity in the frequency range between omega naught and infinity.

So, we cannot make the sensitivity dirt go away everywhere along the frequency axis. If we dig a trench in a certain frequency range in the interest of improving our control performance there, then we have to pile up the sensitivity dirt at a different location along the frequency axis, so that the net area of under the curve or the net quantity of sensitivity dirt that we have dug within one frequency range is going to be equal to the net quantity of sensitivity dirt that we have added on in some other frequency range.

So, these two areas if I were to call this area as A 1, if I want and this area as A 2 we have A 1 should be equal to A 2 in magnitude. Now, it so happens that if we have A open loop system which exhibits unstable dynamics. So, in other words if the loop gain of our system has poles on the right half of the complex plane, then the expression for the sensitivity integral gets modified somewhat. So, if you have an open loop system that is unstable, then the sensitivity integral namely ln of magnitude of S d omega will not be 0, but instead would be equal to pi times the sum of the real parts of all the poles that are on the right of the complex plane.

So, if you have k poles on the right half of the complex plane, then the Bode sensitivity integral would be equal to pi times summation from I equal to 1 to k the real parts of the poles pi of the open loop system which are on the right of the complex plane. So, this is how the Bode sensitivity integral equation would look, if you have unstable dynamics as part of your open loop system. Now, in the particular case that we have been discussing so far; we have assumed that there is an unstable pole at s is equal to a. So, for our particular case you would have the Bode sensitivity integral or in the case when you have one unstable pole a one pole on the right half of the complex plane, you would have the Bode sensitivity integral to be given by integral ln of magnitude of S d omega is equal to pi times a.

Now, what is means is that if we dig a trench, up to some frequency omega naught or in other words is a reduce sensitivity dirt ln of magnitude of S over some frequency range up to some frequency omega naught, then the pile of dirt that we have to add on in the rest of the frequency range which is a between frequencies omega naught and infinity is larger in height or is larger in magnitude than the case for the for the minimum phase system. So, the dirt pile is taller in case of the unstable system and a total dirt that we have to have in frequency range between omega naught and infinity is larger than the dirt that we would have in case of the minimum phase system by the amount pi times a.

So, in this case for instance in the case of a unstable open loop system. We would have that A2 this new area will be equal to A 2 dash equal to A 1 plus pi times A that is the total amount of dirt that we would have in the frequency range beyond omega naught. Now, the question we want to answer is the following. So, we unavoidably have to increase the sensitivity over some frequency range, when we reduce sensitivity over another frequency range.

And a extent of increase in case of the unstable system is higher by the amount pi times a. The question is how do we distribute sensitivity dirt in the frequency range between omega naught and infinity. So, there are two possibilities. So, if I were to draw once again the ln of magnitude of S capital S as function of frequency, up to omega naught we are expecting feedback performance, and therefore we have reduced the we have dug up a pit there, and reduce the sensitivity dirt in this frequency range up to omega naught.

Now, beyond omega naught what do we do, now that we have this excess dirt namely the dirt of quantity A 1 plus pi times a how do we distribute it over the rest of the frequency range. There are two possibilities one is we can make a tall hip of this dirt at some particular frequency. So, we can basically get all this dirt to be located within some frequency range and make a tall hip of this dirt this is one possibility. The other is we can distribute as a thin layer over the entire frequency range from omega naught to infinity. So, this is a thin layer this is a tall hip. Now, which of these two should be chose as control engineers, should be chose to distribute this sensitivity dirt in the form of a tall hip we have some frequency what should be distributed uniformly as a thin layer over the entire frequency range.

Now, to answer this question of whether to distribute it in the form of a tall hip or a thin layer let us, look at the consequences of in the first case distributing it in the form of a tall hip. So, if you have a tall hip there, it means there is some frequency omega one at which ln of magnitude of S is a very large number or equivalently the magnitude of s itself is a very large number. So, let me call that number as k which is much greater than one. So, we would have that magnitude of S at j omega 1 is going to be equal to k which is much greater than 1.

Now, what is the consequence of having a very large sensitivity in the neighbourhood of some frequency omega 1. So, we note that that by definition sensitivity is given by 1 1 by 1 plus L hence we would have the magnitude of 1 by 1 plus L of j omega 1 to be equal to k which is much greater than 1 or equivalently we would have that the magnitude of 1 plus L of j omega 1 is going to be equal to 1 by k, and since k is much greater than 1 we would have that 1 by k is going to be much less than L.

Now, what does it mean if you have a loop gain of such a manner of such a kind, that 1 plus l is very small is a very small number. So, to see the implication of this let us go back to our Nyquist plot. So, in the Nyquist plot we are plotting, of course the real part of L versus the imaginary part of L. And we have the critical point minus 1 at this location here. So, some point here in the complex plane represents the complex number L. And hence 1 plus L is represented by the complex number starting at the point minus 1 and ending at the point ending at this particular location namely that of L of j omega. So, this complex number is going to be equal to 1 plus L of j omega.

Now, the magnitude of 1 plus L is equal to 1 by k specifies that at the frequency omega 1. The distance of L of j omega 1 from the point minus 1 is going to be equal to 1 by k or in other words the loop gain of L of j omega 1 lies on a circle of radius 1 by k centre at the point minus 1. And if k is a very large number, then this radius, so this magnitude of 1 plus L by j omega is 1 plus L L of j omega is going to be equal to 1 by k. So, if k is a very large number the radius of this circle is going to be a extremely small, and what that in turn means is that the loop gain L of j omega 1 will be extremely close to the critical point

And if the loop gain is extremely close to the critical point, then we know that our closed loop system is on the verge of instability or in another words power stability margins are going to be extremely small. Now, we can show that if we have the distance of our loop gain L of j omega 1 to be equal to 1 by k from the critical point, then the best achievable phase margin which I shall call as P M max can be given by 2 times sin inverse of 1 by 2 k, so this is the best achievable phase margin. And the best achievable gain margin G M max is given by 20 log of k by k minus 1.

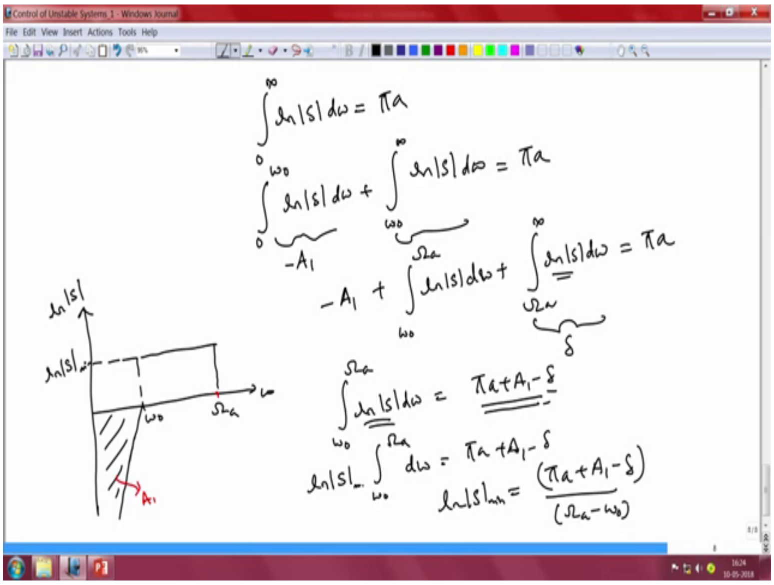
So, if k is a very large number, we note that our phase margin is going to be a very small number, because 2 times sin inverse of 1 by 2 k will be approximately equal to simply 1 by k. When k is a very large number, and therefore phase margin will be a very small number. And likewise the gain margin also will be a very small number, because when k is very large k and k minus 1 will be very nearly equal in magnitude.

And hence you would have 20 log of k by k minus 1 to be very close to 0 dB. So, a very large value k is therefore synonymous with very very small phase margin, and very small gain margin, and the fact that our closed loop system would be on the threshold of instability. Hence, if you come back to this decision that we have to take namely whether to distribute the sensitivity dirt in the form of a tall heap or as a thin layer, the answer from this analysis is quite clear. We do not want tall heaps, because tall heaps are synonymous with low stability margins.

So, we want to distribute this sensitivity dirt in as thin a layer as it is practically possible. Now, what is the frequency range that we have available for us to distribute a sensitivity dirt is infinity, because we have the entire frequency range from omega naught to infinity in which to distribute is dirt of amount A 1 plus pi times a. And since we have this infinite stretch of frequency in which to disturb this dirt you might argue that we have no issues associated with its distribution, because we can distribute it as an infinite assembly thin layer over the entire frequency range. But, this makes an important assumption which is not valid.

If you wanted distribute a sensitivity dirt as a uniform layer of infinite assembly small thickness over the entire frequency range, it means implicitly that we can control the thickness of the sensitivity dirt over the entire frequency range from omega naught to infinity, but in practice that is not the case. There will be a frequency which we shall call as capital omega a which is decided by the bandwidth of the actuator beyond which we cannot control the loop shape or in other words we cannot control the way in which the sensitivity function would vary as function of frequency. So, let us now rewrite the Bode sensitivity integrals taking into account this particular constraint.

(Refer Slide Time: 34:38)



So, we have the Bode sensitivity integral to be that integral 0 to infinity ln of magnitude of S d omega is equal to pi times a. And we can split this integral as integral 0 to omega naught ln of magnitude of S d omega plus integral omega naught to infinity l n of magnitude of S d omega. And we noted that integral 0 to omega naught ln of magnitude of S d omega we called it as A 1, and there is a negative sign, because we have reduced the sensitivity to less than 0 dB in this frequency range. And the rest of it together should add up to pi times a.

Now, the second integral can further be split up in to two integrals, one is the integral from omega naught to capital omega a which is the bandwidth of the actuator and that. So, we would have integral from omega naught to capital omega a ln S d omega plus integral capital omega a to infinity ln of magnitude of s d omega to together comprise the integral ln of omega integral omega naught to infinity ln of magnitude of S d omega. So, this minus A 1 should be equal to pi times a.

Now, as we discussed it is not possible for us to control the thickness of the sensitivity that namely ln of magnitude of S for frequencies beyond omega a, because of actuator bandwidth limitations. So, if the actuator bandwidth is only omega a, there is nothing we can do to control the shape of the loop L of j omega for frequencies beyond omega a. And hence this integral namely omega a to infinity ln of magnitude of S d omega is something whose value we cannot control by using whatever controller we might choose.

Hence, I shall call this particular term as some constant delta which is going to be determined by the particular controller or the actuator that we choose for controlling our unstable system.

So, what we would have, therefore is the integral omega naught to capital omega a ln of magnitude of S d omega is equal to pi a plus A 1 minus delta. So, we have this quantity of sensitivity dirt namely pi a plus A 1 minus delta which has to be distributed between the frequencies omega naught and capital omega a. So, let us mark that out here graphically. So, the y axis is l n of magnitude of s the x axis is linear omega. So, there is a frequency omega naught below that we have dug it trench and reduce sensitivity dirt, and the total quantity that we have reduced is as we discussed A 1.

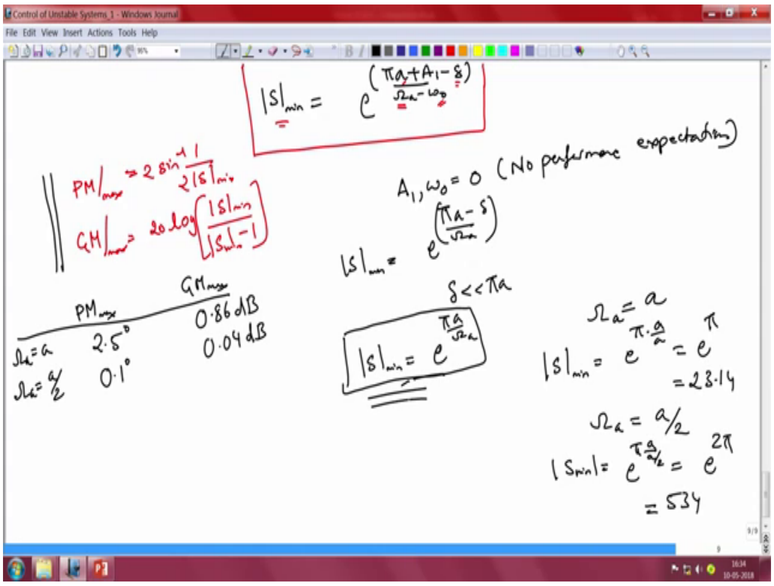
Now, and beyond the frequency omega a we cannot control the thickness of the sensitivity dirt, and that the integral of sensitivity dirt our total quantity of dirt beyond a frequency omega a we called as delta. So, what we have to distribute between the frequency is omega naught, and capital omega a is this particular quantity pi a plus A 1 minus delta.

Now, the question is how do we distribute it in this frequency range? We should distribute it in such a manner that the maximum height of the pile of dirt is minimised, because if we have you know dirt getting piled up any particular frequency in this frequency range, it means that we will have a larger sensitivity at that frequency, and larger sensitivity as we discussed is synonymous with poor gain margin, and phase margin of stability specifications.

So, we want to distribute the dirt between these two frequencies in such a manner that the maximum height of the dirt pile is minimized. Now, a moment thought would reveal that the only way to do it is to make sure that the height of the dirt over this entire frequency range is the same which is equal to some ln S minimum ln of magnitude of S minimum. So, this is going to be the minimum height of the dirt pile between the frequencies omega naught and omega a. And what is its value, its value, it is easy to obtain assuming that we have distributed the dirt such that it is of constant height between omega naught and omega a.

So, this term inside the integral is going to be constant. So, we can remove it from the integral, and write it as ln of S minimum times integral omega naught to omega a d omega is equal to pi times a plus A 1 minus delta or in other words ln of magnitude of S minimum is going to be equal to pi times a plus A 1 minus delta divided by capital omega a minus omega naught.

(Refer Slide Time: 39:56)



And this in turn can be further simplified to give us a expression that the magnitude of S minimum is going to be equal to e to the power pi times a plus A 1 minus delta divided by capital omega a minus omega naught. Now, this is an important expression, because this expression will allow us to compute the best achievable gain margin and phase margin specifications.

If you are given the bandwidth of the actuator omega a, and the location of the unstable pole namely small a, the extent of benefit in terms of feedback performance we want to achieve in terms of capital A 1, and the frequency range up to which you want to achieve this performance namely the frequency omega naught. And we also have this extra term delta which represents the integral of the sensitivity function or sensitivity dirt for frequencies beyond capital omega or beyond the bandwidth of the actuator.

Now, to see how we can put this to use we first need to note that we would like this magnitude of S minimum to be as small as possible. The smaller it is then the better or is a phase margin and gain margin specification. So, for instance the phase margin maximum as we discussed is equal to 2 times sin inverse of 1 by 2 k or in another words 2 times sin inverse of 1 by 2 times magnitude of S minimum. Likewise, the best achievable gain margin was given by 20 times log of magnitude of S minimum divided by magnitude of S minimum minus 1.

So, we would like to have a high gain margin and a high phase margin which implies that our magnitude of S minimum has to be as small as possible. And this equation here which we have obtained by using Bode sensitivity integrals reveals the dependence of magnitude of minimum on the different parameters of our control system. To see how we can minimise it, we note that if we stop expecting any performance what so ever from our control system, in another words we are not really interested in performance at all, let us focus simply on stability.

And let us see what needs to be done to improve the stability margins of our closed loop system, if that is all that our goal is going to be. We note that we can make the frequency omega naught up to which we are expecting performance. And the quantity A 1 which is what we would have in order to get the particular control performance we want, we can set them both to 0, if you are not expecting any control performance at all.

If you do that immediately, we see that the term on the right hand side will reduce. So, we would have therefore, that if our A1 and omega naught are set equal to 0, so in other words no performance expectation, we would have that the magnitude of S minimum is going to be equal to e to the power pi a minus delta by capital omega a.

Now, for the sake of simplicity I shall assume that the term delta is much less than pi a. And for this case we would have that the magnitude of S minimum is going to be this very simple expression given by e to the power pi a divided by capital omega a. So, this expression will tell us what is the smallest possible value for a magnitude of S that we can achieve for a specified actuator bandwidth omega a, and the specified location of the unstable port.

And if we have this information, we can plug it into the two equations that I have written here on the left, and compute the best possible phase margin and gain margin specifications. So, let us compute the best possible phase margin and the gain margin for some typical cases. For instance, if we choose an actuator which is just as fast as that of the unstable dynamics or in other words, if we choose an actuator whose band width omega a is exactly equal to a, then we would have that the magnitude of S minimum would be equal to e to the power pi times a by a or in other words e to the power pi.

And we can show that e to the power pi is equal to 23.14. So, if the magnitude of S minimum is equal to 23.14, we can use the two equations on the left top corner to calculate the best possible phase margin and the best possible gain margin. So, let me write that down here. So, P M max and G M max and when we have that omega a is equal to a. So, for this particular case the best possible phase margin can be computed to be equal to 2.5 degrees. And a best possible gain margin is going to be equal to 0.86 dB. So, you notice the startlingly low values of the best possible phase margin and a best possible gain margin. When we choose an actuator, there is exactly as fast as our unstable dynamics.

Suppose, however we made the mistake and chose an actuator whose bandwidth omega a was actually half of the location of the unstable pole or in another words omega a was equal to a by 2 small a by 2, in which case we would have that the magnitude of s minimum would be equal to e to the power pi times a divided by a by 2 or in other words e to the power 2 pi. If you compute the numerical value of e to the power 2 pi, you get it to be approximately equal to 534.

And if you substitute this value of 534 in the equations on the top left corner to compute the best possible phase margin and a gain margin, we find that that omega a is equal to a by 2. Or in another words our actuator bandwidth is half the value of our unstable dynamics, then we would have that the best possible phase margin is going to be merely 0.1 degrees, and the best possible gain margin is going to be equal to 0.4 dB.

So, what this therefore, reveals is that if we choose actuators whose bandwidth is comparable to the location of the unstable pole, then for all practical purposes we would have closed loop systems are going to be unstable. It is there is no way we can practically achieve a feedback control system whose phase margins are going to be this small namely 2.5 degrees or 0.1 degrees or gain margins being that small, and still manage to get stable closed loop operation.

If you recollect in the past when we performed control design, we chose to have phase margins of at least 40 degrees in the interest of stability, on the other hand with actuators that are of this particular bandwidth namely of bandwidth comparable to the unstable poles value. We would have phase margin and gain margin are exceedingly smaller than the typical and a minimum values that we assumed earlier in the interest of safety of operation.

So, what I also want to point out is the elegance, and a power of the technique of Bode sensitivity integrals in deriving is best achievable stability specifications. If you notice this equation that has been enclosed in the black bracket, there are only 2 bits of information that we are using as far as the open loop system is concern. We are just depending on our knowledge of the actuator bandwidth and the location of the unstable pole to quickly determine, what the best achievable phase margin and gain margin specifications are. It does not matter what controller we use, it does not matter where the other plants poles and 0s are.

The best achieved performance specifications almost entirely get determined by the location of the unstable pole and the bandwidth of the actuator. And if you use simple controllers, then our best achievable phase margin and gain margin will be even worse than this. Because, the kind of sensitivity as function of frequency that gives this particular equation here as the magnitude of S minimum is shown here in this graph on the bottom left corner.

And we note that the dependence of sensitivity on frequency is some fairly strange dependence which cannot be realised using simple controllers. Hence, if one way to use simple controllers, then the best possible phase margin and the gain margin are going to be even worse than what we have accomplished through this analysis. Unfortunately, though the importance of Bode sensitivity integrals has reduced with type to the point where professional control engineers are often not even exposed to the motion of Bode sensitivity integrals. Even though these integrals are extremely powerful, and help us to compute very quickly, and with very with bare minimum information about the plant the best possible stability specifications. And this has had some disastrous consequences.

(Refer Slide Time: 49:58)

Figure 1. *Gripen JAS39 prototype accident on 2 February 1989. The pilot received only minor injuries.*

For instaces, what you see in this photograph is the event of an accident that happened to this particular fighter jet Gripen JAS39. So, fighter jets are designed to be open loop and stable, because it helps them with manoeuvrability. However, the design was done in such a manner that it was not possible for the actuators that were chosen to stabilize this open loop unstable aircraft to actually stabilize them with adequate phase margin and gain margin.

Since, the actuators are chosen to be rather slow and the fighter jet was design to be open loop unstable in the interest of improved manoeuvrability, we had a problem on our hands. The phase margin and a gain margin were too low, and as a result even before this aircraft took off the aircraft crashed, and it caught fire. Fortunately, though the pilot of this aircraft escaped unhurt. So, I want to remind you again that the control system for is aircraft was designed by professional control engineers, but unfortunately though these control engineers were presumably not exposed to Bode sensitivity integrals, and its importance, when it comes to computing the best possible performance. Since people have tended to forget the significance of this result, since the time that Bode first proposed this.

(Refer Slide Time: 51:31)

Figure 2. *Chernobyl nuclear power plant shortly after the accident on 26 April 1986.*

The second example is far more tragic, and this has to do with the Chernobyl nuclear power plant meltdown. So, nuclear power plants are also unstable systems. And nuclear power plants have a water jacket that takes away the heat generator by the nuclear reaction. However, if we are not able to take away the heat fast enough, then the entire nuclear reactor can over heat and meltdown. Now, this can happen, if you have bubbles in the water that is used to take away the heat. If you have air bubbles or air pockets forming in the water, then the thermal conductivity of the overall liquid will get reduced, because air is a bad conductor of heat.

And hence, our heat cannot be taken away as fast enough as one would like it to be. So, if you have a pump that cannot take away heat as fast enough or in other words an actuator that cannot remove heat away as quickly as is necessary from a nuclear reactor, then once again we will have an unstable closed loop system on our hands. And the in this case has resulted in nuclear meltdown with disasters consequences to the people and animals living near it.

So, people have called the Chernobyl nuclear disaster as one of the worst industrial disasters in the history of humanity. And it is a sobering thought that the blame for this disaster can be laid almost entirely at a feet of feedback control engineers. And the ignorance of the limitations of feedback control systems, especially when such control systems have to control unstable plants such as the nuclear power plant.

So, it is well worth our effort to remember, therefore the fundamental limitations that exist when we are trying to control certain types of plants. And the possible danger associated with controlling especially unstable plants such as a nuclear power plant or an unstable aircraft or some other such unstable system. And be aware of the tools that can be employed in order to compute the best possible performance that one can achieve, if one chooses a particular actuator to control the systems.

Thank you.