Slovak University of Technology in Bratislava
Institute of Information Engineering, Automation, and Mathematics

PROCEEDINGS

17$^{\mathrm{th}}$ International Conference on Process Control 2009

Hotel Baník, Štrbské Pleso, Slovakia, June 9 – 12, 2009

ISBN 978-80-227-3081-5

http://www.kirp.chtf.stuba.sk/pc09

Editors: M. Fikar and M. Kvasnica

Šebek, M., Hurák, Z.: An often Missed Detail: Formula Relating Peek Sensitivity with Gain Margin Less Than One, Editors: Fikar, M., Kvasnica, M., In *Proceedings of the 17th International Conference on Process Control '09*, Štrbské Pleso, Slovakia, 65–72, 2009.

Full paper online: http://www.kirp.chtf.stuba.sk/pc09/data/abstracts/115.html

AN OFTEN MISSED DETAIL: FORMULA RELATING PEEK SENSITIVITY WITH GAIN MARGIN LESS THAN ONE

M. Šebek and Z. Hurák

Department of Control Engineering, Faculty of Electrical Engineering
Czech Technical University in Prague
fax: + 420-224-918-6460, e-mails: m.sebek@polyx.cz, hurak@fel.cvut.cz

Abstract: An inequality relating gain margin with sensitivity peek value is presented in numerous basic control textbooks. In fact, this inequality fails to hold as soon as the open-loop Nyquist plot crosses the negative real axis on the left of the critical point. This opposite case is usually ignored by the textbook authors. A simple alternative inequality is derived in the paper to cover the not so popular opposite case. This fills a small gap one often encounters in basic control courses.

Keywords: Gain Margin. Sensitivity. Complementary sensitivity. Nyquist Plot.

1 INTRODUCTION

Several modern control textbooks provide a simple inequality relating gain margin *GM* to the peak of sensitivity function M_S in their sections devoted to frequency domain design specifications. Along with another inequality relating similarly phase margin and M_S, this approach makes it possible to express traditional design specs in a unified manner using only M_S.

This inequality can be found in modern control textbooks e.g. (Skogestad *et al*. 2005), (Seborg *et al*. 2004) etc. Unfortunately, it appears to work only in the case of *GM* > 1 when the open-loop transfer function Nyquist plot crosses the negative real axis on the right of the critical point (-1,0). In fact, it is invalid in the opposite case, when the open-loop transfer function Nyquist plot crosses the negative real axes on the left of the critical point so that *GM* < 1.

In such an opposite case, the standard inequality must be replaced by similar yet different one. Its derivation is so simple that the authors consider their contribution minor. On the other hand, they believe that this minor but often encountered gap should be filled. The authors are not aware of any paper or textbook presenting the opposite-case inequality but would not be surprised to learn that it has been published anyway.

2 GAIN MARGIN AND PEAK SENSITIVITY

As usually, we denote the open-loop transfer function by $L(s)$ and the closed loop sensitivity function by

$$S(s) = \frac{1}{1+L(s)} \tag{1.1}$$

throughout the paper. For a particular frequency $\omega \geq 0$, the distance d_ω of the corresponding open-loop Nyquist plot point $L(j\omega)$ from the critical point $(-1,0)$ in the complex plain reads

$$\begin{aligned} d_\omega &= \operatorname{dist}\left(L(j\omega),-1\right) = \left|L(j\omega)-(-1)\right| = \\ &= \left|L(j\omega)+1\right| = 1/\left|S(j\omega)\right| \end{aligned}$$

while the minimum distance $d_{\min}$ of the whole Nyquist plot of $L(s)$ from the critical point $(-1,0)$ is well known to be

$$\begin{aligned} d_{\min} &= \inf_\omega \left(\operatorname{dist}\left(L(j\omega),-1\right)\right) \\ &= \inf_\omega \left(\left|L(j\omega)+1\right|\right) \\ &= \inf_\omega \left(1/\left|S(j\omega)\right|\right) \\ &= 1/\sup_\omega \left|S(j\omega)\right| = 1/M_S \end{aligned} \tag{1.2}$$

Here, as often, M_S stands for the peek sensitivity function value or its H_∞ norm

$$M_S = \sup_\omega |S(j\omega)| = \|S(s)\|_\infty$$

It is evident that for any particular frequency ω is

$$d_\omega \geq d_{\min} . \tag{1.3}$$

If only the open-loop gain happens to be uncertain, the frequency $\omega = \omega_{180}$ at which the open-loop Nyquist plot $L(j\omega)$ crosses the negative real axis plays a crucial role. Then the distance of the negative real axis crossing point $L(j\omega_{180})$ from the critical point $(-1,0)$ denoted here by $d_{180} = \mathrm{dist}(L(j\omega_{180}), -1)$ gives rise to the classical concept of gain margin *GM*.

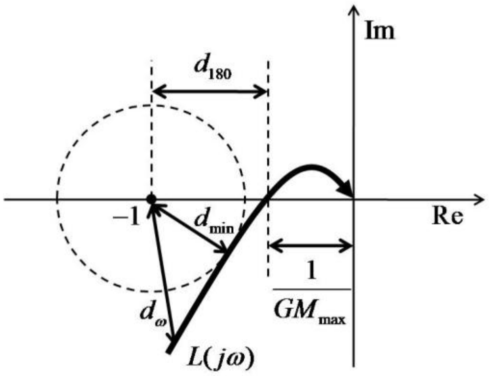

Fig. 1: The case of $GM > 1$ frequently encountered in textbooks

3 THE CLASSICAL CASE

Only the classical case is encountered in textbooks when $GM > 1$ because the open-loop transfer function Nyquist plot crosses the negative real axis on the right of the critical point as in Fig 1. Note the gain margin depicted in Fig 1 according to its standard definition. It such a situation, we denote the gain margin by $GM_{\max}$ to emphasize its physical meaning: A nominally stable closed loop remains stable even when the open-loop gain is multiplied by any factor k such that

$$k < GM_{\max}$$

The figure reveals immediately that

$$d_{180} + \frac{1}{GM_{\max}} = 1$$

which further implies

$$d_{180} = 1 - \frac{1}{GM_{\max}} \tag{1.4}$$

Putting together (1.2) and (1.4) with (1.3) yields

$$d_{180} = 1 - \frac{1}{GM_{\max}} \geq d_{\min} = \frac{1}{M_S} \tag{1.5}$$

and finally

$$GM_{\max} \geq \frac{M_S}{M_S - 1} \tag{1.6}$$

results. This is the inequality frequently encountered in textbooks. It turns out, however, that (1.6) fails to hold as soon $GM < 1$ as well as in other more complex cases.

4 THE OPPOSITE CASE

Let us now investigate the opposite case when $GM < 1$ and the open-loop Nyquist plot crosses the negative real axis on the left from the critical point. If this happens, the inequality (1.6) is not valid any longer. However, a similar yet different formula can easily be derived. The situation is illustrated on Fig 2.

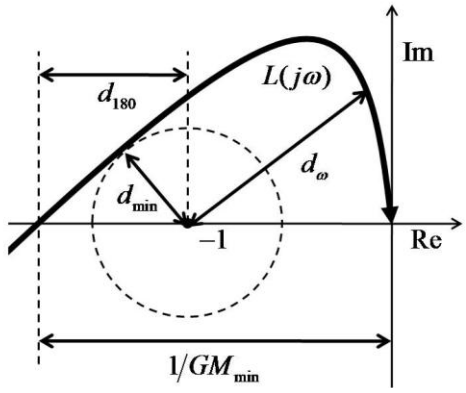

Fig. 2: The opposite case of $GM < 1$

In such a case, we denote the gain margin by $GM_{\min}$ to emphasize that a nominally stable closed loop remains stable even when the open-loop gain is multiplied by any factor k such that

$$GM_{\min} < k$$

By definition, $0 \leq GM_{\min} \leq 1$. It is clear from Fig 2, that now $d_{180} + 1 = 1/GM_{\min}$ so that

$$d_{180} = \frac{1}{GM_{\min}} - 1 \tag{1.7}$$

Combining (1.2) and (1.7) with (1.3) again implies

$$d_{180} = \frac{1}{GM_{\min}} - 1 \geq d_{\min} = \frac{1}{M_S} \tag{1.8}$$

This finally gives rise to the "opposite" formula

$$GM_{\min} \le \frac{M_S}{M_S+1} \tag{1.9}$$

This is the missing inequality to replace (1.6) in the opposite case. Although (1.9) is as simple and as practical as (1.6), this inequality is missing in the textbooks. Much worse, even the concept of two-sided margin as well as the fact of the two-sided crossing itself is ignored by the most of textbooks. One of notable exceptions is (Zhou, Doyle and Glover, 1996).

5 THE TWO-SIDED CASE

It is even possible that the open-loop Nyquist plot crosses the negative real axis both on the left form the critical point and on the right of it as in Fig 3.

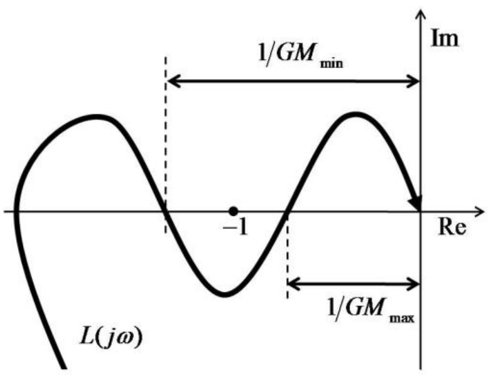

Fig. 3: The two-sided case

Whenever this happens, we must employ two different gain margins at the same time: the minimum gain margin $GM_{\min}$ as well as the maximum gain margin $GM_{\max}$. The case of two-sided crossing is illustrated by Fig. 3 where the both margins are indicated. Physical meaning of the two margins is as follows: A nominally stable closed loop remains stable even when the open-loop gain is multiplied by any factor k such that

$$GM_{\min} < k < GM_{\max} \tag{1.10}$$

However, the closed loop stability is lost as soon as either $k = GM_{\min}$ or $k = GM_{\max}$.

We proceed as above applying simultaneously (1.6) for $GM_{\max}$ and (1.9) for $GM_{\min}$. Putting (1.6), (1.9) and (1.10) together yields

$$GM_{\min} \le \frac{M_S}{M_S+1} < k < \frac{M_S}{M_S-1} \le GM_{\max} \tag{1.11}$$

Hence one can replace (1.10) by another, possibly narrower, interval described only by means of M_S as follows

$$\frac{M_S}{M_S+1} < k < \frac{M_S}{M_S-1} \tag{1.12}$$

6 THE GENERAL CASE

Should the open-loop Nyquist plot cross the negative real axis several times on the left and/or right side of the critical point (-1,0), only one crossing on each side counts (the closest one to the critical point, of course). The above reasoning holds true when applied to these closest neighbors. Such a situation is illustrated by Fig. 4.

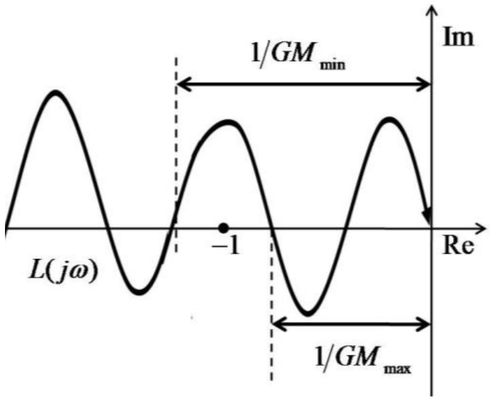

Fig. 4: Multiple crossings on the both sides

7 USING COMPLEMENTARY SENSITIVITY

Similar formulae can be derived the peak value of complementary sensitivity

$$T(s) = \frac{L(s)}{1+L(s)} \tag{1.13}$$

Comparing definitions (1.1) and (1.13), it is easy to see that

$$T(L) = \frac{L}{1+L} = \frac{1}{1+1/L} = S(1/L) \tag{1.14}$$

So one can simply repeat all the derivations above, replacing $S(s)$ by $T(s)$ and, at the same time, exchanging the Nyquist plot of $L(s)$ by the Nyquist plot of its reciprocal function $1/L(s)$. Such a way, it can be proved that

$$GM_{\max} \ge 1 + \frac{1}{M_T} \tag{1.15}$$

Here, as usually, M_T stands for the peek value of the complementary sensitivity function or for its H_∞ norm

$$M_T = \sup_{\omega} |T(j\omega)| = \|T(s)\|_\infty$$

The inequality (1.15) is encountered in textbooks as often as (1.6). However, it again holds true only for $GM > 1$. In the opposite case of $GM < 1$, the inequality (1.15) must be replaced by is "opposite counterpart"

$$GM_{\min} \leq 1 - \frac{1}{M_T} \qquad (1.16)$$

Finally, for the case of "two-sided crossing", a two-sided margin applies similarly to (1.10) giving rise to the "complementary version" of (1.12), which is

$$1 - \frac{1}{M_T} < k < 1 + \frac{1}{M_T} \qquad (1.17)$$

8 EXAMPLES

Example 1:
To prove that standard inequalities (1.6) and (1.15) indeed fail in the opposite case, just consider a trivial unstable open-loop transfer function

$$L(s) = \frac{2}{s-1} \qquad (1.18)$$

with Nyquist plot on Fig 5. The drawing reveals that

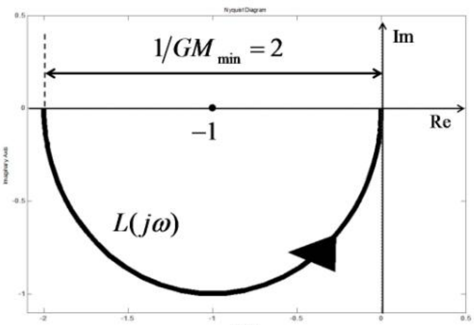

Fig. 5: The Nyquist plot of (1.18)

the negative axis crossing appears to be on the left of the critical point while $GM = GM_{\min} = 1/2$. The corresponding closed-loop sensitivity and complementary sensitivity functions are show on Fig 6 from which it is clear that $M_S = 1$ and $M_T = 2$. It is easy to check that both (1.6) and (1.15) fail

$$GM = \frac{1}{2} \ngeq \frac{M_S}{M_S - 1} = \infty$$

$$GM = \frac{1}{2} \ngeq 1 + \frac{1}{M_T} = \frac{3}{2}$$

while (1.9) and (1.16) do hold

$$GM_{\min} = \frac{1}{2} \leq \frac{M_S}{M_S + 1} = \frac{1}{2}$$

$$GM_{\min} = \frac{1}{2} \leq 1 - \frac{1}{M_T} = \frac{1}{2}$$

This result in fact justifies of the current paper.

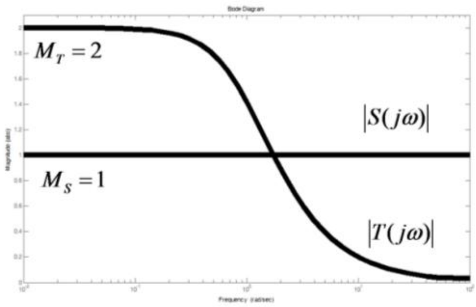

Fig. 6: Sensitivity and complementary sensitivity related to (1.18)

Example 2:
The case of two-sided crossing can be demonstrated by another quite elementary unstable open-loop transfer function

$$L(s) = \frac{2-s}{2s-1} \qquad (1.19)$$

Its Nyquist plot on Fig 6 indeed crosses the negative real axis twice and this happens on the right as well as on the left of (-1,0) giving rise to two-sided margin with $GM_{\min} = 1/2$ and $GM_{\max} = 2$. As $M_S = 2$

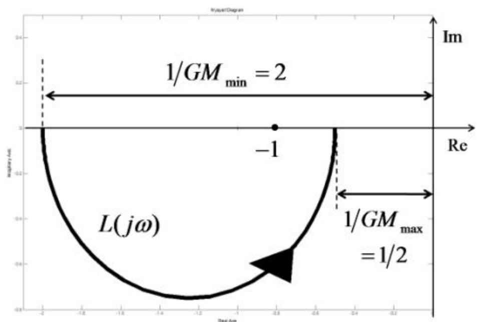

Fig. 7: The Nyquist plot of (1.19)

and $M_T = 2$ (see Fig. 8), we can apply (1.12) to learn that the closed-loop system remains stable even if the open-loop transfer function is multiplied by a factor k such that

$$\frac{M_S}{M_S+1}=\frac{1}{2}<k<\frac{M_S}{M_S-1}=2\,.$$

Alternatively, we can employ (1.17) to conclude that it remains stable for any multiplicative factor k such that

$$1-\frac{1}{M_T}=\frac{1}{2}<k<1+\frac{1}{M_T}=\frac{3}{2}$$

Note that a narrower interval for k results by chance when using M_T in this example.

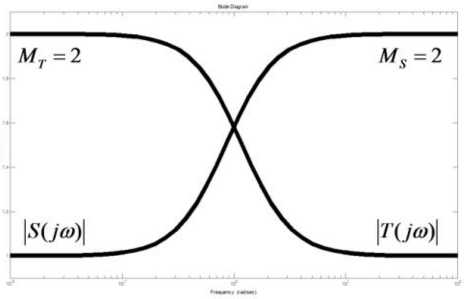

Fig. 8: Sensitivity and complementary sensitivity related to (1.19)

Example 3:
We end the section with example of a multiple negative real axis crossing. Nyquist plot of a complicated open loop transfer function

$$L(s)=\frac{7.9555(2-s)^3(s^2-0.5403s+0.1252)}{(2s-1)^3(s^2-4.9511s+7.3022)}$$

has been plotted by Matlab as follows

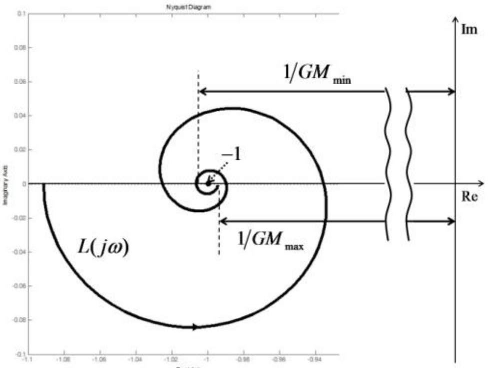

Fig. 9: Multiple crossing of the negative real axis

It crosses negative real axis six times: there times on the left and three times on the right of the critical point. To find the two-sided gain margin, only one crossing point should be considered on each side.

9 HOW CAN WE BENEFIT FROM ALL THE INEQUALITIES?

At first, (Skogestad *et al*. 2005) proposed using the peaks M_S or M_T to replace traditional measures for design specifications *GM (*and *PM*). For instance, requiring $M_S<2$ implies requiring $GM_{\max}>2$ (by (1.6)) and $PM>30°$ (by another formula not discussed here). Thank to the development above, we can tweak this claim by adding that requiring $M_S<2$ implies requiring $GM_{\min}<2/3$ as well (by (1.9)). Hence by specifying M_S one guarantees resulting robust stability gain interval to be at least (1.12), i.e.

$$\frac{M_S}{M_S+1}<k<\frac{M_S}{M_S-1}$$

This interval may be narrower than the classical

$$GM_{\min}<k<GM_{\max}$$

In reward, it is more lucid as it is using one variable M_S only. In addition, it is easier to handle and guarantee by modern loop-shaping techniques.

As another outcome, the above discussion sheds more light on the relation between gain margins and sensitivity peaks. Given M_S (as the design has already been made or for other reasons), what does it mean for the two-sided gain margin?

First the inequalities

$$GM_{\min}\le\frac{M_S}{M_S+1} \qquad (1.20)$$
$$0\le GM_{\min}\le 1$$

imply that

$$0\le GM_{\min}\le\frac{M_S}{M_S+1}\le 1 \qquad (1.21)$$

To put it in words, once M_S is given, the $GM_{\min}$ cannot be worse than $M_S/(M_S+1)$. In fact, it can only range the interval

$$\left[0, M_S/(M_S+1)\right] \qquad (1.22)$$

where the left bound is a dream while the right one is the worst case. This, however, tells us nothing about where within the interval $GM_{\min}$ is actually located. This evidently depends on other properties of the sensitivity function rather than just on its peak.

To have an idea, look at the following plot on Fig 10.

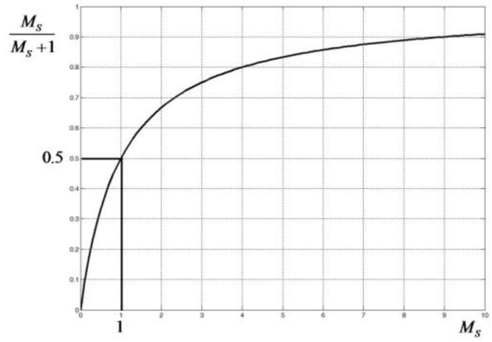

Fig 10: Dependence of the upper bound of $GM_{\min}$ on the sensitivity peak M_S

For example, $M_S = 1$ implies $GM_{\min} \in [0, 1/2]$. This means that $GM_{\min}$ cannot be worse than 1/2, but can be better, perhaps even 0. As another example, take a larger $M_S = 2$. This results in a larger interval $GM_{\min} \in [0, 2/3]$ allowing $GM_{\min}$ to rang up to $2/3$, that is larger and hence worse than before.

To summarize, for a given M_S, the inequality

$$GM_{\min} \le \frac{M_S}{M_S + 1} \qquad (1.23)$$

provides the worst case bound for $GM_{\min}$.

The inequalities

$$\frac{M_S}{M_S - 1} \le GM_{\max}, 1 \le GM_{\max} \qquad (1.24)$$

can be investigated similarly. Just consider the plot on Fig 11, which is a hyperbola with an asymptote at $M_S = 1$.

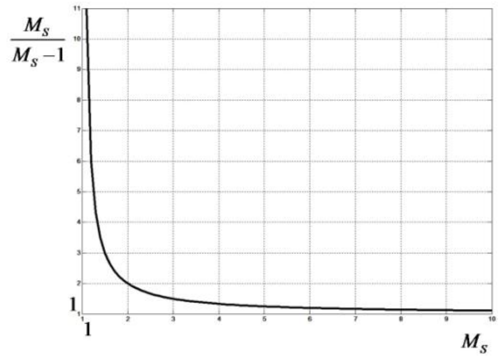

Fig 11: Dependence of the lower bound of $GM_{\max}$ on the sensitivity peak M_S

Since $1 \le GM_{\max}$ by definition and since $M_S \ge 1$ for any physical system, only the right branch of the hyperbola counts. So for $M_S \in [1, \infty)$ is always

$$\frac{M_S}{M_S - 1} \in [\infty, 1]$$

Here again (1.24) provide the worst case for $GM_{\max}$, this time the lower bound. For a given $M_S \in [1, \infty)$, $GM_{\max}$ cannot be worse (lower) than $M_S/(M_S - 1)$. Then it must range

$$[M_S/(M_S - 1), \infty) \qquad (1.25)$$

For example, if $M_S = 1$, then even the lower bound reaches ∞ and $GM_{\max}$ is fixed to be $GM_{\max} = \infty$. For a rising M_S, the lower bound becomes smaller. It falls down to 1 for large M_S until the interval (1.25) blows up to the definition interval

$$GM_{\max} \in [1, \infty)$$

and thereby, in fact, loses its purpose.

10 PEAK SENSITIVITY AND NONLINEAR ACTUATOR

Yet another benefit from the inequalities developed above is that they help to explain the impact of eventual actuator nonlinearity. Thanks to them, one can apply the Circle criterion for stability of nonlinear systems, as we show in this section.

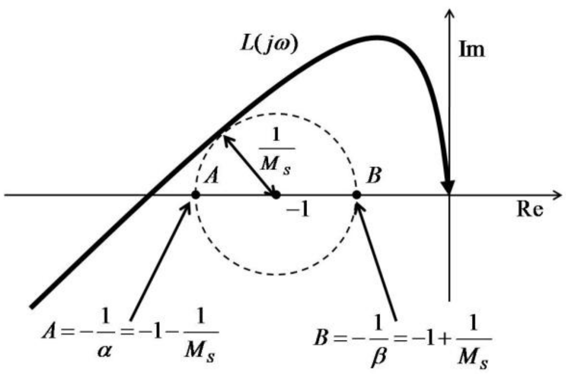

Fig 12: The circle definitely avoided by Nyquist plot

Regardless of where and how many times it crosses the negative real axis, the open-loop Nyquist plot never gets inside the circle around the critical point with radius $1/M_S$. This is guaranteed by the very definition of M_S and illustrated on Fig 12.

Assuming that the closed loop is stable, we are sure that the open-loop Nyquist plot encircles the critical point properly as many times as the Nyquist stability criterion requires. Then, however, it properly encircles the whole circle as well. This enable us to apply the Circle criterion for stability of nonlinear systems, see e.g. (Glad *at al.* 2001). Next, the criterion requires to find conditions that the nonlinearity must satisfy. They naturally depend on the position and circle size. To this end, we calculate points where the circle crosses the negative real axis. To be consistent with nonlinear control textbooks, we denote A, the left crossing point, by $-1/\alpha$ and B, the right crossing point, by $-1/\beta$. Then it is clear from Fig 12 that

$$-\frac{1}{\alpha}=-1-\frac{1}{M_S} \quad \text{and} \quad -\frac{1}{\beta}=-1+\frac{1}{M_S} \qquad (1.26)$$

yielding

$$\alpha=\frac{M_S}{M_S+1} \quad \text{and} \quad \beta=\frac{M_S}{M_S-1} \qquad (1.27)$$

This brings our favorite inequalities into the game.

Before stating the final result of this section, let us remind what was proven so far. We already know that if the closed loop is stable, then it remains stable even if the open-loop transfer function $L(j\omega)$ is replaced by $kL(j\omega)$ with k such that by (1.12)

$$\frac{M_S}{M_S+1}<k<\frac{M_S}{M_S-1} \qquad (1.28)$$

Now, using the Circle criterion (Glad *at al.* 2001) we claim even more. If the closed loop is stable, the it remains (globally asymptotically) stable even if we insert into the loop a nonlinearity $f(x)$ such that $f(0)=0$ and that for any $x\neq 0$ satisfies

$$\frac{M_S}{M_S+1}<\frac{f(x)}{x}<\frac{M_S}{M_S-1} \qquad (1.29)$$

The geometric meaning of (1.29) is clear from Fig 13

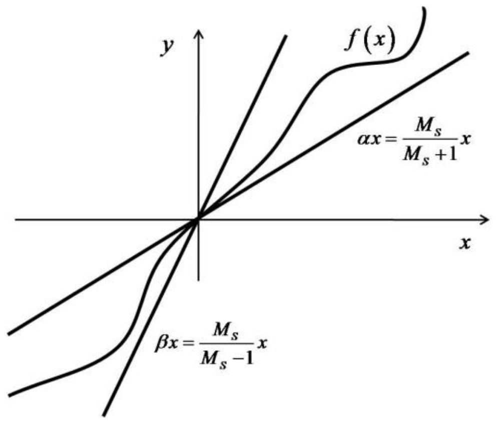

Fig 13: Nonlinearity and its bounds

The nonlinearity graph is confined to a cone shape region bound by two straight lines passing through the origin and having slopes

$$\alpha=\frac{M_S}{M_S+1},\beta=\frac{M_S}{M_S-1}.$$

respectively. The smaller is the sensitivity peak, the more spacious is the region and the more complex nonlinearity fits. It is then less likely that a nonlinear actuator violates overall stability.

11 "DO NOT ENTER" CIRCLE

As a byproduct, we win even better geometrical insight. It is crystal clear form Fig 14 what happens when applying the inequalities. In fact, we just trade the original interval given by the points where $L(j\omega)$ crosses negative real axis for another interval defined by the points where the circle crosses negative real axis. The new interval may be smaller, but it is nicer,

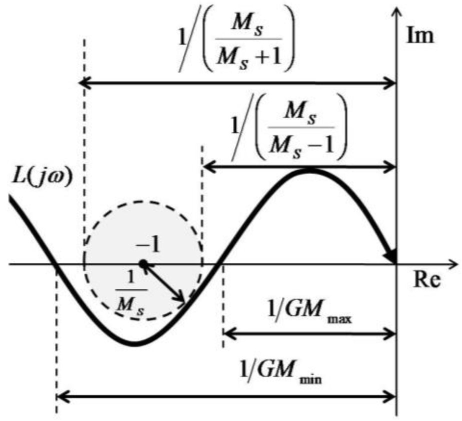

Fig 14: The final picture

easier to express and design. However, its main advantage is that not only the negative real axis segment is untouched by $L(j\omega)$ but the whole circle is intact. This is important as it guarantees true robust stability, regardless whether gain, phase or both are uncertain.

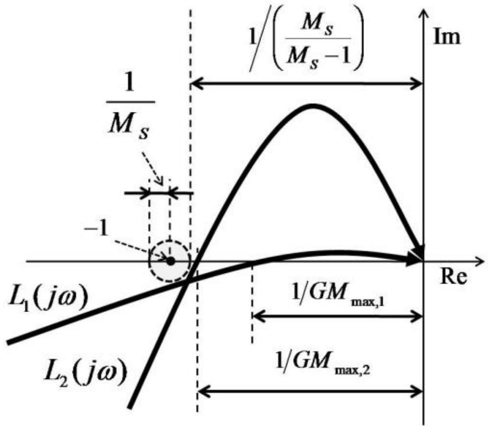

Fig 15: High sensitivity peak

Example 4:
Consider first the case of a high sensitivity peak as on Fig 15 where two different open-loop transfer functions are plotted having the same closed-loop sensi-

tivity peak M_S. As the peak is high, the "do not enter" circle is small and the stability robustness is poor. Note that although $L_1(j\omega)$ offers much larger gain margin than $L_2(j\omega)$ (indeed $GM_{\max,1} > GM_{\max,2}$) the same lower bound $M_S/(M_S-1)$ results from (1.6) leading to exactly the same new interval (1.25) or (1.11). Here one could wrongly conclude that more is lost for $L_1(j\omega)$ than for $L_2(j\omega)$. However, just the reverse is true: Robustness is evenly poor for $L_1(j\omega)$ and for $L_2(j\omega)$ and the larger gain margin $GM_{\max,1}$ really means nothing. The correct explanation for the difference is simply that the gain margin $GM_{\max,1}$ is a shoddy measure to evaluate the closed-loop stability robustness for $L_1(j\omega)$.

Example 5:
For change, consider now a perfect design with the smallest reasonable peak $M_S = 1$, for instance an LQR design for a double integrator that results in

$$L(s) = \frac{1+1.4s}{s^2} \qquad (1.30)$$

plotted on Fig 16. As $M_S = 1$, the "do not enter" circle is large enough to touch the origin. In such a case,

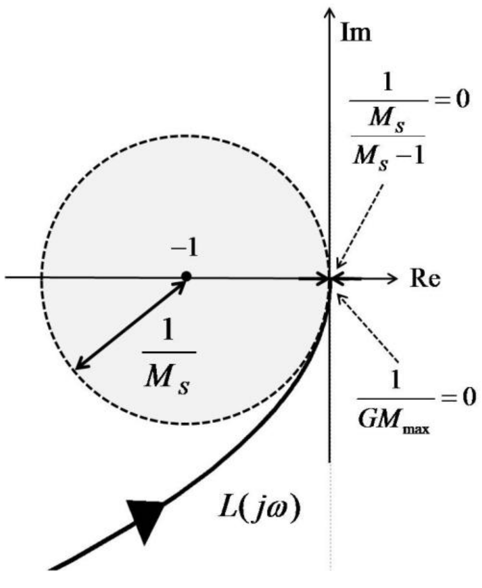

Fig 16: Low sensitivity peak

the inequalities (1.9) and (1.6) reveal good robustness with

$$GM_{\min} \le \frac{1}{2}, GM_{\max} = \infty$$

Since physical systems do not transfer infinitely high frequencies, their open-loop Nyquist plot always ends in the origin. Hence the "do not enter" circle can never cross the imaginary axis and the sensitivity peak M_S never drops below 1 for a physical system.

Example 6:
For a non-physical system or a reduced model or alike, it may well happen that $M_S < 1$. For instance

$$L(s) = \frac{s+2}{2s+1} \qquad (1.31)$$

gives rise to $M_S = 2/3$ and the "do not enter" circle becomes even larger having $1/M_S = 3/2$ in diameter. This is illustrated on the following Fig. 17.

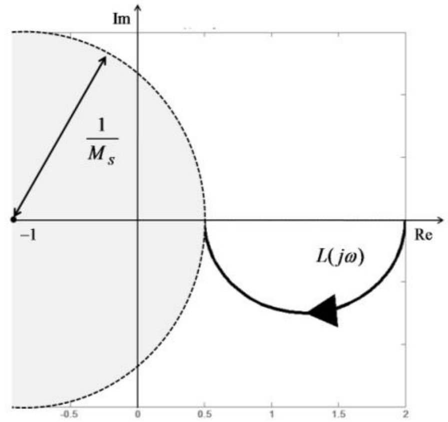

Fig 17: Sensitivity peak less than one

When the circle goes beyond ten imaginary axis, however, gain margin would be negative. This would require modifying the definitions. We postpone this to another paper and finish the exposition right here.

ACKNOWLEDGMENTS

The work has been supported by the by the Grant Agency of the Czech Republic grant No. 102/08/0186 and by the Ministry of Education of the Czech Republic under contract No. LA09039.

REFERENCES

Glad, T. and Ljung, L (2001). Control Theory: Multivariable and Nonlinear Methods. Taylor & Francis. London.

Seborg, D.E., Edgar T.F. and Mellichamp D.A. (2004). Process Dynamics and Control. John Wiley & Sons. New York.

Skogestad, S. and Postlethwait, I. (2005). Multivariable Feedback Control-Analysis and design. John Wiley & Sons. Chichester.

Zhou, K., Doyle, J.C. and Glover, K. (1996). Robust and Optimal Control. Prentice Hall. Upper Saddle River, NJ.