Describing functions and prediction of limit cycles

ZDENĚK ÚŘEDNÍČEK
Department of Automation and Control
Tomas Bata University in Zlin
Jižní Svahy, Nad Stráněmi 4511, 760 05, Zlin
CZECH REPUBLIC
urednicek@fai.utb.cz http://www.utb.cz/fai-en/structure/zdenek-urednicek

Abstract: - This paper deals with the so called describing functions method description as a simplified method of describing certain types of nonlinear systems, using the complex function of frequency response. The first part showed its using as an example of some so-called hard nonlinear systems (e.g. the mechanical chain of robots). In the actual last part, the formalization of the limit cycle prediction process is performed based on the representation of the non-linear element by describing function. The basic approach for this prediction is based on the application of an extended version of the well-known Nyquist criterion from linear control theory to a description of the systems with describing function utilization

Key-Words: - Motion control, describing function, mechatronics systems, multiport model, limit cycles, Nyquist criterion

1 Introduction

Inherent and inseparable part of the symptomatic subsystems of mechatronic systems (mechanical subsystem, subsystem of actuators, sensors subsystem) - including robotic systems – they are the specific nonlinearities occurring in their mechanical and regulatory subsystems. These specific, so-called hard nonlinearities, exemplified by non-viscous friction, saturation, backlash, hysteresis, etc., are often the cause of unwanted behavior of the system, but in some cases they are used as a tool for introducing specific desired system properties.

In the publication [11], a basic description of the using possibility of some powerful tool for linear control-frequency systems analysis and designing was performed. Linear system description utilization by a complex function, frequency response, instead of differential equations, is a tool that cannot be directly applied to a non-linear system because the frequency response cannot be defined for a non-linear system. However, for some nonlinear systems, an extended version of the frequency response called function-describing method can be used for the approximate analysis and prediction of non-linear behavior.

The main use of this method is to predict the limiting cycles of nonlinear systems, which will form the main content of this contribution. But the method has a number of other applications such as sub-harmonic prediction, jump phenomenon and non-linear system response to sinusoidal input.

2 Describing functions applications possibility

Let's first discuss briefly on what types of nonlinear systems the method is applicable and what type of information about the nonlinear system can provide.

2.1. On what types of nonlinear systems the method can be used

Simply put any system that can be transformed into the arrangement of Fig.1 can be studied using descriptive functions. There are at least two important classes of systems in this category.

The first important class is "almost" linear systems. By "almost" linear systems we mean systems that contain so-called "hard" nonlinearities in the control loop but are otherwise linear. These systems arise in the design of control law using a linear approach, but its implementation includes "hard" nonlinearities such as motor torque saturation, actuator (or sensor) backlash (dead band), Coulomb friction, or hysteresis in a controlled system.
Fig. 1 Non-linear system.

An example is the system of Fig. 2, containing hard non-linearity in the actuator.

Fig. 2 Control system with one "hard" non-linearity

The regulated system is linear as well as the controller. But the actuator contains hard nonlinearity. This system can be reconfigured to the form of Fig. 1 with

\[ G(s) = G_{reg}(s) \cdot G_{sys}(s) \cdot G_{sens}(s) \]

An "almost" linear system containing non-linearity in a sensor or controlled system can also be reconfigured to the shape of Fig. 1.

The second class of systems are systems containing real nonlinear subsystems whose dynamic equations can be converted to the structure of Fig. 1. We have seen an example of such a system in the [11].

2.2 Describing function applications

For systems as shown in Fig. 1, the limit cycle may occur as a result of non-linearity. But linear control cannot predict this problem. On the other hand, descriptive functions can conveniently be used to detect the existence of limit cycles and to determine their stability. Regardless of whether they are "hard" or "soft" nonlinearities. The use for limit cycle analysis is due to the fact that the shape of the system signal on the limiting cycle is commonly approximated by a sinusoidal one.

This can be conveniently explained on the system of Fig. 1. Suppose the linear part in Fig. 1 has the characteristics of the low pass filter (which is the case for many physical systems). If there is a limit cycle in the system, then the system signals must be all periodic. Because the periodic signal, as input to the linear part in Fig. 1, can be decomposed as the sum of many harmonic oscillations, and because the linear member, due to its low pass filter properties, fuses higher frequencies, the output \( y(t) \) must in most cases consist of the lowest harmonic oscillations. It is therefore reasonable to assume that the signals throughout the system are essentially sinusoidal, thus allowing the technique used in the previous section.

Limit cycle prediction is very important because limit cycles can occur in physical nonlinear systems. Sometimes the limit cycle may be desirable. This is the case for limit cycles in electronic oscillators. Another case is the so-called vibration technique to minimize the negative effect of Coulomb's friction in mechanical systems. On the other hand, in most control systems, the limit cycles are undesirable. This can be for several reasons:

1. The limit cycle is the path to instability, causing poor accuracy of regulation.
2. Constant oscillations associated with the limit cycle may cause increased wear or mechanical failure in the hardware of the control system.
3. The limit cycle may also cause other undesirable effects such as passenger discomfort during autopilot flight.

In general, although precise knowledge of the shape of the limiting cycle curve is not necessary, the knowledge of its existence or non-existence, as well as its approximate amplitude and frequency, is necessary. The method describing functions is applicable for these purposes. Knowledge of this kind can also lead to the design of compensators in order to avoid limiting cycles.

2.3. Basic assumptions of describing functions using

Let us consider the non-linear system in the general form of Fig. 1. In order to be able to use the basic version of the method describing functions, the system must meet the following conditions:

1. There is only one non-linear member.
2. A non-linear member is time-invariant.
3. In sinusoidal input \( x(t) = \sin(\omega t) \), only the fundamental harmonic can be considered in the output \( w \).
4. Non-linearity is odd function.

The first condition means that if there are two or more non-linear components in the system, they can either be joined to one (such as parallel pairing of two nonlinearities) or, only one the nonlinearity is under consideration and the other is neglected.
**The second condition** means that we only consider autonomous non-linear systems. This is sufficient for much practical nonlinearity, such as saturation of amplifiers, transmission backlash, Coulomb friction between surfaces and hysteresis in relay systems. The reason for this assumption is that the Nyquist criterion, on which the describing function is broadly based, requires linear time-invariant systems.

**The third condition** is essential for the describing function. It represents an approximation because the output from the nonlinear element at the sinus input usually contains, besides the basic, even higher harmonics. The assumption means that higher harmonics can be neglected in the analysis in comparison to the basics one. In order for this assumption to be fulfilled, it is important that the next linear element has the character of the low pass. I.e.

$$|G(i \cdot \omega)| >> |G(n \cdot i \cdot \omega)| \quad \text{for } n = 2, 3, \ldots$$

This means that higher harmonics in the output of nonlinearity will be significantly filtered. Thus, the third assumption is often called a filtering hypothesis.

**The fourth condition** means that the graph of the nonlinear relation $f(x)$ between the input and the output of this member is symmetrical with respect to the origin of the coordinate system. This assumption is introduced for simplicity, i.e. that Fourier development can neglect the DC component. Note that most of the nonlinearities occurring in our systems (robot motion system) meet this condition.

Failure to meet the above conditions has been widely studied in the literature on the use of general context descriptors such as multiple nonlinearities, time-dependent nonlinearity, or multiple sinusoids. However, these conditions relaxation-based methods are usually much more complicated than basic versions based on the above four conditions.

### 3 Basic definition

Consider the sinus input of a non-linear element with amplitude $A$ and frequency $\omega$. I.e.

$$x(t) = A \cdot \sin(\omega \cdot t)$$

as is shown in Fig. 3.

![Diagram of a non-linear element and its representation](Image 53x146 to 282x180)

**Fig. 3** The non-linear element and its representation by the describing component

Nonlinear element output is often a periodic but generally non-sinusoidal function. Note that this occurs whenever the non-linearity $f(x)$ is a uniquely invertible function because the output is

$$f[A \cdot \sin(\omega \cdot t + 2\pi)] = f[A \cdot \sin(\omega \cdot t)].$$

By using the Fourier series, the periodic function $w(t)$ can be expanded as

$$w(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cdot \cos(k \cdot \omega \cdot t) + b_k \cdot \sin(k \cdot \omega \cdot t)] \quad (1)$$

where Fourier coefficients are function of $A$ and $\omega$.

It is valid:

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} w(t) \cdot \cos(k \cdot \omega \cdot t) \cdot d(\omega \cdot t); \quad k = 0, 1, \ldots \infty$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} w(t) \cdot \sin(k \cdot \omega \cdot t) \cdot d(\omega \cdot t); \quad k = 1, \ldots \infty$$

As a result of the fourth of the above assumptions, it is $a_0 = 0$. Furthermore, the third condition means that you only need to consider the basic harmonics $w_i(t)$.

So

$$w(t) = w_i(t) = a_1 \cdot \cos(\omega \cdot t) + b_1 \cdot \sin(\omega \cdot t) = M \cdot \sin(\omega \cdot t + \varphi) \quad (3)$$

where

$$M(A, \omega) = \sqrt{a_1^2 + b_1^2}$$

$$\sin \varphi(A, \omega) = \frac{a_1}{\sqrt{a_1^2 + b_1^2}}; \cos \varphi(A, \omega) = \frac{b_1}{\sqrt{a_1^2 + b_1^2}}$$

because

$$a_1 \cdot \cos(\omega \cdot t) + b_1 \cdot \sin(\omega \cdot t)$$

$$b_1 = M \cdot \cos \varphi; \quad a_1 = M \cdot \sin \varphi \Rightarrow M = \sqrt{a_1^2 + b_1^2}$$

$$a_1 \cdot \cos(\omega \cdot t) + b_1 \cdot \sin(\omega \cdot t) =$$

$$= M \cdot [\cos(\omega \cdot t) \cdot \sin \varphi + \sin(\omega \cdot t) \cdot \cos \varphi] =$$

$$= M \cdot \sin(\omega \cdot t + \varphi)$$

The term (3) means that the basic harmonic corresponding to the sinus input is the sinusoidal function of the same frequency. Representing in a complex variable, it is possible to write this sinus as

$$w_i(t) = M \cdot e^{i(\omega t + \varphi)} = (b_1 + i \cdot a_1) \cdot e^{i\omega t}$$

Similarly to the frequency response concept of the linear system, which is the ratio of sinusoidal output to the sinus input in the frequency domain, we define the describing function of the nonlinear element as the **complex ratio of the fundamental harmonic output to the sinus input**. I.e.

$$N(A, \omega) = \frac{M \cdot e^{i(\omega t + \varphi)}}{A \cdot e^{i\omega t}} = \frac{M}{A} \cdot e^{i \varphi} = \frac{1}{A} \cdot (b_1 + i \cdot a_1) \quad (4)$$
By describing function which describes a non-linear element, this element - for the sinusoidal input - can be presented as a linear element with frequency transmission. This is shown in Fig.3.

The describing function concept therefore can be understood as extending of the frequency response term. For linear dynamic system the frequency transition function is independent of the amplitude of the input signal. But describing function of the non-linear element differs from the frequency transition function of the linear element by being dependent on the amplitude of the input signal. Thus, the representation of the non-linear element of Fig. 3 is sometimes called quasi linearization.

Generally the describing function depends on the frequency and amplitude of the input signal. There are, however, several special cases. If the non-linearity is a odd function, describing function is real and does not depend on the input frequency. Real describing function \( N(A, \omega) \) is the consequence of \( a_1 = 0 \) in this case.

4 Examples of discontinuous nonlinearities

Nonlinearities can be divided into continuous and discontinuous. Since discontinuous nonlinearities cannot be approximated locally by linear functions, they are often referred to as "hard" nonlinearities. These "hard" nonlinearities often occur in regulatory systems, both in small scale and large scale operations. Whether it can be considered as nonlinear or linear- when it is operating in a small scale of activity- the size of the "hard" nonlinearity arbitrates and the also the application of its effect on the performance of the system.

Due to the frequent occurrence of "hard" nonlinearities, let's briefly discuss the characteristics and effects of two important.

4.1 Describing function of the saturation

If the input of the physical device increases, it is often possible to see the following phenomenon. If input is small, its magnification leads (often proportionally) to increasing output. But when it reaches a certain value, its further magnification leads to little or no increase in output. The output simply stays close to its maximum value. We say the device is saturation in this state. A simple example is a transistor and a magnetic amplifier. Saturation type of non-linearity is commonly caused by limitations in component size, material properties, and limitation of available power. In Fig.8 there is a typical non-linearity of saturation, where the stronger line represents real non-linearity and thinner is its idealization - partial linearization.

![Fig. 4 Saturation nonlinearity](image)

Most actuators show saturation nonlinearity. For example, the output torque of the servo motor can not grow to infinity and exhibits saturation not only due to the properties of the magnetic material. Similarly, the torque (pressure) hydraulic servo motor controlled by valve is limited by the maximum accessible system fluid pressure.

Saturation may have a complicating effect on the properties of the control system. Simply, the occurrence of saturation reduces device gain (e.g. of amplifier) when the input signal increases. As a result, if the system is unstable in its linear part, the divergent behavior can be suppressed to permanent oscillations through signal. On the other hand, in a linear stable saturation system, the system's response drops because saturation reduces effective gain.

The input-output relationship for saturation non-linearity is illustrated in Fig. 5 with \( a \) and \( k \) as the determining parameters of non-linearity.

![Fig. 5 Non-linearity of saturation and its input-output relationship](image)
Because this non-linearity is an odd function, we assume that its descriptive function will be real and will be only a function of the input amplitude.

Consider the input
\[ x(t) = A \cdot \sin(\omega \cdot t) \]

If \( A \leq a \), then, the input remains all the time in the linear region and therefore the output is
\[ w(t) = k \cdot A \cdot \sin(\omega \cdot t) \]

The describing function is
\[ b_i = \frac{1}{\pi} \int_{-\pi}^{\pi} k \cdot A \cdot \sin(\omega \cdot t) \cdot \sin(\omega \cdot t) \cdot d(\omega \cdot t) = k \cdot A \]

And so
\[ N(A, \omega) = \frac{1}{A} \cdot (b_i + i \cdot a_i) = k \]

Let's think that \( A > a \). The input and output are then drawn in Fig. 5. Output on interval \( (-\pi, \pi) \) is
\[
\begin{cases}
  k \cdot A \cdot \sin(\omega \cdot t) & -\pi \leq \omega \cdot t \leq -\pi + \gamma \\
  k \cdot a & -\pi + \gamma < \omega \cdot t \leq -\gamma \\
  k \cdot A \cdot \sin(\omega \cdot t) & -\gamma < \omega \cdot t \leq \gamma \\
  k \cdot A \cdot \sin(\omega \cdot t) & \gamma < \omega \cdot t \leq \pi - \gamma \\
  k \cdot A \cdot \sin(\omega \cdot t) & \pi - \gamma < \omega \cdot t \leq \pi 
\end{cases}
\]

where for the angle \( \gamma \) we obtain
\[ \sin \gamma = \frac{a}{A} \]

Due to the oddity of the function \( w = f(x) \), it is \( a_i = 0 \). Determine \( b_1 \).

\[ b_1 = \frac{1}{\pi} \int_{-\pi}^{\pi} [A \cdot \sin(\omega \cdot t)] \cdot [\sin(\omega \cdot t) \cdot d(\omega \cdot t) =
\]
\[ = \frac{2 \cdot k \cdot A}{\pi} \left[ \gamma - \frac{a}{A} \sqrt{1 - \left(\frac{a}{A}\right)^2} \right] \]

So describing function is
\[ N(A, \omega) = \frac{1}{A} \cdot (b_i + i \cdot a_i) = b_i = \frac{2 \cdot k}{\pi} \left[ \arcsin \left( \frac{a}{A} \right) - \frac{a}{A} \sqrt{1 - \left(\frac{a}{A}\right)^2} \right] \]

Dividing by \( k \) we get the so-called normalized description function \( N(A) / k \)
\[ \frac{N(A)}{k} = \frac{2}{\pi} \left[ \arcsin \left( \frac{a}{A} \right) - \frac{a}{A} \sqrt{1 - \left(\frac{a}{A}\right)^2} \right] \]

In Fig. 6 its shape is plotted according to the ratio \( \frac{A}{a} \).

It is possible to see the three properties of this describing function:

1. If the input amplitude is in the linear region, then \( N(A) = k \).
2. As the input amplitude increases, \( N(A) \) decreases.
3. There is no phase shift.

The first feature is evident from the above mentioned. When signal is low, saturation does not occur. The second is also intuitively obvious. Saturating reduces the ratio of output to input. The third property is also understandable. Said symmetric saturation does not cause phase shift in the output.

As a special case of saturation we can consider the saturation with \( k = \infty \) i.e. non-linearity on-off-comparator. Its dependence \( w = f(x) \) is in Fig. 7.

Thus, this case corresponds to the limiting case of a linear saturation function \( a \rightarrow 0; k \rightarrow \infty \), which does not exclude that \( a \cdot k = M \). Although \( b_1 \) can be obtained from the (6) by limit, it is easier to calculate it directly.
\[ b_1 = \frac{1}{\pi} \int_{-\pi}^{\pi} f[A \cdot \sin(\omega \cdot t)] \cdot \sin(\omega \cdot t) \cdot d(\omega \cdot t) = \]
\[ = \frac{1}{\pi} \left\{ - \int_{0}^{\pi} M \cdot \sin(\omega \cdot t) \cdot d(\omega \cdot t) + \int_{-\pi}^{0} M \cdot \sin(\omega \cdot t) \cdot d(\omega \cdot t) \right\} = \]
\[ = \frac{1}{\pi} \left\{ - M \left[ - \cos(\omega \cdot t) \right]_{0}^{\pi} + M \left[ - \cos(\omega \cdot t) \right]_{-\pi}^{0} \right\} = \]
\[ = \frac{1}{\pi} \left\{ - M \left[ -1 - (-1, -1) \right] + M \left[ 1 - (-1, -1) \right] \right\} = \]
\[ = \frac{4}{\pi} \cdot M \]

So, describing function is
\[ N(A, \omega) = N(A) = \frac{1}{A} \cdot (b_1 + i \cdot a_1) = \]
\[ = \frac{b_1}{A} = \frac{4}{\pi} \cdot M \]

Normalized describing function
\[ N(A) = \frac{4}{\pi} \cdot \frac{1}{A} \cdot f(A) \text{ is on Fig.8.} \]

Fig. 8 Normalized describing function of on-off non-linearity.

4.2 Describing function of backlash and hysteresis non-linearity

In systems with mechanical force (torque) transmission, the backlash often occurs. This is due to the small airspace in the transmission mechanism. In a transmission chain consisting of, for example, a gearbox with front wheels (with parallel shafts), there is always some airspace between a pair of adjacent wheels. It is not just a consequence of inaccuracies in production and assembly. This is a prerequisite for reasonable transfer efficiency.

Fig. 9 shows a typical situation. The result of the dead zone in torque transfer between the teeth is that when the driving wheel is rotated, there is a certain path on the contact wheel circle, within which the torque is not transmitted to the driven wheel. Let the angle of rotation of the output wheel corresponding to this path be denoted as \( b \). Let the gear ratio be
\[ k = \frac{r_2}{r_1} \]

Thus, the torque is transmitted only when
\[ |r_1 \cdot \phi_1(t) - r_2 \cdot \phi_2(t)| \geq r_2 \cdot b \Rightarrow \]
\[ \Rightarrow r_2 \cdot \left| \frac{r_1}{r_2} \cdot \phi_1(t) - \phi_2(t) \right| \geq r_2 \cdot \frac{1}{k} \cdot \left| \phi_1(t) - \phi_2(t) \right| \geq 0 \Rightarrow \]
\[ \Rightarrow \left| \frac{1}{k} \cdot \phi_1(t) - \phi_2(t) \right| \geq 0 \Rightarrow \]
\[ \Rightarrow \left| \frac{1}{k} \cdot \phi_1(t) \geq \phi_2(t) \text{ then } \frac{1}{k} \cdot \phi_1(t) - \phi_2(t) \right| \geq b \]
\[ 2) \left| \frac{1}{k} \cdot \phi_1(t) - \phi_2(t) \right| < 0 \Rightarrow \]
\[ \frac{1}{k} \cdot \phi_1(t) < \phi_2(t) \text{ then } \left| \frac{1}{k} \cdot \phi_1(t) - \phi_2(t) \right| \leq -b \]

Thus, while the angular rotation between the recalculated driving wheel at the output and the driven wheel with different size (gear ratio is \( k \)) is less than \( b \), the torque between the wheels is not transmitted.

Fig. 10 Torque transfer dynamic model with backlash between teeth

Based on the dynamic model shown in Fig. 10 and the response behaviour of this non-linearity to the sinus input \( x(t) = A \cdot \sin(\omega \cdot t) \) (for \( A > k \cdot b \) ) shown in Fig. 11 and Fig. 12, can be written
Fig. 12 Input-output context for backlash type non-linearity in gearing

\[
\begin{align*}
\frac{1}{k} & \cdot A - b & -\pi \leq \omega \cdot t \leq -\pi + \gamma \\
\frac{1}{k} & \cdot \sin(\omega \cdot t) + b & -\pi + \gamma < \omega \cdot t \leq -\frac{\pi}{2} \\
\frac{1}{k} & \cdot \sin(\omega \cdot t) - b & -\frac{\pi}{2} < \omega \cdot t \leq \gamma \\
\frac{1}{k} & \cdot A - b & \gamma < \omega \cdot t \leq \frac{\pi}{2} \\
\frac{1}{k} & \cdot A - b & \frac{\pi}{2} < \omega \cdot t \leq \pi
\end{align*}
\]

For \( A \leq k \cdot b \Rightarrow \frac{k \cdot b}{A} \geq 1 \) the output is zero. We determine the angle \( \gamma \), i.e. the angle at which the backlash influence ends and the output wheel begins to move: Is valid

\[
\frac{1}{k} \cdot A \cdot \sin \gamma - \left[ -\left( \frac{1}{k} \cdot A - b \right) \right] = b \Rightarrow \\
\frac{1}{k} \cdot A \cdot \sin \gamma + \frac{1}{k} \cdot A - b = b \Rightarrow \sin \gamma = \frac{2 \cdot k \cdot b}{A} - 1
\]

Now, unfortunately, \( w(t) \) is not odd. Thus neither \( a_1 \) nor \( b_1 \) is zero and we have to determine them. The calculation is lengthy.

So

\[
|N(A, \omega)| = \frac{1}{A} \cdot \left| b_1 + i \cdot a_1 \right| = \frac{1}{A} \cdot \sqrt{a_1^2 + b_1^2}
\]

\[
\angle N(A, \omega) = \arctg \left( \frac{a_1}{b_1} \right)
\]

The describing function amplitude for backlash is shown in Fig. 13 and its phase is in Fig. 14.

Here they are some interesting facts:

1. \( |N(A, \omega)| \rightarrow \frac{1}{k} \) for \( b \rightarrow 0 \)
2. \( |N(A, \omega)| \) grows when \( \frac{k \cdot b}{A} \) decreases
3. \( |N(A, \omega)| \rightarrow 0 \) for \( b \rightarrow \frac{A}{k} \)

Phase shift (from 0° to -90°) is due to the effect of a given non-linearity. It is the result of the time shift caused by the backlash \( b \) [rad] on the output side of
the gear. Higher $b$ leads naturally to greater phase shifting, which may cause a stability problem with the feedback control system.

5 Analysis of nonlinear systems by the describing functions utilisation

In the following section we show the prediction of the limit cycles of the selected nonlinear system type based on the nonlinear element representation by the describing function. The basic approach for this prediction is based on the application of an extended version of the known Nyquist criterion from linear control to an equivalent system.

5.1 Nyquist criterion and its extension

Let us consider the linear system from Fig. 15.

![Fig.15 Closed linear control circuit](image)

The characteristic equation of this system is

$$\delta(s) = 1 + H(s) \cdot G(s) = 0$$  \hspace{1cm} (10)

Let's recall that $\delta(s)$ - sometimes called as an antiparallel circuit transmission function - it is the rational function of the complex parameter $s$ with the roots of its numerator forming the poles of the closed system. The roots of its denominator are the poles of the open regulatory system with transmission $G_0(s) = H(s) \cdot G(s)$.

The Nyquist criterion serves to determine the stability of the closed loop. As is well known, the closed loop will be stable if all the poles of the closed circuit transmission are in the left half of the Gauss plane. Nyquist, however, has shown that closed loop stability can be determined based on the frequency response of the open loop and the position of its poles. This is advantageous, since open loop transmission is, as opposed to closed loop transmission, mostly available. It is neither necessary to know the analytical shape, just experimentally detected data. It is also possible to use it for systems with transport delays where algebraic criteria fail.

5.1.1 Cauchy phase theorem

Consider a closed, negative-oriented curve $\Gamma_S$ (a clockwise curve) in a plane $s$ that does not pass through any root (zero or pole) of the $G(s)$ transmission. If we gradually set the points from this curve in the selected direction to $G(s)$, we will obtain another oriented, closed curve $\Gamma_G$ in the $G$ plane.

This process is called **mapping** of the closed curve $\Gamma_S$ from plane $s$ to the plane $G$. E.g. picture Fig. 16 shows the curve for the mapping process in case the mapping function (transmission) is equal $G(s) = \frac{s + 0.5}{(s + 2) \cdot (s + 3)}$.

![Fig.16 The curve $\Gamma_S$](image)

Its parametric equations are

$$\begin{align*}
\text{Re}(s) &= -2.5 + 1.5 \cdot \cos \varphi \\
\text{Im}(s) &= 1.5 \cdot \sin \varphi
\end{align*} \hspace{1cm} \varphi \in (0; -2\pi)$$

Thus, the curve envelops the two poles of the given transmission -2, -3, and does not cover its zero - 0.5. After any point substitution of this curve into the $G(s)$ transmission, we get the curves in Fig. 17.

![Fig.17 The curve $\Gamma_S$ and $\Gamma_G$ for transmission $G(s) = \frac{s + 0.5}{(s + 2) \cdot (s + 3)}$](image)

There is a relationship between the number of nulls and poles inside the closed curve $\Gamma_S$ and the change of the phase of the curve $\Gamma_G$, in other words, how much turns the vector begins at the beginning and passing successively the points on the curve $\Gamma_G$ in the direction obtained by mapping the curve $\Gamma_S$.

This relationship tells **Cauchy's theorem** about the phase:

If the closed, negative-oriented curve in the plane $s$ encompass $n_b$ zeros and $n_A$ poles of the transmis-
sion $G(s)$ and doesn’t pass through its pole or zero, then the closed curve formed by mapping the curve to the plane $G$ by the functions $G(s)$, circulate the origin of this plane $n_B - n_A$ in the negative direction.

For the case in Fig. 17, there are two poles in the closed curve and no zero. Thus $n_B = 0$ and $n_A = 2$. The curve also does not go through zero or pole. According to the previous statement, the number of turns should be in the negative directions, which are two circles in the positive direction.

### 5.1.2 Nyquist stability criterion

Create a negative-oriented curve $\Gamma_s$ that encircles the entire right half of plane $s$. It consists of an imaginary axis and a semicircle with an infinite radius $r \to \infty$ across the right half (see Fig.18).

![Fig.18 The Nyquist curve](image)

This shape is advantageous since the points on the imaginary axis correspond to the frequency characteristic $G(i\omega)$ after the mapping and the curve through the infinity corresponds to the point at the beginning of the plane $G$. The closed curve $\Gamma_G$ is therefore the frequency characteristic for $\omega \in (-\infty: \infty)$.

In the Nyquist criterion, therefore, we map to the $G_0$ plane. Since the open circuit transmission $G_0$ is usually given as the product of the transfer of the base type cells, the construction of the frequency characteristic is not a major problem.

### The Nyquist criterion in the final form

The Nyquist criterion in the final form is

**Closed feedback circuit** is stable if the frequency response of the open circuit in the complex plane circulates point (-1,0) in a positive direction for frequency changes from $-\infty$ to $\infty$ as many times as the number of open loop transmission poles lie in the right half of the plane $s$.

### Example

The open circuit is created by a two time-invariant inertial cells in series with a proportional gain $K_p$ regulator. Open loop transfer is

$$G_0(s) = \frac{K_p}{(1+\tau_1 s)(1+\tau_2 s)} \Rightarrow G_0(i\omega) = \frac{K_p}{(1+i\tau_1 \omega)(1+i\tau_2 \omega)} =$$

$$= -\frac{\tau_1 \cdot \omega + 1}{\tau_1 \cdot \omega^2 + 1} - \frac{\tau_2 \cdot \omega + 1}{\tau_2 \cdot \omega^2 + 1} + \tau_1 \cdot \omega^2 + 1 + \tau_2 \cdot \omega^2 + 1$$

![Fig.19 The frequency characteristic for $G_0(i\omega)$](image)

The frequency characteristic of this system for specific values $K_p, \tau_1, \tau_2$ is in Fig. 19. How does the change in gain $K_p$ of the proportional controller occur?

![Fig.20 The frequency characteristic for a) $K_p = 5$ and b) $K_p = 100$](image)

In Fig. 20 a), b) the critical part of the frequency characteristic for the two gains $K_p$ is given. Changing the gain occurs as a scale change on both real and imaginary axes. Thus, it is clear that by any gain $K_p > 0$ it is not possible to shift the frequency response $G_0(i\omega)$ so that it circulates the point (-1; 0). Therefore, the number of turns of the point (-1; 0) is equal to 0 for all possible gains. Since the transmission $G_0(i\omega)$ does not contain any unstable pole, we
can say on the basis of the Nyquist criterion that the closed circuit will be stable for $\forall K_p > 0$.

**Example.** The regulated system has a transmission

$$G_S(s) = \frac{1}{s^2 - 4 \cdot s + 1} = \frac{1}{[1 - (2 + \sqrt{3})][1 - (2 - \sqrt{3})]}$$

So it has two poles $s_1 = 2 + \sqrt{3}; s_2 = 2 - \sqrt{3}$ in the right part of the Gaussian plane. The controller is a real PD controller (with 1st order filtering)

$$G_R(s) = \frac{K_D \cdot s + 1}{0.1 \cdot s + 1}$$

Open loop transfer is

$$G_0(s) = \frac{K_D \cdot s + 1}{0.1 \cdot s + 1} \Rightarrow G_0(j\omega) = \frac{K_D \cdot \omega + j}{0.1 \cdot \omega + j}$$

For a sufficiently large derivative gain, the intersection with the negative real axis is more negative than point $-1$, as shown in Fig. 21.

Fig.21 Frequency response for $K_D = 5$

The frequency characteristic will make two circles around the point (-1.0) in the positive direction. This case is stable, since according to the Nyquist criterion, the number of turns in the positive direction coincides with the number of open loop unstable poles. If the gain $K_D$ is greater than any limit value $K_{D_{min}}$, then the closed circuit will be stable (Fig.22).

Fig.22 Frequency response for $K_D = K_{D_{min}} = 4.24$.

For small gain $K_D$, the intersection with the negative real axis is to the right of the point (-1.0). The number of rounds around the point (-1.0) is in this case 0. The closed circuit is unstable (Fig. 23).

Fig.23 Frequency response for $K_D = 3$.

Summarize the findings so far:

1. Create a negative-oriented curve - the so-called **Nyquist curve** that encircles the entire right half of plane $s$. It consists of an imaginary axis and a half-circle with an infinite radius over the right half-plane (see figure Fig.18).
2. Map this closed curve from plane $s$ to plane $G$ by transfer of an open loop. I.e. create the frequency response of the open loop $G_0(s) = H(s) \cdot G(s)$.
3. Determine the number of turns of this curve $N$ around the point (-1, 0) of the Gaussian plane $G$.
4. Determine the number $n_B$, the number of zero of the characteristic equation $F(s) = \delta(s) = \frac{C(s)}{A(s)} = 0$ in the right part of the plane with help $N = n_B - n_A \Rightarrow n_B = N + n_A$, where $n_A$ the
number of poles in the right part of plane s is (unstable poles). Then the number \( n_B \) is the number of unstable poles of the closed-loop system.

5.2 Existence of limit cycles

Let us suppose that in the system of Fig. 24 there is a spontaneous oscillation with amplitude \( A \) and frequency \( \omega \). Then, the variables in the loop must meet the following relationships:

\[
\begin{align*}
\dot{x} &= -y \\
\dot{w} &= N(A, \omega) \cdot x \\
y &= G(\omega) \cdot w 
\end{align*}
\]

So we get

\[
y = G(\omega) \cdot w = G(\omega) \cdot N(A, \omega) \cdot x = G(\omega) \cdot N(A, \omega) \cdot (-y) \Rightarrow y = -G(\omega) \cdot N(A, \omega) \cdot y \Rightarrow [1 + G(\omega) \cdot N(A, \omega)] \cdot y = 0
\]

Because \( y \neq 0 \), it has to pay

\[
1 + G(\omega) \cdot N(A, \omega) = 0,
\]

which we can write as

\[
G(\omega) = -\frac{1}{N(A, \omega)} \tag{11}
\]

Therefore, the amplitude \( A \) and the frequency \( \omega \) of the system’s limit cycle must meet (11). If this equation has no solution, then the nonlinear system does not create a limit cycle.

The relation (11) represents two non-linear equations (one for the real part and the other for the imaginary part of the equation) of the two variables \( A \) and \( \omega \). There is usually a definite number of solutions.

In general, it is difficult to solve these equations by analytical methods, especially for higher order systems. Generally, a graphical approach based on drawing both sides of equation (11) in the Gaussian plane is used looking for intersections of two curves.

5.2.1 Frequency-independent describing function

Let’s first consider the simplest case when describing function \( N \) is only a function of the amplitude \( A \), i.e. \( N(A, \omega) = N(A) \) is valid.

This case includes all one-valued non-linearity (expressed as a function of one variable) and some non-linear non-linearity expressible as a function of two variables – e.g. backlash.

Equation (10) moves to form

\[
G(\omega) = -\frac{1}{N(\omega)} \tag{12}
\]

We can draw the frequency response \( G(\omega) \) (variable \( \omega \) and the negative function inverse to describing the function \( -\frac{1}{N(\omega)} \) (variable amplitude \( A \)) to the complex Gaussian plane.

Consider the following example:

Draw the frequency characteristic of the linear part and \( -\frac{1}{N(\omega)} \) in the Gaussian plane for the following circuit

\[
\begin{align*}
G_s(s) &= \frac{K}{(s+1) \cdot (s+2)} \Rightarrow G_s(i\omega) = \frac{K}{(i\omega+1) \cdot (i\omega+2)} = \\
&= \frac{K \cdot (2 - \omega^2)}{(\omega^2 + 1) \cdot (\omega^2 + 4)} - i \cdot \frac{3 \cdot K \cdot \omega}{(\omega^2 + 1) \cdot (\omega^2 + 4)}
\end{align*}
\]

Backlash and hysteresis describing function is (see (9))

\[
N(A, \omega) = N(\omega) = \frac{1}{A} \cdot \left( b_i + a_1 \right)
\]

The angle \( \gamma \) applies

\[
\sin \gamma = \frac{2 \cdot k \cdot b}{A} - 1
\]

That is true

\[
-\frac{1}{N(\omega)} = -\frac{b_i + a_1}{A} = \frac{2 \cdot A^2 + k \cdot b}{A^2 - 1}
\]

\[
\begin{align*}
&= \left[ 2 \cdot A^2 - \left( \frac{2}{A} \right)^2 \right] + 4 \cdot A \cdot \left( \frac{1 - 2 \cdot k \cdot b}{A} \right) \cdot \left( 1 - \frac{k \cdot b}{A} \right) - i \cdot 8 \cdot k \cdot b \cdot A \cdot \left( 1 - \frac{k \cdot b}{A} \right) \\
&= \left[ A \cdot \left( \frac{2}{A} \right)^2 + 2 \cdot \left( 1 - \frac{k \cdot b}{A} \right) \right] \cdot \left( 1 - \frac{k \cdot b}{A} \right) - i \cdot 4 \cdot k \cdot b \cdot A \cdot \left( 1 - \frac{k \cdot b}{A} \right) \\
&= \left[ 4 \cdot A \cdot k \cdot b \cdot \pi \cdot \left( 1 - \frac{k \cdot b}{A} \right) \right] - i \cdot 4 \cdot k \cdot b \cdot A \cdot \left( 1 - \frac{k \cdot b}{A} \right)
\end{align*}
\]
Fig. 26 shows both characteristics in Gauss plane.

If the curves intersect, then there is a limit cycle with the values $A$ and $\omega$ corresponding to the intersections that are the solution of equation (11). If the curves intersection is $n$ times repeated (in example 2 times), then the system has $n$ possible limit cycles. Which limit cycle will be achieved depends on the initial conditions.

In Fig. 26 both curves intersect at two points. The first one answers to the $A_1 = 0.95$, $\omega_1 = 1.95$, the other one to the $A_2 = 2.525$, $\omega_2 = 6.82$.

Note that for nonlinearities with $a_1 = 0$ (one-valued non-linearity - expressed as odd function of one variable) the describing function is with $\text{Im} \{N(A,\omega)\} = 0$, it is the real function, i.e. $-\frac{1}{N(A)}$ lies on the real axis.

It is also useful to point out that the above procedure gives only the prediction of the existence of limit cycles. The validity and accuracy of this prediction should be confirmed by simulation.

5.2.2 Frequency–dependent describing function

In general, the describing functions are dependent on both amplitude and frequency, i.e. $N = N(A,\omega)$.

The described method can be used here, but with greater generality. The right side of (11) corresponds to the whole set of curves in the complex plane with the variable amplitude $A$ and $\omega$ as fixed for each curve - see Fig. 27. There is generally an infinite number of intersections between the curves $G(\omega)$ and $\frac{1}{N(A,\omega)}$.

Only the intersections in the corresponding $\omega$ indicate limit cycles.

To avoid the complexity of finding identical frequencies at the intersections, it may be advantageous to consider the graphical solution (10) directly, based on the plot $G(\omega) \cdot N(A,\omega)$.

For a fixed $A$ and a change $\omega$ from 0 to $\infty$, we obtain the curves representing $G(\omega) \cdot N(A,\omega)$. Different values $A$ correspond to the set of curves as in Fig. 28. A curve passing through a point $(-1,0)$ on the Gaussian plane indicates the existence of a limit cycle with the curve $A$ as the amplitude and frequency $\omega$ corresponding to this point is the frequency of the given limit cycle.

Although this procedure is much easier than in the previous case, it requires repeated calculations $G(\omega) \cdot N(A,\omega)$ to generate a set of curves. Therefore, there is a clear advantage of computer tools using.
5.3 Limit cycles stability

As we know, the limit cycles can be stable or unstable. In the previous we discussed how to detect their existence. Let us now discuss how to determine the stability of the limit cycle by extending the Nyquist criterion detailed in 5.1.

Let us consider the frequency characteristics and characteristics \(-\frac{1}{N(A)}\) of Fig. 26. There are two intersections, predicting that the system has two limit cycles. The amplitude value \(A_1\) of the first intersection \(L_1\) is smaller than the amplitude value \(A_2\) of the second intersection \(L_2\). For simple discussion, let's assume in general that linear transmission function \(G(s)\) does not have unstable poles, which is true in our example.

Let's first discuss the stability of the limit cycle at the first intersection. Let us assume that the system first operates at the \(L_1\) point with limiting amplitude \(A_1\) and its frequency will be \(\omega\). As a result of a small disturbance, the amplitude of the input to the non-linear member slightly increases and the working point moves from point \(L_1\) towards point \(L_2\). Because the new point is surrounded by a curve \(G(i\omega)\), according to the extended Nyquist criterion in 6.1, the system at this operating point is unstable and the system signal amplitudes will increase. Therefore, the operating point will continue to move along the curve \(-\frac{1}{N(A)}\) towards the second point of the \(L_2\) limit cycle.

If, on the other hand, a system disturbance causes a reduction in the amplitude, i.e., the movement on the red curve from \(L_1\) in the direction from \(L_2\), then the amplitude of the oscillations \(A\) will be further reduced as the new working point is not surrounded by a curve, and according to the extended Nyquist criterion the system at this operating point is stable.

Thus, a small disturbance changes the \(L_1\) working point, and this point is unstable. A similar analysis can be performed to evaluate the stability of the limit cycle at point \(L_2\). We will find that this limit cycle is stable.

Summarizing the discussion for this example and using the results of the previous section, we can formulate a criterion for the existence and stability of the limiting cycles:

**Limit Cycle Criterion:** Each intersection of the frequency characteristics of the system linear part \(G(i\omega)\) without the poles in the right-hand side of the Gaussian plane and the curve \(-\frac{1}{N(A,\omega)}\) where \(N(A,\omega)\) is the non-linearity describing function corresponds to the limit cycle.

If the points near the intersection and along the curve \(-\frac{1}{N(A,\omega)}\) with the increasing amplitude \(A\) are not surrounded by a curve \(G(i\omega)\), then the corresponding limit cycle is stable. If the points near the intersection and along the curve \(-\frac{1}{N(A,\omega)}\) with the increasing amplitude \(A\) are surrounded by a curve \(-\frac{1}{N(A,\omega)}\), then the corresponding limit cycle is unstable.

**Example.** Consider a simple one-dimensional satellite model at constant height and having one degree of freedom - angle \(\phi\), indicating the angular position of the antenna with an inertia moment \(J = 1\text{kgm}^2\) controllable by double rockets generating with constant positive or negative mechanical torque \(Q\).

![Fig.29 Simplified control of the satellite angular position](image)

**Fig.29** Simplified control of the satellite angular position

**Fig. 30** shows the block structure of the angle control.

![Fig.30 Block diagram of simplified satellite angular position control](image)

**Fig.30** Block diagram of simplified satellite angular position control

For the frequency characteristic of the linear part is valid

\[
G_{lin}(s) = \frac{1}{s^2} \Rightarrow G_{lin}(i\omega) = -\frac{1}{\omega^2}
\]

Comparator describing function is (8)
\[ N(A, \omega) = N(A) = \frac{1}{A} \cdot (b_i + i \cdot a_i) = \frac{b_i}{A} = \frac{4}{\pi} \cdot \frac{M}{A} \]

That is true

\[ -\frac{1}{N(A)} = -A \cdot \frac{1}{b_i} = -\frac{\pi \cdot A}{4 \cdot M} \]

For nonlinearities with \( a_i = 0 \) (one-valued nonlinearity - expressed as odd function of one variable) the describing function is with \( \text{Im}[N(A, \omega)] = 0 \), it is the real function, i.e. \(-\frac{1}{N(A)}\) lies on the real axis.

In our case also the frequency characteristic of the linear part lies on the real axis! Both curves merged. Thus, there are infinitely many intersections of both curves, but individual points correspond to different \( A \) and \( \omega \). E.g. the intersection with the value on the real axis \(-0.438\) corresponds to the parameters of both lines \( A = 0.558, \omega = 0.477 \). Therefore, the estimation is that there are infinitely many limit cycles corresponding to different amplitudes and frequencies.

Can we decide whether or not the intersection points are surrounded by a curve \( G(\omega) \) near the intersections and along a curve \( \frac{1}{N(A, \omega)} \) with increasing amplitude \( A \)? These are two merging lines! We know the system is on the line of critical stability!

Let us consider the variant of the example with added some linear friction with the coefficient \( b \). Thus the new block diagram is

\[ \text{Fig.31 Block diagram of simplified satellite with linear friction angular position control} \]

Now for the frequency characteristic of the linear part is valid

\[ G_{\text{lin}}(s) = \frac{1}{s \cdot (s + b)} \Rightarrow G_{\text{lin}}(\omega) = \frac{1}{i \omega \cdot (\omega + b)} = \]

\[ = -\frac{i \cdot \frac{1}{\omega}}{\frac{1}{\omega} \cdot (\omega + b)} = -\frac{i}{\frac{1}{\omega} \cdot (\omega^2 + b^2)} = -i \cdot \frac{1}{\omega} \cdot (\omega^2 + b^2) \]

Comparator describing function is again

\[ N(A, \omega) = N(A) = \frac{1}{A} \cdot (b_i + i \cdot a_i) = \frac{b_i}{A} = \frac{4}{\pi} \cdot \frac{M}{A} \]

That is again true

\[ -\frac{1}{N(A)} = -A \cdot \frac{1}{b_i} = -\frac{\pi \cdot A}{4 \cdot M} \]

\[ \text{Fig.32 Limit cycle detection for the situation in Fig.30} \]

On Fig. 32 is part of the frequency characteristic at change \( \omega \) from 0 to \( \infty \) and the line indicating change of \( A \) from 0 to 1. We see that the only point of intersection is the point [0, 0] of Gaussian plane. Thus, the ”limit cycle” will have amplitude of 0 and an infinite frequency. It is a point. The curve does not enclose it when \( A \) rises, the point will be stable.

\[ \text{Fig.33 The Physical multiport diagram of the situation from Fig.31.} \]

Fig. 33 shows the physical multiport diagram of the rotating satellite with linear friction and the angular variance control to zero by means of a comparator.

\[ \text{Fig.34 Time dependencies under non-zero initial conditions.} \]

Fig. 34 shows the time dependencies of the system response to non-zero initial conditions. They can see confirmation of analysis made by frequency analysis (modified Nyquist criterion). Under any initial conditions different from zero, there is an oscillating
transient phenomenon that ends at a stable "cycle" point.
It is better to see this in Fig. 35, where the phase portrait is presented.

![Phase portrait of system response under non-zero initial conditions](image)

Fig.35 Phase portrait of system response under non-zero initial conditions

### 6 Conclusions

The general mathematical description of the mechatronic systems dynamic behavior as artificial systems with purposeful motion control, in which one part is a subsystem with the motion of interconnected bodies with non-zero resting mass, necessarily leads to a nonlinear system.

The primary cause of its nonlinearity is the existence of the Coriolis type forces (forces dependent on the product of the bonded bodies’ motion speeds). But even if in the case of slow movements these elements of the dynamic description are neglected in the design of control laws (we consider these forces as disturbances), in the real systems remain the effects of the so-called hard nonlinearities that are part of both mechanical subsystems (friction, backlash, hysteresis) and the control system (saturation, hysteresis).

These nonlinearities can cause both desirable and undesired phenomena where their most significant manifestation is the existence of limit cycles.

This article follows up on the previous paper [11], where it’s describe how we can obtain a describing function for a non-linear system containing one such non-linear element. This will allow us to further analyse the existence of limit cycles based on the representation of the non-linear element by describing function.

The basic approach for this prediction is based on the application of the extended version of the criteria based on Cauchy's lemma from complex analysis (Nyquist criterion known from the linear control theory) to the equivalent system obtained by a describing function application.

The paper is second part of an analysis of non-linear systems, where in this part is the application of describing functions—the so-called **frequency linearization**—to the existence and basic parameters of the limit cycles analysing.

### References:


