Cancellation of Non-Linear Effects using Inverse Describing Functions

George Page, Dingli Yu and Barry Gomm
Mechanical Engineering and Materials Research Centre
Liverpool John Moores University

26th June 2017

Abstract

Many of the features of the behaviour of non-linear systems can be investigated by superimposing describing functions onto inverse Nyquist diagrams. For instance, the existence and properties of limit-cycles can be predicted in this way. It naturally follows that if other non-linear systems could be created whose describing functions could cancel out the effects of the original functions then an effective method of controlling non-linear behaviour would have been developed. In this paper a technique for creating these 'inverse' describing functions is presented and it is shown that when such functions are placed in series with the original non-linearities they can completely cancel out their effects.

1 Introduction

When a system possesses both linear and non-linear features the classical approach has been to treat the two areas quite separately, as shown in Fig.1. Usually the characteristics of the non-linearity or non-linearities, have been carefully determined so that the input parameters of the system can be chosen so that the plant or process operates entirely outside the region in which they appear. The arsenal of control procedures are then brought into play on the linear part of the system. The non-linearity has been effectively isolated and the system sanitized so that it, hopefully, operates like a well-behaved purely linear system. Unfortunately such sanitization does not always work. Sometimes a linearization procedure has to be undertaken [1] and/ or several control rules determined for strictly specified linear ranges.

In all of the above, no attempt has been made to control or modify the behaviour of the non-linearity itself. In this paper it is shown that by using describing function techniques it is possible to modify the characteristics of the non-linearity so that the undesirable features are either reduced in effect or else moved to a less sensitive operating region.

Certain types of non-linear system [2] can exhibit a phenomenon known as limit-cycle or spontaneous oscillation and this effect has been of particular interest as far as this research is concerned. There is a related effect known as jump-resonance but this is not as serious a problem and is more easily controlled. Limit-cycling is quite distinct from ordinary resonance and is independent of

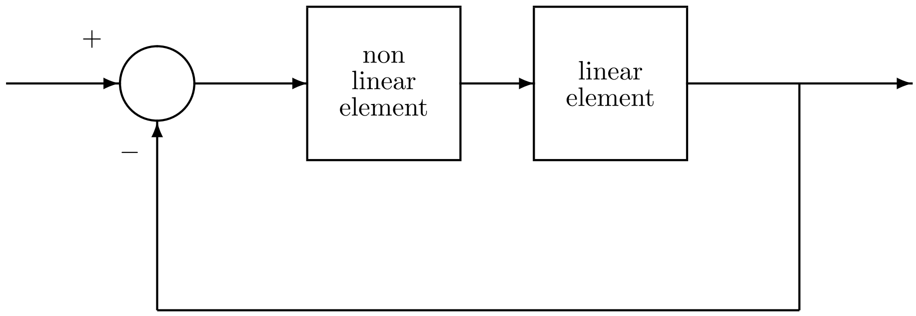

Figure 1: System in which the linear and non-linear sections are treated separately

the frequency content of any input signal. It is a property of the architecture of the nonlinear system itself and occurs spontaneously when the amplitude of the signal enters a certain critical range. This critical range may exist at low amplitudes but not at high, at high amplitudes but not at low; or it may even exist for one or more intermediate ranges of amplitudes. When a limit-cycle occurs, the affected system goes into a stable sinusoidal, or even non-sinusoidal, oscillation with an amplitude and frequency which are characteristics of the architecture of the individual system itself and are in no way connected to the frequency and amplitudes of any inputs.

The fact that the linear and non-linear elements could be treated separately gradually led to the development of the describing function approach [3] and when this was combined with classical control methods, in particular Nyquist diagrams [3] [4] [5], it became possible to predict when limit cycles might occur and also to predict their amplitudes and frequency of oscillation. The method is not foolproof [6] but has proved sufficiently reliable to enable its use in our non-linear control paradigms. Once a limit-cycle starts, control of the affected system is lost. Changing the amplitude of the input signal will often have no effect. ‘Even removing all input signals will usually have no effect. Once it has started, a limit cycle is self-sustaining and usually the only way to stop it is to remove all power and shut down the system.

Up to now, the only way in which the unwanted oscillation could be avoided was to make sure that the amplitude of the input or feedback systems reside outside certain critical ranges or, more drastically, to redesign the system itself. Having to use signals within limited ranges of amplitude can seriously curtail the performance of a system but sometimes it has been a price that has had to be accepted. Re-designing the very architecture of a system has not usually been an option.

A new method is presented in this paper which avoids the necessity of re-designing the system itself by enabling the characteristics of the non-linearity to be adjusted to either remove the conditions in which the unwanted oscillations occur or to move them into regions of error signal amplitude in which they no longer present operational problems.

2 The Basic Cancellation Method

The basic method can be divided into three parts: (i) the derivation of the non-linear function that describes the undesirable non-linear effects which are to be removed, (ii) combining the Nyquist graph of the linear part of the system with the describing function in order to compare the predicted non-linear behaviour with that of the actual (to check that the derived describing function is correct) and (iii) the design of the inverse function which can then be inserted into the original system and so remove the unwanted non-linear effects.

2.1 The describing function

To simplify the approach only those non-linearites which produce real describing functions have been considered in this paper. Although this has excluded systems which possess memory the method is still applicable to them. However, as such non-linearities produce complex describing functions they require a three dimensional graphical approach which would complicate this initial presentation unnecessarily. To derive the describing functions an algorithm [7] was used, which is based on the standard frequency-domain graphical techniques [8]. This approach has been summarized in the appendix.

2.2 Predicting the behaviour of the non-linear system

If the linear part of the system is represented by the transfer function $G(j\omega)$, where ω represents the frequency of the signal, and the describing function of the non-linear part is represented by $N(X, \omega)$, in which X represents the amplitude of the signal then, for the system to remain stable, either the locus $G(j\omega)$ must keep the entire locus $-1/N(X, \omega)$ on the right or the inverse locus $1/G(j\omega)$ must keep the locus $-N(X, \omega)$ on the left (or must completely enclose the whole of the locus) [3]. For our work, the use of the inverse Nyquist diagram turned out to be far more convenient because it avoided the graphs going off to infinity at crucial positions.

If a non-linearity was described by a function which was entirely real, opposed to complex,then its locus would lie entirely along the real axis. If an inverse Nyquist diagram was superimposed on this graph then the two loci could only intersect at one frequency position along the Nyquist locus - where they crossed on the real axis. This means that a system which possessed a describing function which was real but so convoluted that it crossed the real axis several times could have multiple limit-cycles but they would all possess the same frequency. However, they would occur at different signal amplitudes. The traditional way of showing the superposition of the two loci, by plotting their amplitudes on real/imaginary axes does not show the multiple crossings of the loci. The alternative graph showing the signal amplitude against the Nyquist and describing function amplitudes, Fig.2, gives a much clearer picture of what is happening. On this diagram it can be seen that the loci crossed three times. Limit-cycles occurred at points A and C. Point B is a critical point, if the signal amplitude x exceeds this value then the system will immediately sweep around to the second limit-cycle even though the signal amplitude is lower than the apparently required value.

2.2.1 Interactions between limit-cycles

The switching between limit-cycles takes place in an entirely random fashion and cannot be predicted. It is an example of chaotic behaviour; in fact the phase-plane portrait of this interaction between two limit-cycles is an example of the Lorenz butterfly effect. Furthermore, it is perfectly possible to design a system which has multiple limit-cycles. Provided that these limit-cycles are sufficiently close to each other there will be spontaneous cascaded switching between them. The phase-plane portrait in this case will be a butterfly with as many wings as there are limit-cycles.

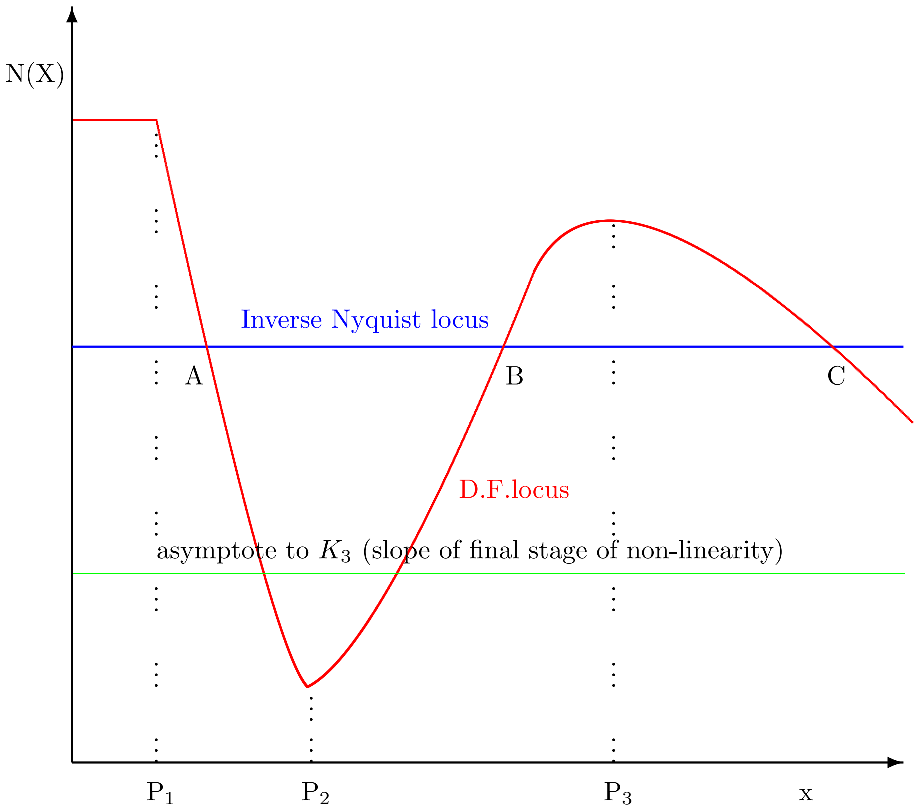

Figure 2: Superposition of real describing function and Inverse Nyquist loci

2.3 Can an inverse describing function be designed?

It has often been stated that it is impossible to develop a perfect inverse of a non-linear system,[9],[10],[11] because the principle of superposition did not apply. Whilst it is true that an analytical solution may not be easy to find, it is certainly possible to find a geometrical solution to any problem which can

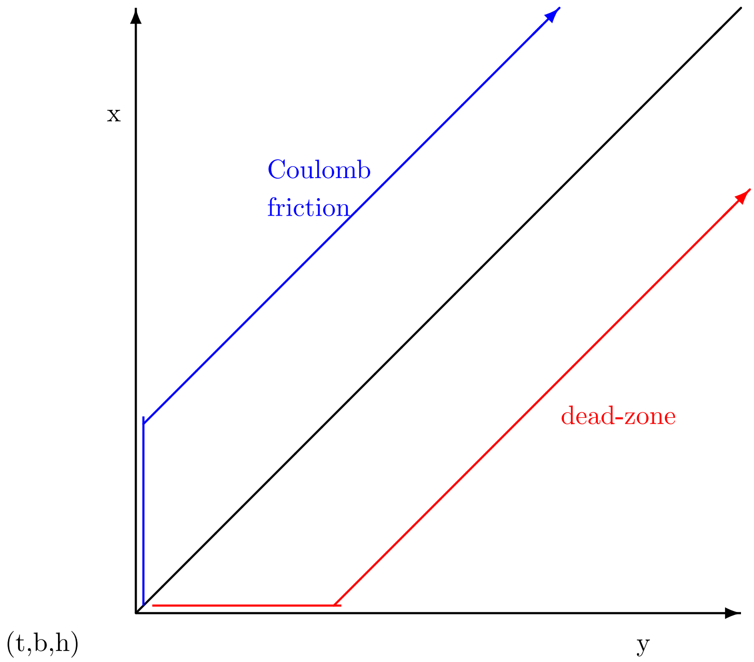

Figure 3: Dead-zone plus linearity with its mirror image, Coulomb friction, about the 45^o line

be posed graphically. In any case, often a complete inverse function did not have to be found all: that was necessary was that the shape of the native describing function should become modified sufficiently for it to no longer cross the inverse Nyquist locus. If a ramp input was applied to an open-loop system which contained an embedded non-linearity then the output signal would be representative of that non-linearity. If a perfect inverse of the non-linearity was then placed in series in the open-loop the the output should simply be the same as the original ramp input. All such inverse have to do is to nullify the instantaneous gain introduced by the original non-linearity at each point along its path. So such inverses are simply the mirror images of the original non-linearities about the unit ramp. One result which follows immediately from this realization is that Coulomb friction, or the ideal relay effect, is the inverse of dead-zone and vice-versa, as shown in Fig.3. Furthermore,since the inverse is simply the mirror image about the 45^0 line it follows that the inverse function is simply the original function with the x and y coordinates interchanged, Fig 4.

2.4 The design of an inverse describing function

In its simplest form a real non-linearity can be represented by a series of straight lines, with gradients K_0, K_1, K_2, K_3..., with K_0 being the straight line starting at, or passing through the origin. The junctions between the lines can be represented by breakpoints P_0, P_1, P_2, P_3, ... Breakpoint P_0 would lie at the origin and would only exist if Coulomb friction or the pure relay effect were present.

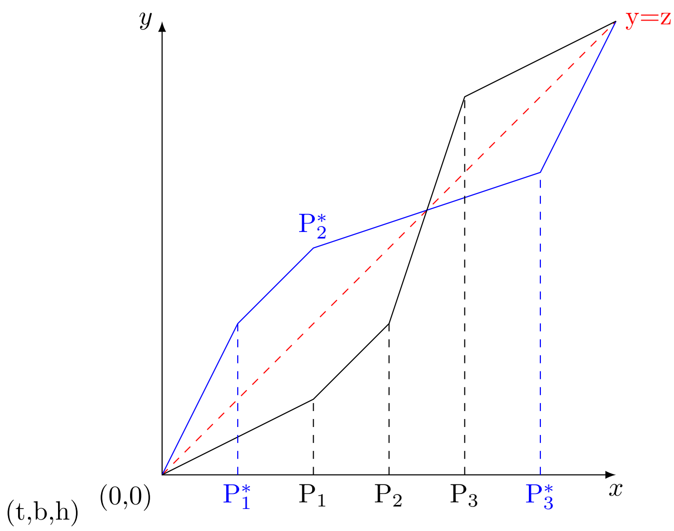

Figure 4: approximation to a non-linearity and its inverse: mirror-images about $y = z$

From the premise that the inverse of a non-linearity can be represented by its mirror image about the 45^0 line, the inverse can be drawn on the same axes as the original non-linearity, Fig.4. The corresponding slopes of the linear sections will be represented by $K_0^*, K_1^*, K_2^*, K_3^*, ...$ and the break-points will occur at $P_0^*, P_1^*, P_2^*, P_3^*, ...$ The crux of the problem resides in finding the values of K_i and P_i. If theses values can be determined then the inverse values follow automatically because, from simple geometry:

$$K_i^* = 1/K_i \tag{1}$$

$$P_0^* = P_0 \tag{2}$$

$$Q^* = Q \tag{3}$$

and

$$P_n^* = \sum_{j=1}^{n-1} (K_{j-1} - K_j)\, P_j + K_{n-1} P_n \tag{4}$$

2.5 Determining the parameters

If a unit-ramp input can be applied to the open-loop of the system containing the non-linearity then it is a simple matter to measure the pseudo-linear slopes and also the positions of intervening break-points. If the open-loop measurements are not available then information has to be obtained indirectly by means of the describing functions.

From Fig.2 certain basic features of a non-linearity are immediately obvious: the gain of the non-linearity near the origin can be measured, as can the position of the first break point P_1. Also, break-points must occur singly between limit cycle positions and critical points. If a break-point occurs between positions A

and B on the Inverse Nyquist diagram then if the linear gain is adjusted until A and B coincide the position of the break point will be found. Similarly, if a break-point occurs between B and C and the linear gain is adjusted in the opposite direction until B and C coincide the next break point can be located. In this way, for a system with multiple limit-cycles,and consequently multiple break-points, the general shape of the non-linearity can be determined. In the case where the value of the describing function locus is less than that of the inverse Nyquist value as the magnitude of the system value x approaches infinity then the position of the final break-point can be located by adjusting the linear gain and locating the point at which the magnitude of the final limit-cycle starts to reduce. This will give the position of the final break-point. The magnitude of the Inverse Nyquist at these positions will also give an approximate value for the adjacent lower pseudo-gain of the non-linearity [7] but is likely to be out by about 10%. If it is preceded by a limit-cycle position (e.g.: point A, Fig.2) then the reading is likely to be too high; if is is preceded by a critical position (e.g.: point B, Fig. 2) then the reading will be too low.

3 Discussion and Conclusions

Limitations of the method Although linear graphs have been used to demonstrate the method it has to be noted that this graphical approach is very much a simplification which is ONLY valid when dealing with REAL describing functions. The system only works when there is a one-to-one correspondence between the input and output stages of the non-linear effect. So graphs which have a negative slope at some point are not valid because they can lead to multiple correspondences between input and output. Furthermore if the system experiences saturation then all reconstruction systems must fail because once information about a system has been truly lost, as happens at saturation, then it is gone for ever and any attempt at reconstruction is only ever going to be guesswork. The full describing function approach when applied to systems which exhibit memory, which produces complex describing functions, cannot be demonstrated either by the simple graphical approach or by the simplified Nyquist/describing function interaction shown in figure 2. The full approach requires, at the very least, 3-dimensional representations and certainly higher-order dimensions when rotating non-linearities (as found, for example, in optical systems) are present. The case of complex, but not rotating, non-linearities is the subject of a subsequent paper.

References

[1] Dutton,K., Thompson, S. and Barraclough, B., *The Art of Control Engineering*, Prentice-Hall 1997

[2] Khalil, H.K.,*Nonlinear Systems 3rd ‘Ed.* Prentice-Hall, pp54-59, 280-288,2002

[3] Kochenburger, R.J., *A frequency response method for analysing and synthesising contactor servomechanisms*, Trans. AIEE, Vol. 69, pp 270-285, 1950

[4] Atherton, R.J.,*Stability of Non-linear Systems*, Research Studies Press, Wiley,1981

[5] Nagrath, I.J. and Gopal, M. *Control Systems Engineering, revised 5th Ed.*, Ansham, 2013

[6] Engelberg, S., *Limitations of the describing function for limit-cycle prediction*, IEEE Trans. Aut. Control, Vol. 47, Issue 11, pp 1887-1890, Nov. 2002.

[7] Page, G.F., Gomm, J.B. and Douglas, S. *An Algorithm for Generating Real Describing Functions*, UKACC, Cardiff, 2012

[8] Atherton, D.P.,*Nonlinear Control Engineering*, Van Nostrand, 1975

[9] Lee, H.J. Park, H.J. and Chen, G.*Robust Fuzzy Control of Non-linear Systems with Parametric Uncertainties*,IEEE Trans. Fuzzy Systems, Vol. 9, No. 2, pp 369-379, 2001

[10] Vashkov, G. and Fukuda, T., *Fuzzy Control Based on Self-Tuning Inverse Fuzzy Model*, Proc 6th European Congress on Intelligent Techniques and Soft Computing, Aachen Germany, pp 839-843, 1998

[11] Foo, T.J., *Stable Model Reference Adaptive Fuzzy Control of a Class of Nonlinear Systems*, IEEE Trans.Fuzzy Systems, Vol.9, No.4. pp 624-636, 2001

A Appendices

A.1 An algorithm for generating real describing functions

For ease of reference, the basic algorithm for rapidly calculating real describing functions is presented here. More complete details can be found in [7].

If Coulomb friction or relay action is present start at stage one, otherwise start at stage two.

Stage One:

(a) If Coulomb friction or relay action is present make $4Q/\pi X$ the first term of the describing function, where Q is the value of the Coulomb friction term.

(b) If dead-zone is also present multiply the above result by

$$\sqrt{\left(1 - \left(\frac{P}{X}\right)^2\right)}$$

where P is the dead-zone break-point.

Stage Two:

(a) If saturation is not present make K_n the first term of the describing function, where K_n is the gain of the last stage of the non-linearity. Or else add it to the result of stage one.

(b) If saturation is present then omit this term.

Stage Three:

(a) If there are n breakpoints then add n terms of the form

$$\frac{2}{\pi}\left(K_{i-1} - K_i\right)\left(\arcsin\left(\frac{P_i}{X}\right) + \left(\frac{P_i}{X}\right)\sqrt{\left(1 - \left(\frac{P_i}{X}\right)^2\right)}\right) \tag{5}$$

where

$$i = 0 \rightarrow n$$

Go to end

(b) If saturation is present then change the last of the terms in stage 3(a) to:

$$\frac{2}{\pi}K_{i-1}\left(\arcsin\left(\frac{P_{i-1}}{X}\right) + \left(\frac{P_{i-1}}{X}\right)\sqrt{\left(1 - \left(\frac{P_{i-1}}{X}\right)^2\right)}\right) \tag{6}$$

where

$$i = 0 \rightarrow n$$

Go to end

end

THE INVERSE ALGORITHM

If dead-zone is present then start at stage one, otherwise start at stage two

Stage One:

(a) If dead-zone is present make $4P_0^*/\pi X$ the first term of the describing function, where P_0^* is the value of the dead-zone term.

(b) If Coulomb friction or relay action is also present multiply the above result by

$$\sqrt{\left(1 - \left(\frac{Q^*}{X}\right)^2\right)}$$

where Q^* is the value of the Coulomb friction or the relay action.

Stage Two:

(a) If saturation is not present make K_n^* the first term of the describing function. (K_n^* is the gain of the last stage of the non-linearity) or add it to the result of stage one.

(b) If saturation is present then omit this term.

Stage Three:

(a) If there are n breakpoints then add n terms of the form

$$\frac{2}{\pi}\left(K_{i-1}^* - K_i^*\right)\left(\arcsin\left(\frac{P_i^*}{X}\right) + \left(\frac{P_i^*}{X}\right)\sqrt{\left(1-\left(\frac{P_i^*}{X}\right)^2\right)}\right) \tag{7}$$

where

$$i = 1 \to n$$

Go to end

(b) If saturation is present then change the last of the terms in stage 3(a) to:

$$\frac{2}{\pi}K_{i-1}^*\left(\arcsin\left(\frac{P_{i-1}^*}{X}\right) + \left(\frac{P_{i-1}^*}{X}\right)\sqrt{\left(1-\left(\frac{P_{i-1}^*}{X}\right)^2\right)}\right) \tag{8}$$

where

$$i = 1 \to n$$

Go to end
end
and it follows from the definitions of the terms and from Fig.4 that:

$$K_i^* = 1/K_i$$

$$P_0^* = P_0$$

$$Q^* = Q$$

and

$$P_n^* = \sum_{j=1}^{n-1}\left(K_{j-1} - K_j\right)P_j + K_{n-1}P_n$$

Finally, substituting these values into the preliminary iteration gives the general result for the inverse algorithm.