# Early Developments in Nonlinear Control

Derek P. Atherton

This article presents a review of some of the early developments in nonlinear control engineering. It begins by briefly examining the status prior to World War II and then concentrates on the developments in approximately the following two decades. A significant amount of the work will be seen to have as its foundation the need to obtain methods to solve wartime problems related to servomechanisms, fire control, and missile control. The major analytical techniques of the phase plane, the describing function and Tsypkin's method for relay systems, are discussed at length. Two important aspects in obtaining control engineering solutions to problems, which are often forgotten by theoreticians, are the existence of simulation facilities and the capabilities of the available hardware to implement specific solutions at an allowable cost. Later sections therefore comment on these aspects. Also, a brief review of some of the early papers on nonlinear control is given, and finally some comments are provided on the continuing relevance of these early methods and some still unsolved problems.

## Introduction

The purpose of the article is to review developments in nonlinear control engineering up to approximately 1960. This is a suitable end point for many reasons. First, many people might define the start of the "modern era" as the first IFAC Congress which was held in Moscow in 1960, and second, the first publication from what became our Society was the May 1956 issue of the IRE PGAC. It is also particularly convenient from my point of view, since I started my graduate studies in 1956. Perhaps most important, however, is the fact that significant work took place on nonlinear control engineering during the wartime period of the early 1940s. It took about a decade for these activities to become generally known to what at the time was a relatively small control engineering community compared with the situation today. Further, the page limit would not allow justice to be done to later work even if the author felt such a review was possible.

The article is organized as follows. The next section reviews the situation that existed prior to 1940 and during the war with respect to the problems that needed to be solved and the approaches used. The next three sections then cover the three main theoretical approaches used for studying nonlinear systems in the period 1940-1960, namely the phase plane, the describing function, and special methods for relay systems. Simulation has always played, and still does play, a major role in the study of nonlinear systems, but the hardware available during this period both for system simulation and implementation was far removed in many ways from today's technology. This situation is looked at in some detail under the heading "Hardware Technology." In the following section, some comments are given on the publications available around 1960 to give an idea of the size of the control community and its concern with nonlinear problems. I have called the final section "Reflections and Conclusions." There, some personal opinions are given on the current relevance of these early methods and the difficulties that still remain in dealing with nonlinear systems.

*The author is with School of Engineering, University of Sussex, Brighton, BN1 9QT U.K., telephone: +44 (0)1273 678046, fax: +44 (0)1273 678399, email: atherton@sussex.ac.uk.*

## The Early Years

### Prior to 1940

Although nonlinear control systems were being successfully applied, for example the Tirrill regulator and the flyball governor, they were made to work without any significant theoretical understanding. Control engineering, as a subject, was in its infancy in the late '30s, and some of the techniques that could be used for studying nonlinear systems had been, and were being, developed by mathematicians and physicists interested in nonlinear differential equations that occurred in modeling problems in such areas as celestial mechanics, mechanical vibrations, acoustics, and electronic oscillators. The major activity concentrated on second-order nonlinear differential equations of the form

$$\ddot{x} + \omega_0^2 x = mg(x, \dot{x}) \tag{1}$$

which were satisfactory models for the study of many of these physical problems. The main approaches used were perturbation techniques, harmonic balance, and the phase plane.

In the perturbation technique $\mu$ is a small parameter and a solution of the form

$$x(t) = \sum_{j=0}^{\infty} \mu^j x_j(t) \tag{2}$$

is sought. During this period there was considerable interest in finding limit cycle solutions for some of these differential equations, and difficulties were encountered in using the perturbation approach until the problem was solved by Lindstedt [1] and Poincaré [2] They replaced t by $\tau/\omega$, where $\omega$ is the unknown limit cycle frequency, and used the series expansions



0272-1708/96/$05.00©1996IEEE

$$x(\tau) = \sum_{j=0}^{\infty} \mu^j x_j(\tau) \tag{3}$$

and

$$\omega = \sum_{j=0}^{\infty} \mu^j \omega_j \tag{4}$$

where $\omega_0$ is the oscillation frequency with $\mu = 0$, to obtain solutions.

Averaging methods for obtaining solutions to Equation (1) were given by Van der Pol [3] and Krylov and Bogoliubov [4]. The formulation used by the latter authors, which was slightly different from that of Van der Pol, is the procedure that has been adopted more frequently in subsequent literature. In their method of slowly varying amplitude and phase, Krylov and Bogoliubov showed that an approximate solution to Equation (1) is

$$x = a(t)\sin[\omega_0 t + \phi(t)] \tag{5}$$

where $a(t)$ and $(t)$ are the solutions of the differential equations

$$\dot{a} = (\mu / 2\pi\omega_0)\int_0^{2\pi} g(a\sin\theta,\ a\omega_0 \cos\theta)\cos\theta d\theta \tag{6}$$

$$\dot{\phi} = (\mu / 2\pi a\omega_0)\int_0^{2\pi} g(a\sin\theta,\ a\omega_0 \cos\theta)\sin\theta d\theta \tag{7}$$

Various forms of harmonic balance technique were used by many authors, after the original work of Duffing (5), to study both free and forced oscillations. Both the perturbation and harmonic balance approaches are tedious algebraically, since they involve the collection of similar terms in what becomes a large equation for complex expressions for $g(x, \dot{x})$. They were therefore only applied to systems with simple nonlinear functions $g(x, \dot{x})$. The works of Duffing on nonlinear vibrations and Van der Pol on electronic oscillators resulted in specific nonlinear second-order equations being named after them. Using these methods, techniques were also found for predicting phenomena unique to nonlinear systems, such as limit cycles, jump phenomena, subharmonic oscillations, and synchronization or frequency entrainment.

The phase plane method was introduced by Poincaré [2] and is undoubtedly one of the most useful early methods which has proved of value for control systems design. Poincaré recognized that a solution for the second-order nonlinear differential equation:

$$\begin{aligned} \dot{x} &= P(x, y) \\ \dot{y} &= Q(x, y) \end{aligned} \tag{8}$$

could be sketched in a phase plane—a plot of $y$ against $x$—using the fact that the slope of the solution curve, or trajectory, at a point in the phase plane is:

$$\dot{y} / \dot{x} = dy / dx = Q(x, y) / P(x, y) \tag{9}$$

Many contributions [6] following Poincaré's work were made to topics in phase plane topology, such as information on singular points and the structure of trajectories near to them, conditions for the existence of limit cycles, and so on, in subsequent years. The phase plane approach became a valuable tool of the control engineer from the late '30s primarily because many control problems of interest in subsequent years were concerned with servomechanisms, which could often be approximated by second-order dynamics. Another major reason was the recognition that the method could be used effectively for the type of nonlinearities encountered in these systems, which could be approximated reasonably well by linear segmented characteristics, unlike the continuous functions previously considered.

Although a stability theory for linear differential equations had been established about the time of Poncaré's work, little seems to have been done on general nonlinear differential equations as distinct from second-order ones. Although Lyapunov's [7] original work was first published in 1892 in Russia, it appears to have been neglected for many years and certainly only became known to scientists outside the USSR toward the end of the 1950s.

The major effort at this time on investigations into the effects of nonlinearity in control systems was being undertaken at MIT [8]. Here Bush and his co-workers were developing differential analyzers, using mechanical integrators, for the study of nonlinear differential equations and required accurate servomechanisms for curve-following applications. Hazen [8] implemented various designs, including some using relays, and was aware of the limitations on the performance caused by backlash in the gears. This early link between nonlinear control and simulation is particularly interesting with respect to the continued importance of the latter in control engineering.

### Wartime Problems

The requirement for accurate fire-control systems led to significant work on servomechanisms in both the U.K. and the U.S. The problems associated with the major inherent nonlinear effects in these systems, such as friction, backlash, and saturation in amplifiers, were soon recognized, and techniques for analyzing their effects and innovative ideas for reducing them, such as the use of dither and anti-backlash gears, were developed.

In Germany, nonlinear control problems were also being encountered in the design of the controls for guided weapons, whose development at that time was far in advance of anything elsewhere. In the U.K. a servo-panel was formed in 1942 which functioned as might a committee of a learned society by organizing meetings for the general exchange of information on weapons control. Similarly, in the U.S. a National Defense Research Committee (NDRC) was set up in 1940, and one division was eventually established under Hazen's chairmanship devoted to the study of fire control problems [8]. The methods being used to study nonlinear effects were primarily the phase plane and later the describing function approach, with relay systems being studied by both methods. Minorsky [9] did, however, in a paper published in 1941, make a brief reference to nonlinear control problems and the possibility of using Lyapunov's method.

## The Phase Plane Method

As previously mentioned, the fact that many servomechanisms could be approximated by second-order dynamics led to the use of the phase plane for analyzing the effects of various nonlinearities, such as nonlinear operations on the error, torque saturation, friction, and backlash in these systems. The major difference from the "classical" approach was that the nonlinearities were often approximated by linear segmented characteristics, with the result that the phase plane could be divided up into several regions with trajectories in each region described by different linear differential equations. Models of nonlinear phenomena can often be obtained using linear segmented characteristics, but in some cases they may involve significant simplifications of the actual effect. Backlash, for example, is often modeled using the nonlinear characteristic of Fig. 1, which assumes a high friction-to-inertia ratio for the load shaft, and is never satisfied precisely in practice. In some situations, when acceleration or deceleration is taking place in a system with backlash, multiple impacts may occur between gear teeth so that the modeling of the precise phenomenon is extremely difficult.

Other advantages of the phase plane method are that systems that have changes in their parameters or have more than one nonlinear element can be studied. A simple example of the latter is the block diagram of Fig. 2, which is an approximate model of a servomechanism with nonlinear effects due to torque saturation and Coulomb friction.

The differential equation of motion in phase variable form is

$$\dot{x}_2 = f_S(-x_1) - (1/2)\operatorname{sgn} x_2 \qquad \text{with } \dot{x}_1 = x_2 \tag{10}$$

and where $f_S$ denotes the saturation nonlinearity and sgn the signum function, which is +1 for $x_2 > 0$ and -1 for $x_2 < 0$. It is easily seen that there are six linear differential equations that describe the motion in different regions of the phase plane. For $x_2$ positive Equation (1) can be written

$$\ddot{x}_1 + f_S(x_1) + 1/2 = 0$$

so that for

(a) $x_2 + ve, x_1 < -2$

we have

$$\dot{x}_1 = x_2, \dot{x}_2 = 3/2,$$

a parabola in the phase plane.

(b) $x_2 + ve, |x_1| < 2$

we have

$$\dot{x}_1 = x_2, \dot{x}_2 + x_1 + 1/2 = 0$$

a circle in the phase plane.

(c) $x_2 + ve, x_1 > 2$

we have

$$\dot{x}_1 = x_2, \dot{x}_2 = -5/2,$$

a parabola in the phase plane. Similarly, for $x_2$ negative

(d) $x_2$ - *ve*, $x_1 < -2$

we have

$$\dot{x}_1 = x_2, \dot{x}_2 = 5/2,$$

a parabola in the phase plane.

(e) $x_2$ - *ve*, $|x_2| < 2$

we have

$$\dot{x}_1 = x_2, \dot{x}_2 + x_1 - 1/2 = 0,$$

a circle in the phase plane.

(f) $x_2$ - *ve*, $x_1 > 2$

we have

$$\dot{x}_1 = x_2, \dot{x}_2 = -3/2,$$

a parabola in the phase plane. Since all the phase plane trajectories are described by simple mathematical expressions, it is straightforward to calculate specific phase plane trajectories for any initial conditions. In other cases the equations of motion in the various regions may be more complicated, if, for example, in Fig 2 viscous friction is included and $x_2$ is still taken equal to $\dot{x}_1$, but can often be easily sketched. Knowledge of the phase plane approach is obviously also of value in obtaining a "feel" for the system's behavior and interpreting simulation results.

Many papers and books were written during the 1950s that analyzed problems using the phase plane technique but they involved no new fundamental theoretical concepts [10-14]. The major contributions investigated different nonlinear effects in specific second-order systems, for example, to provide an understanding of the effects of torque saturation, nonlinearity in the error channel, backlash, friction, and relay control in second-order systems. Optimum control using relays was studied, and the phenomena of chattering in relay systems was also understood.

A significant number of books written in the period 1950-1960, such as references [15-19], give a detailed coverage of the phase plane approach.

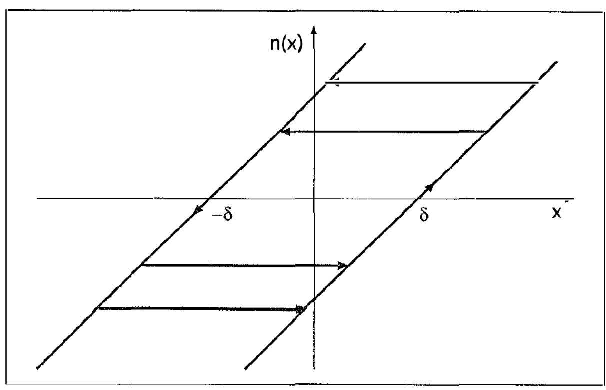



*Fig. 1. Backlash characteristic.*

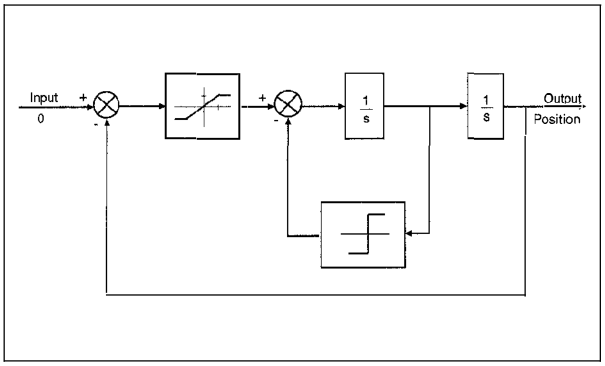



*Fig. 2. Block diagram of servomechanism.*

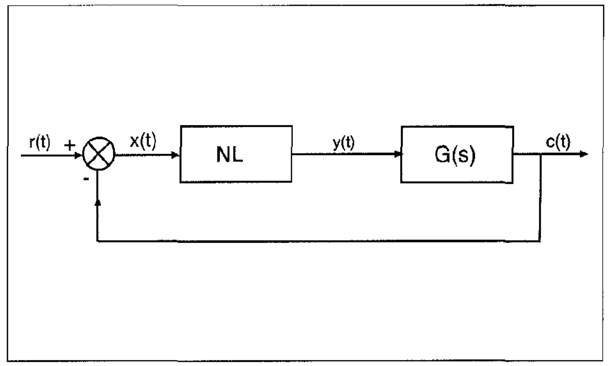



*Fig. 3. Simple feedback loop.*

## The Describing Function Method

The describing function (DF) method appears to have been used independently during the wartime period by Goldfarb in Russia, Dutilh in France, Oppelt in Germany, Kochenburger in the U.S., and Daniell in the U.K. [20], but publications describing their work did not appear until later. The method is identical to a harmonic balance approach, where only the first harmonic is balanced, but was developed in a way more suitable for use in feedback control, where typically nonlinear systems were being modeled in terms of interconnected blocks of static nonlinear and dynamic (transfer function) elements. The DF, N(a), of a nonlinear element was defined as the ratio of the fundamental output to the magnitude of an applied sinusoidal input. Thus, when using the DF in analysis the higher harmonics produced by a sinusoidal input are neglected, and both the idea of using the method and its justification arose from the fact that observations of the output of servomechanisms in which limit cycles occurred often revealed them to be nearly sinusoidal, indicating good low-pass filtering had taken place. Most of the early uses of the DF were directed at the study of feedback systems with a single nonlinear effect, such as backlash, dead zone, saturation, and friction or containing a relay. Many papers were written in the '50s using the DF to study the occurrence of limit cycles and to study system stability, since it was widely assumed, for the situation of Fig. 3, of a single odd symmetrical nonlinear element plus dynamics, but had not yet been proven, that if the linear system was stable for all gains within the sector occupied by the odd symmetric nonlinearity, illustrated in Fig. 4 for positive values only, any instability was only possible in the form of a limit cycle.

Aizermann [21] had presented his famous conjecture, which was soon invalidated by the innovative choice of counterexamples, that the nonlinear system of Fig. 3 would be stable if the linear system with gain K replacing the nonlinear element was stable for all values of K within the sector occupied by the nonlinearity.

Few of the aforementioned papers contributed anything significantly new to describing function theory but provided insight into the effects of nonlinearity in the behavior of a large number of practical systems. Two exceptions were extensions of the approach to determine the stability of any predicted limit cycle [22] and use of the DF method to determine the forced harmonic response of a nonlinear system [23,25]. Graphical methods developed for studying the latter situation enabled a much clearer

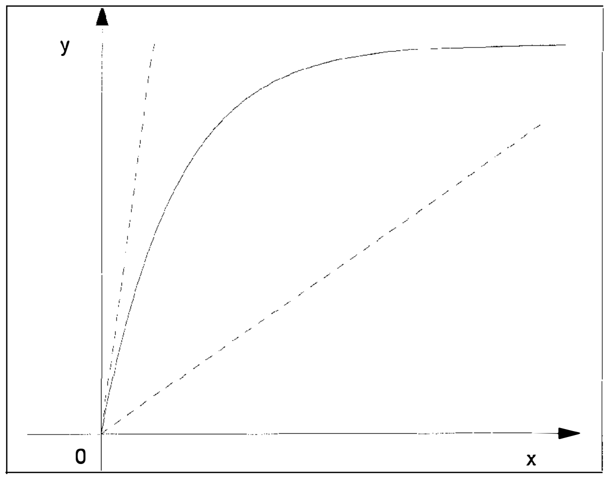



*Fig. 4. Sector bounded nonlinearity.*

appreciation of the jump phenomenon, which may occur in the frequency response of a nonlinear system.

Considering a nonlinear element n(x) with input $x = a \sin\theta$ and corresponding output $y(\theta)$ then if n(x) has odd symmetry the fundamental component in $y(\theta)$, namely $b_1 \sin\theta + a_1 \cos\theta$, has

$$b_1 = \frac{2}{\pi}\int_0^{\pi} y(\theta)\sin\theta d\theta \tag{11}$$

$$a_1 = \frac{2}{\pi}\int_0^{\pi} y(\theta)\cos\theta d\theta \tag{12}$$

The describing function, N(a), is then given by

$$N(a) = (b_1 + ja_1)/a \tag{13}$$

which will be real, that is, $a_1 = 0$, if the nonlinearity is single-valued. To investigate the possibility of a limit cycle in the system of Fig. 3 the characteristic equation

$$1 + N(a)G(s)|_{s=j\omega} = 0 \tag{14}$$

is then examined. Typically this is done using a Nyquist diagram where the loci $G(j\omega)$ and $C(a) = -1/N(a)$ are plotted, and any intersection of the loci gives the amplitude and frequency of a possible limit cycle. The stability of the limit cycle is normally obtained from the direction of crossing of the two loci using the criterion due to Loeb [22], which is a necessary but not sufficient condition. Tables of describing functions for various nonlinearities are given in many books, and the method is easy to apply as shown in Fig. 5, which is for the case of an on-off relay with hysteresis, having output h of $\pm 1$ and hysteresis $\Delta$ of $\pm 1$, controlling a linear plant with transfer function $5e^{-0.5s}/(s+1)^2$. It is easy to show for this relay that

$$C(a) = -\frac{\pi}{4h}\left[\left(a^2 - \Delta^2\right)^{1/2} - j\Delta\right] \tag{15}$$

which is a line parallel to the negative real axis. The limit cycle solution is given from the intersection of $C(a)$ and $G(j\omega)$ in Fig. 5.

Several papers discussed compensation of nonlinear systems to avoid limit cycles predicted by the DF method. The most common procedure was to design a linear compensator, $G_c(j\omega)$, to change the open-loop dynamics so that no intersection existed between the $C(a)$ and $G_cG(j\omega)$ loci on the Nyquist diagram. Other approaches included nonlinear compensators which, essentially, modified the inherent system nonlinearity by placing another nonlinearity effectively either in series or parallel. Some nonlinear integrators were also proposed and used successfully in specific problems [26, 27].

In the late '50s many papers developed extensions to the basic DF method so that more complicated behavior of nonlinear systems could be examined. The first extension was to allow for the situation where the input signal to the nonlinearity also contained a bias signal. This is typically the case when the nonlinearity is asymmetrical or the loop contains constant disturbance or reference signals. This situation requires a sine plus bias describing function with the nonlinearity being represented by two gains, one to the bias $\gamma$ and one to the sinusoid $a$, both of which are functions of $a$ and $\gamma$. To solve for a possible limit cycle the bias and fundamental are balanced around the loop resulting

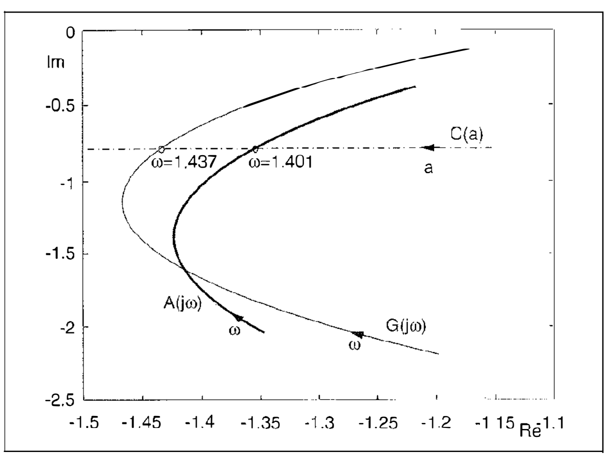



*Fig. 5. Graphs for the solutions of the relay oscillation frequency by the DF and Tsypkin methods.*

in a bias balance equation as well as an equation similar to Equation (14).

As mentioned previously, control engineers interested in nonlinear problems knew of the earlier studies that revealed the existence of specific nonlinear phenomena such as jump phenomena, subharmonic oscillations, and frequency entrainment in simple nonlinear differential equations, and they were therefore interested in extending the describing function approach to study the possibilities of the occurrence of these phenomena in control loops, particularly servomechanisms. It was realized that to do this the response of nonlinear elements to two harmonic inputs would need to be investigated. West et al. [28] called this the dual input describing function and examined the response of single-valued nonlinearities to inputs of the form $a\sin\theta + b\sin(n\theta + \varphi)$, where $n$ was an odd integer. The resulting gains for the nonlinearity to the fundamental and nth harmonic are functions of the three parameters a, b, and $\varphi$, as well as the chosen value of n. Analytical results are only easily found for simple nonlinearities such as a cubic so that to obtain results for saturation West et al. [28] had to use computational techniques. From their computations for the dual input describing function they were able to obtain conditions for the existence of subharmonic oscillations and jump phenomena in saturating servomechanisms. The solution to the latter problem requires the dual input describing function for the special case of n = 1, when it is usually referred to as the incremental describing function. It was also shown how the incremental describing function could be used to assess the stability of a limit cycle [29].

Similar work was being done in the U.S. by Oldenburger [30], who concentrated primarily on the situation of a system with two unrelated sinusoidal inputs. In this case the dual input describing function for the nonlinearity depends upon the input amplitudes $a$ and $b$ only. He used his formulation to examine such effects as the introduction of dither into a system to quench a limit cycle and the possibilities for changing the behavior of a nonlinear system by the injection of another signal. In problems involving dual input describing function methods it was often found that the accuracy of the approach was not as good as that for the single input case due to the increased distortion components present at the nonlinearity input caused by cross modulation products.

It was not surprising that after the considerable interest aroused in control problems involving random inputs during the 1940s by Wiener [31] that engineers should start to examine the problem of nonlinear systems with random inputs. The pioneering work was done by Booton [32], who used the term equivalent gain for the nonlinearity describing function when the input was Gaussian. His approach was to approximate the nonlinearity by a linear gain, such that the error between the nonlinearity output and that from the linear gain with the same random input applied to both was a minimum. He used minimization of the mean squared error as his criterion and obtained an "amplitude" dependent gain, $K_{eq}$, which is a function of the root mean square value of the random input. Other contributions were made by Barrett and Coales [33] and West and Nikiforuk [34]. It was soon realized, due to the properties of Fourier Series, that Booton's definition of equivalent gain was consistent with that of the sinusoidal DF, and it was later shown that the equivalent gain is not only the best gain but also the best linear filter approximation for a nonlinearity for sinusoidal or Gaussian inputs [35,36]. The approach was soon extended to Gaussian plus bias and multiple sinusoidal and Gaussian inputs, with similar work being donc in the U.K. [37], U.S. [38], and Japan [39]. Good coverage of describing function methods can be found in references [40] and [41].

## Relay Systems

During the late '40s and early '50s, relays were finding increasing application in control systems as a relatively cheap and reliablc power amplifier. It was rcalized that, unlikc other continuous nonlinear elements, the output from a relay, once it had switched, became independent of the input. This led to the development by Hamel in France [42] and Tsypkin in Russia [43] of techniques for the accurate evaluation of limit cycles in relay systems. Both started by assuming a periodic form for the relay output, with Hamel then working in the time domain and Tsypkin in the frequency domain to produce the same solution. For a relay with no dead zone, the resulting graphical approach to find any limit cycle from both methods, although the curves are labeled differently, is essentially the same. The Tsypkin approach is briefly reviewed for a system controlled by an on off relay with hysteresis, and the reader is referred elsewhere for more detailed information [40, 41, 44]. Consider the system of Fig. 6, where the relay output has the basic square waveform shown. Thc output can be expressed as the Fourier series

$$y(t)=\frac{4h}{\pi}\sum_{n=1(2)}^{\infty}\frac{1}{n}\sin n\omega t \tag{16}$$

and correspondingly the plant output $c(t)$ and its derivative $\dot{c}(t)$ are given by

$$c(t)=\frac{4h}{\pi}\sum_{n=1(2)}^{\infty}\frac{1}{n}\left|G(jn\omega)\right|\sin(n\omega t+\phi_n) \tag{17}$$

$$\dot{c}(t)=\frac{4h\omega}{\pi}\sum_{n=1(2)}^{\infty}\frac{1}{n}\left|G(jn\omega)\right|\cos(n\omega t+\phi_n) \tag{18}$$

provided $\lim_{s\to\infty}sG(s)=0$, where $\phi_n=\angle G(jn\omega)$. The analysis is possible [41] provided $G(s)$ is proper but the aforementioned assumption is made here to simplify the presentation.

For this output to be generated as assumed the relay input $-c(t)$ must satisfy the conditions

$$-c(0)=\Delta, \quad -\dot{c}(t)>0 \tag{19}$$

By substituting into these conditions the expressions for $c(0)$ and $\dot{c}(0)$ from Equations (17) and (18) it is easy to show that the necessary condition for a limit cycle becomes the locus $\Lambda(\omega)$, must satisfy the conditions.

$$Re\Lambda(\omega)<0 \quad \text{and} \quad Im\Lambda(\omega)=-\Delta \tag{20}$$

where

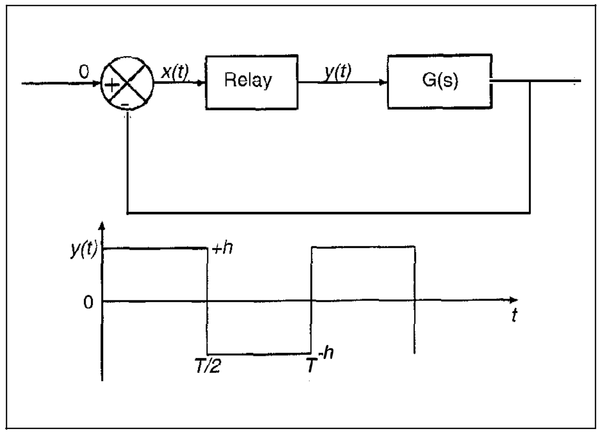



*Fig. 6. Relay system and relay output waveform.*

$$\Lambda(\omega)=\frac{4h}{\pi}\left[\sum_{n=1(2)}^{\infty}U_G(n\omega)+\frac{j}{n}V_G(n\omega)\right] \tag{21}$$

Herc $G(j\omega)=U_G(\omega)+jV_G(\omega)$ and closed forms for thc summations are available for specific transfer functions $G(j\omega)$. Solutions for any limit cyclc frequency can thcn be found by finding the intersections of $A(\omega)=(\pi/4h)\Lambda(\omega)$ and the relay DF on a Nyquist diagram as shown in Fig. 5 for the examplc given in the previous section.

Since in the case considered the relay has no dead zone, only one nonlinear equation has to bc solvcd to obtain thc solution for the limit cycle frequency. If, however, the relay has dead zone, the method yields two equations to solve for the two unknown parameters, namely the limit cycle frequency and the pulse width. These equations are only necessary conditions, so to ensure the solution is valid the relay input waveform must be evaluated to check that it does produce the relay output initially assumed, that is, there are no "false" switchings. It was also shown by Tsypkin that the stability of any predicted limit cycle can be determined exactly, which is a major advantage of this approach. The method was also extended to determine asymmetrical limit cycles in an autonomous system and the forced response to a sinusoidal input, while in recent years several further extensions have been made.

## Hardware Technology

It is appropriate to comment briefly on the hardware used in simulation and also that available for implementing control systems during the 1950s. Simulation studies of complex nonlinear systems were greatly affected by the time taken to set up the simulation, the slow speed of the simulation, and the reliability of the hardware. Similarly, the hardware that was used to implement control systems was, by today's standards, bulky, heavy, noisy, and expensive, and limited the complexity of control algorithms that could be implemented within an allowable cost.

### Simulation

As mentioned earlier, differential analyzers were built in the 1930s to perform simulations, and these were succeeded in the wartime period by electronic simulators. Reference [45], by

Williams and Ritson, gives an excellent account on the status of the simulators toward the end of that period. The heart of the simulator was the vacuum tube DC amplifier, the forerunner of today's operational amplifier. Williams did a lot of work in this area during the wartime period and maintained close contact with the work done at MIT. He was a co-author of the volume *Vacuum Tube Amplifiers* published in the Radiation Laboratory series by McGraw-Hill. Compared to today's operational amplifiers they were significantly inferior in terms of size, reliability, cost, power consumption, and accuracy. Typically the amplifiers were supplied from ±300 volts DC and of course required 6.3 volts AC for the heaters. Because of drift problems, sophisticated techniques of chopper stabilization had to be used in the amplifiers, especially when they were required to be used as integrators.

To implement nonlinear functions was also extremely tedious. Early function generators used biased vacuum tube diode circuits to approximate specific characteristics by adjusting the break point and slope for the contribution from each individual diode. Even operations which we now take for granted, such as multiplication, were very difficult. Several forms of multiplier were available, but the most popular one was probably the quarter squares technique, that is, the product of $xy$ was determined from $(x+y)^2 - (x-y)^2$. This required operational amplifiers to add and subtract signals, as well as square law characteristics which were normally again implemented using diode circuitry. Costs were also a problem; good simulators used chopper stabilized DC amplifiers that, in today's valuation, would cost several thousand dollars. Put another way, a good operational amplifier cost more than my month's salary when I first started as an assistant lecturer, as so did a good multiplier.

Simulation of process control problems invariably required a time delay. Before the days of hybrid computers, which appeared after the time considered in this presentation, time delays had to be approximated using sophisticated transfer functions, as, for example, given in the Padé table. Implementation of these, because of the cost of amplifiers and components, made them very expensive. Simulators were required by the government defense agencies and industry, in particular, for studies of high-speed flight of aircraft and guided missiles. Typical of one such installation was TRIDAC, a large analog computing machine [46] installed at the Royal Aircraft Establishment in the U.K. (The acronym stands for *tridimensional analog computer*.) The project took four years from its conception in 1950 to completion in 1954. The machine used electronic, mechanical, and hydraulic components, and was housed in its own building. The power consumption was 600kW, of which 200kW was for the electronic components, which consisted of more than 8,000 vacuum tubes. It cost around £0.5 million, which, in today's terms, would be of the order of £20 million, and you can now handle the same problems better with a standard simulation language on a good PC. Two pictures relating to TRIDAC are shown in Figs. 7 and 8. The former shows the huge plotter used to provide graphical output being viewed by Princess Margaret at the official opening ceremony. The latter shows part of the control console, with an old Cossor oscilloscope clearly visible at the left-hand side.

### System Implementation

The technology used in control system hardware was again far removed from today's situation. Vacuum tubes were used for both operational amplifiers and power amplifiers, and apart from requiring ±300 volts DC and 6.3 volt AC heaters, as mentioned previously, they were also bulky, unreliable, and expensive when compared to the situation today, and also not readily usable in harsh environments. This meant that the implementation of quite simple active transfer functions for compensators and filters was expensive, as was any algorithm which might be required to include a nonlinear function.

Research, in fact, was carried out on synthesizing high-order transfer functions using complex networks around a single operational amplifier simply because at that time the operational amplifiers were much more expensive than resistors and capacitors, although low-loss capacitors were by no means inexpensive. Today, of course, the whole strategy is different, since operational amplifiers are now as cheap as the passive components. Power electronics was in its infancy and vacuum tube power amplifiers could not supply the types of powers that are taken for granted these days. They were typically used only to drive small motors; for larger motors, devices such as mercury arc rectifiers, magnetic amplifiers, metadynes, and amplidynes were used. The technology, therefore, was not available to implement sophisticated control algorithms, even in those cases where economic considerations did not enter into the equation.

## Publications

It is possibly appropriate to reflect on both the number of and topics covered by control engineering publications around 1960.


*Fig. 7. TRIDAC plotting table.*

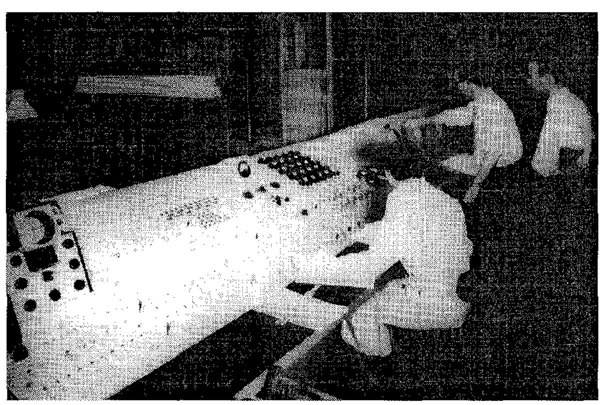



*Fig. 8. TRIDAC operators' console.*

As previously mentioned, the first publication of what has become our Society was the May 1956 issue of the IRE PGAC. The volume contained seven papers in approximately 90 pages, and further issues followed in February 1957 and March 1958. Of the seven papers in the first issue, three dealt with nonlinear problems. Other English-language publications covering control engineering were the IEE proceedings, Parts A and C, the AIEE, the ASME, the I.Mech.E., and some covering process control and aerospace control. In total, these journals were publishing around 30 papers per annum on control engineering, of which, typically, around 30% dealt with nonlinear problems.

The first IFAC Congress was held in Moscow in 1960, and 285 papers were presented. IFAC at the time had four official languages—English, French, German, and Russian—and therefore not all the conference papers are available in the proceedings of the Congress published by Butterworth's in 1963. This is probably not surprising if you have seen the results of the translation into English hy way of Russian of a paper first written in Italian! One of the volumes [47] published by Butterworths from material presented at the first IFAC Congress held in Moscow is entitled *Theory of Nonlinear Control Systems* and contains 14 papers. The first paper, entitled "Theory of Nonlinear Control," is a review paper by Y.H. Ku, which is primarily of interest for the quite extensive list of around 100 references. Unlike the current situation, it was common to publish discussion comments on papers at the time, and many of the papers included in this book have quite interesting discussion contributions. Of interest in the discussion to this first paper are the comments of M. Aizermann on the need for caution in using the describing function approach. The second paper, by E.P. Popov, discusses systems with relatively complex block diagrams containing one or two nonlinearities. The paper uses the approaches of harmonic and statistical linearization in the analysis. The third paper is by V.M. Popov, who had recently published his now well-known stability criterion. The major contribution of the paper is in drawing attention to the applications of this theorem. The next paper, by Shen, considers nonlinear compensation of servomechanisms that have backlash. The work was done using the describing function approach, and an interesting contribution to this discussion is by Kochenburger who discusses the problem of modeling backlash. The next paper also uses the describing function method to discuss the effects of amplifier saturation in process control. Again using the describing function approach, the following paper by Kochenburger examines the effect of power source regulation on the response of a power amplifier in a feedback control system. The example is concerned with a hydraulic pump where the prime mover driving the pump cannot maintain a constant speed as the pump load is increased.

The next paper, by Takahashi et al., looks at the problem of process control when the valve controlling the process has a velocity limit. The analyzis is conducted using both the phase plane and describing function methods and the results compared. Goldfarb provides an interesting comment in the discussion regarding block diagram rearrangement which could be helpful when using the describing function approach. This is followed by a paper by Gille et al., which is written in French and concerns investigations into forced oscillations in nonlinear systems. The work is concerned with a technique for deriving a region in the complex plane which the Nyquist locus must not enter if a specific system is to have no jump phenomena in its frequency response. The results are similar to those given in Reference [28]. The paper by Macura deals with the linearization of nonlinear functions of more than one variable, and this is followed by a paper by Clauser that examines multiple frequency effects primarily in nonlinear electric circuits. Forced oscillations and subharmonic resonances are considered. The paper on signal stabilization of self-oscillating systems by Oldenburger and Nakada discusses the same topic as Reference [30], namely the addition of another deterministic signal into a nonlinear system in order to change its response. The following paper, by West, entitled "Gain-Modulated Control Systems," examines how the performance of some systems can be changed by varying the gain of the loop signal according to some other system signal. This, of course, involves the use of multipliers in the system. The paper by Ferner examines an advantage of using nonlinear controllers for controlling linear plants, and the final paper looks at an interesting practical problem, namely, the control of plants whose response is direction dependent. Perhaps the most interesting aspect about all these 14 papers is that a large number were concerned with practical problems of nonlinearity in specific systems, and some novel solutions were found even though the supporting analyzes were often only approximate. A few papers on nonlinear control also appeared in some of the other volumes published; for example, three papers on relay control were published in the volume on *The Theory of Optimal Control*.

## Reflections and Conclusions

Although several methods have been developed for the study of nonlinear systems in the last three decades, they all have their limitations, and many problems remain unsolved. Perhaps the most important, and easily defined, one is that no necessary and sufficient conditions have been found for the absolute stability of the feedback system Fig. 3 with a given nonlinearity and transfer function G(s). This, in principle, is a relatively simple problem when compared with, say, the design of a controller for a nonlinear plant to ensure that certain performance specifications are met, since this involves two nonlinear dynamic components. Simple linear controllers such as PID or phase lead were developed historically not because they were the optimal solutions to a problem but because they often provided satisfactory solutions at a reasonable cost with the available technology. Today, with microprocessor controllers, the complexity of the control algorithm, whether linear or nonlinear, may have a relatively small effect on the system cost, so the opportunity exists to develop complex control algorithms. Without a general theory of nonlinear systems it is difficult to compare the relative merits of different solutions to a problem even when good specifications are defined for the requirements of the completed design, which is rare in academic papers. Performance specifications for a nonlinear control problem may also be quite different from the linear situation; for example, a major requirement could be that the process is controlled so that a particular mode of operation, say a limit cycle, does not exist, and other factors, such as speed of set point change, may be relatively unimportant. The "true" performance index for better control in industry, assuming it is safe and environmentally friendly, is improved economic benefits, which in many situations is difficult to relate to requirements of an analytical design.

Perhaps this is a roundabout way of saying that nonlinear control problems tend to be unique and a variety of tools are that they use sound physical principles and therefore can be useful even when used as approximations in revealing possible aspects of nonlinear behavior or the effect of changes in parameters, so as to obtain a better understanding of the system behavior. This is particularly important with nonlinear systems, because although modern simulation facilities are excellent, checking the system behavior throughout the entire state space, which is necessary unlike the situation for a linear system, can be an exhaustive task.

## References


[1] A Lindstedt, "Differentialgleichungen der Störungs-theorie," Mem. Acad. Imp. St. Petersburg, 31, 1883.

[2] H. Poincaré, *Les Méthodes Nouvelles de la Méchanique Céleste*, vol. 1, Gauthier-Villars, Paris, 1892.

[3] B. Van der Pol, "On Relaxation Oscillations," *Philos. Mag.* 7(2), pp. 978, 1926.

[4] U. Krylov and N. Bogoliubov, *Introduction to Nonlinear Mechanics*, Princeton University Press, Princeton, NJ, 1943.

[5] G. Duffing, "Erzwungene Schwingungen bei Verländerlicher Eigenfrequenz," *Vieweg*, Braunschweig, 1918.

[6] D.P. Atherton, "Nonlinear Systems: History," *Encyclopedia of Systems and Control*, Pergamon Press, pp. 3,383-3,390, 1987.

[7] A.M. Lyapunov, *Obshchaya Zadacha ob Ustoichivosti Dvizheniya (General Problem of the Stability of Motion)*, Gostekhizdat, Moscow, 1915.

[8] S. Bennett, *A History of Control Engineering 1930-1955*, IEE Control Engineering Series No. 47, Peter Peregrinus, London, U.K., 1993.

[9] N. Minorsky, "Control Problems," *J. Franklin Institute*, Nov.-Dec., pp. 451-487, pp. 519-551, 1941.

[10] J.C. West, J.L. Douce, and R. Naylor, "The Effects of Some Nonlinear Elements on the Transient Performance of a Simple R.P.C. System Possessing Torque Limitation," *Proc. I.E.E.*, 101, pp. 156-165, 1954.

[11] R.E. Kalman, "Phase Plane Analysis of Automatic Control Systems with Nonlinear Gain Elements," *Trans. A.I.E.E.*, 73(II), pp. 383-390, 1954.

[12] T.J. Higgins, "A Resume of the Development and Literature of Nonlinear Control System Theory," *Trans. A.S.M.E.*, 79, pp. 445-449, 1957.

[13] W.H. Pell, "Graphical Solution of Single Degree of Freedom Vibration Problem with Arbitrary Damping and Restoring Forces," *Trans. A.S.M.E.*, 79, 1957, pp. 311-312.

[14] L.S. Jacobsen, "On a General Method of Solving Second Order Differential Equations by Phase Plane Displacements," *Trans. A.S.M.E.*, 74, 1952, pp. 543-553.

[15] A.A. Andronov and C.E. Chaikin, *Theory of Oscillations*, Princeton University Press, NJ, 1949.

[16] I. Flügge-Lotz, *Discontinuous Automatic Control*, Princeton University Press, Princeton, NJ, 1953.

[17] Y.H. Ku, *Analysis and Control of Nonlinear Systems*, Ronald, New York, 1958.

[18] W.J. Cunningham, "Introduction to Nonlinear Analysis," McGraw-Hill, New York, 1958, pp. 36-39.

[19] J.C. West, *Analytical Techniques for Nonlinear Control Systems*, E.U.P., London, 1960, Chapter 6.

[20] D.P. Atherton, *Nonlinear Control Engineering*, Van Nostrand Reinhold Co. Ltd., Berks, UK, 1975.

[21] M.A. Aizerman, *Lectures on the Theory of Automatic Regulation*, Fitzmatgiz, 1958, Moscow.

[22] J.M. Loeb, "Recent Advances in Nonlinear Servo Theory,".in Oldenburger R, ed., *Frequency Response*, Macmillan, New York, pp. 260-68, 1956.

[23] J.C. West and J.L. Douce, "The Frequency Response of a Certain Class of Nonlinear Feedback Systems," *Br. J. Appl. Phys.*, 5, 210-10, 1954.

[24] L.T. Prince, "A Generalised Method for Determining the Closed Loop Frequency Response of Nonlinear Systems," *Trans A.I.E.E.*, 73, pt. II, pp. 217-224, 1954.

[25] P.E.W. Grensted, "The Frequency-Response Analysis of Non-Linear Systems," *Proc I.E.E.*, 1955, 102C, pp. 244-253.

[26] J.C. West and M.J. Somerville, "Integral Control with Torque Limitation," *Proc I.E.E.*, 1956, 103, p. 407.

[27] J.C. Clegg, "A Nonlinear Integrator for Servomechanisms," *Trans A.I.E.E.*, 1958, pt. II, 77, pp. 41-42.

[28] J.C. West, J.L. Douce, and R.K. Livesley, "The Dual Input Describing Function and Its Use in the Analysis of Nonlinear Feedback Systems," *Proc. Inst. Electr. Eng.*, Part B 103, pp. 463-74, 1956.

[29] R. Oldenburger, "Signal Stablization of a Control System," *Trans. ASME*, 79, pp. 1869-72, 1957.

[30] Z. Bonenn, "Stability of Forced Oscillations in Nonlinear Feedback Systems," *I.R.E. Trans.*, 1958, AC-6, pp. 109-111.

[31] N. Weiner, "Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications," Wiley, New York, 1949.

[32] R.C. Booton, "Nonlinear Control Systems with Random Inputs," *IRE Trans. CT-1*, pp. 9-17, 1954.

[33] J.F. Barrett and J.F. Coales, "An Introduction to the Analysis of Non-Linear Control Systems with Random Inputs," *Proc. I.E.E.*, 1955, 103C, pp. 190-199.

[34] P.N. Nikiforuk and J.C. West, "The Describing Function Analyzis of a Nonlinear Servomechanism Subject to Stochastic Signals and Noise," *Proc. I.E.E.*, 1957, 104C, pp. 193-203.

[35] A.H. Nuttall, "Theory and Application of the Separable Class of Random Processes," M.I.T. Res. Lab. Electronics, 1958, Rept. 343.

[36] J.L. Brown, "On a Cross Correlation Property for Stationary Random Processes," *Trans. I.R.E.*, 1957, IT-3, pp. 28-31.

[37] M.J. Somerville and D.P. Atherton, "Multi-Gain Representation for a Single-Valued Nonlinearity with Several Inputs, and the Evaluation of their Equivalent Gains by a Cursor Method," *Proceedings*, IEE, 105C, pp. 537-549, 1958.

[38] R. Oldenburger and R. Sridhar, "Signal Stabilization of a Control System with Random Inputs," *Trans. Am. Inst. Electr. Eng.*, Part 2, 80, pp. 260-67 1961.

[39] Y. Sawaragi and S. Takahashi, "Response of Control Systems Containing Zero-Memory Non-Linearity to Sinusoidal and Gaussian Inputs," *Proc. Heidelberg Conf. Automatic Control*, International Federation of Automatic Control, Laxenburg, pp. 271-74, 1956.

[40] A. Gelb and W.E. Van der Velde, "Multiple Input Describing Functions and Nonlinear Systems Design," McGraw-Hill, New York, 1968.

[41] D.P. Atherton, "Nonlinear Control Engineering: Describing Function Analysis and Design," Van Nostrand Reinhold, London, 627 pp., September 1975; or D.P. Atherton, "Nonlinear Control Engineering," student edition, Van Nostrand, Reinhold, 470 pp., 1982.

[42] B. Hamel, "Contribution a l'Etude Mathematique des Systemes de Reglage par Tout-ou-rien," CEMV, Service Technique Aeronautique, 17, 1949.

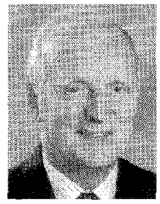
[43] J.A. Tsypkin, *Theorie der Relais Systeme der Automatischen Regelung*, R. Oldenbourg-Verlag, Munich, 1958.

[44] P.A. Cook, *Nonlinear Dynamical Systems*, Prentice-Hall, 1986.

[45] F.C. Williams and F.J.U. Ritson, "Electronic Servo Simulators," *JIEE*, 94 (IIA), pp. 112-24, 1947.

[46] F.R.J. Spearman et al., "TRIDAC: A Large Analogue Computing Machine," IEE monograph, Paper No. 1899M, October 1955.

[47] J.F. Coales et al., eds., "Theory of Non-Linear Control," Butterworths, London, 1963.



**Derek Atherton** was born in Bradford, U.K., in 1934 and studied at the universities of Sheffield and Manchester, from which he obtained B.Eng. and Ph.D. degrees. He spent several years in Canada before taking up his present position as professor of control engineering at the University of Sussex in 1980. He was president of the Institute of Measurement and Control in 1990, president of the IEEE Control Systems Society in 1995, and is currently a member of the IFAC Council. His major research interests are in nonlinear control theory, computer-aided control system design, simulation, target tracking, and parallel processing. He has written two books and is co-author of a third and has approximately 200 publications in journals and conference proceedings.