

Ex. PGS 1037

A Brief History of Automatic Control

Stuart Bennett

Automatic feedback control systems have been known and used for more than 2000 years; some of the earliest examples are water clocks described by Vitruvius and attributed to Ktesibios (circa 270 B.C.). Some three hundred years later, Heron of Alexandria described a range of automata which employed a variety of feedback mechanisms. The word “feedback” is a 20th century neologism introduced in the 1920s by radio engineers to describe parasitic, positive feeding back of the signal from the output of an amplifier to the input circuit. It has entered into common usage in the English-speaking world during the latter half of the century.

Automatic feedback is found in a wide range of systems; Rufus Oldenburger, in 1978, when recalling the foundation of IFAC, commented on both the name and the breadth of the subject: “I felt that the expression ‘automatic control’ covered all systems, because all systems involve variables, and one is concerned with keeping these variables at constant or given varying values. This amounts to concern about control of these variables even though no actual automatic control devices may be intentionally or otherwise incorporated in these systems. I was thinking of biological, economic, political as well as engineering systems so that I pictured the scope of IFAC as a very broad one.”

This diversity poses difficulties for historians of the subject (and for editors of control journals), and this article does not attempt to cover all application areas.

The history of automatic control divides conveniently into four main periods as follows:

- Early Control: To 1900
- The Pre-Classical Period: 1900-1940
- The Classical Period: 1935-1960
- Modern Control: Post-1955

This article is concerned with the first three of the above; other articles in this issue deal with the more recent period.

Early Control: To 1900

Knowledge of the control systems of the Hellenic period was preserved within the Islamic culture that was rediscovered in the West toward the end of the Renaissance. New inventions and applications of old principles began to appear during the 18th century—for example, René-Antoine Ferchault de Réaumur (1683-1757) proposed several automatic devices for controlling the temperature of incubators. These were based on an invention of Cornelius Drebbel (1572-1663). The temperature was measured by the expansion of a liquid held in a vessel connected to a U-tube containing mercury. A float in the mercury operated an arm which, through a mechanical linkage, controlled the draft to a furnace and

hence the rate of combustion and heat output. Improved temperature control systems were devised by Bonnemain (circa 1743-1828), who based his sensor and actuator on the differential expansion of different metals. During the 19th century an extensive range of thermostatic devices were invented, manufactured, and sold. These devices were, predominantly, direct-acting controllers; that is, the power required to operate the control actuator was drawn from the measuring system.

The most significant control development during the 18th century was the steam engine governor. The origins of this device lie in the lift-tenter mechanism which was used to control the gap between the grinding-stones in both wind and water mills. Matthew Boulton (1728-1809) described the lift-tenter in a letter (dated May 28, 1788) to his partner, James Watt (1736-1819), who realized it could be adapted to govern the speed of the rotary steam engine. The first design was produced in November 1788, and a governor was first used early in 1789. The original Watt governor had several disadvantages: it provided only proportional control and hence exact control of speed at only one operating condition (this led to comments that it was “a moderator, not a controller”); it could operate only over a small speed range; and it required careful maintenance.

The first 70 years of the 19th century saw extensive efforts to improve on the Watt governor, and thousands of governor patents were granted throughout the world. Many were for mechanisms designed to avoid the offset inherent in the Watt governor. Typical of such mechanisms were the governors patented by William Siemens (1823-1883) in 1846 and 1853, which substituted integral action for proportional action and hence produced “floating” controllers with no fixed set point. Practical improvements came with the loaded governor of Charles T. Porter (1858); his governor could be run at much higher speeds, and hence greater forces could be developed to operate an actuator. A little later Thomas Pickering (1862) and William Hartnell (1872) invented spring-loaded governors, which also operated at higher speeds than the Watt governor and which had the added advantage of smaller physical size than the Watt and Porter governors.

From the early years of the 19th century there were reports of problems caused by governors “hunting,” and attempts to analyze the governor mechanism to determine the conditions for stable (non-hunting) operation were made. J. V. Poncelet (1788-1867) in 1826 and 1836, and G. B. Airy (1801-1892) in 1840 and 1851 produced papers that showed how dynamic motion of the governor could be described using differential equations, but both met difficulties when they attempted to determine the conditions for stable behavior. Airy, in 1851, stated the conditions for stable operation, but his report is so terse that it is not possible to determine how he arrived at these conditions. In 1868, James Clerk Maxwell (1831-1879) published his now-famous paper entitled “On Governors.” In it he described how to derive the linear differential equations for various governor mechanisms. At this time mathematicians and physicists knew that the

The author is with the Department of Automatic Control & Systems Engineering, The University of Sheffield, Mappin Street, Sheffield S1 3JD, U.K., telephone: +44 (0)114 282 5230, email: s.bennett@sheffield.ac.uk.

stability of a dynamic system was determined by the location of the roots of the characteristic equation, and that a system became unstable when the real part of a complex root became positive; the problem was how to determine the location of the real parts of the complex roots without finding the roots of the equation. Maxwell showed, for second-, third-, and fourth-order systems, that by examining the coefficients of the differential equations the stability of the system could be determined. He was able to give the necessary and sufficient conditions only for equations up to fourth order; for fifth-order equations he gave two necessary conditions. Maxwell's paper, now seen as significant, was little noticed at the time, and it was not until the early years of this century that the work began to be assimilated as engineering knowledge.

The problem formulated by Maxwell was taken up by Edward J. Routh (1831-1907), whose first results were published in 1874. In 1877 he produced an extended treatise on the "Stability of Motion" in which, drawing on the work of Augustin-Louis Cauchy (1789-1857) and Charles Sturm (1803-1855), he expounded what we now know as the Routh-Hurwitz stability criteria. In 1895, the Swiss mathematician Adolf Hurwitz (1859-1919) derived the criteria independently (basing his work on some results of C. Hermite). He had been asked for help with the mathematical problem by his colleague Aurel Boleslaw Stodola (1859-1942), who was working on a turbine control problem.

Most of the inventions and applications of this period were concerned with the basic activities of controlling temperatures, pressures, liquid levels, and the speed of rotating machinery; the desire was for regulation and for stability. However, growth in the size of ships and naval guns, and introduction of new weapons such as torpedoes, resulted in the application of steam, hydraulic, and pneumatic power systems to operate position control mechanisms. In the United States, Britain, and France, engineers began to work on devising powered steering engines to assist the helmsman; on large ships the hydrodynamic forces on the rudder were such that large gear ratios between the helm and the rudder were required and hence moving the rudder took a long time. The first of powered steering engine, designed by Frederick Sickels in the U.S. (patented 1853) was an open-loop system. The first closed-loop steering engine (patented 1866) was designed by J. McFarlane Gray for Brunel's steamship the *Great Eastern*. In France, around the same time, Jean Joseph Farcot designed a range of steering engines and other closed-loop position control systems. He suggested naming his devices "servo-moteur" or "moteur asservi," hence our terms "servo-mechanisms" and "servomotors."

Further applications for control systems became apparent with the growth in knowledge of electricity and its applications. For example, arc lamps required the gap between the electrodes to be kept constant, and generally it was helpful to all users if either the voltage or the current of the electricity supply was kept constant. Electricity also provided additional tools—for measurement, for transmission and manipulation of signals, and for actuation—which engineers began to use. The electric relay, which provided high gain power amplification, and the spring biased solenoid, which provided (crude) proportional control action, were significant devices.

The Pre-Classical Period (1900-1935)

The early years of the 20th century saw the rapid and widespread application of feedback controllers for voltage, current, and frequency regulation; boiler control for steam generation; electric motor speed control; ship and aircraft steering and auto stabilization; and temperature, pressure, and flow control in the process industries. In the twenty years between 1909 and 1929, sales of instruments grew rapidly as Fig. 1 shows. The majority of the instruments sold were measuring, indicating, and recording devices, but toward the end of the period the sales of controllers began to increase. The range of devices designed, built, and manufactured was large; however, most were designed without any clear understanding of the dynamics both of the system to be controlled and of the measuring and actuating devices used for control. The majority of the applications were concerned with simple regulation, and in such cases this lack of understanding was not a serious problem. However, there were some complex mechanisms involving complicated control laws being developed—for example, the automatic ship-steering mechanism devised by Elmer Sperry (1911) that incorporated PID control and automatic gain adjustment to compensate for the disturbances caused when the sea conditions changed. Another example is the electricity supply companies concerned about achieving economic operation of steam-generating boilers. Boiler control is of course a multivariable problem in that both water level and steam pressure have to be controlled, and for efficient combustion the draught to the boiler has also to be controlled. During the 1920s several instrument companies develop complete boiler control systems.

As control devices and systems began to be used in many different areas of engineering, two major problems became apparent: (1) there was a lack of theoretical understanding with

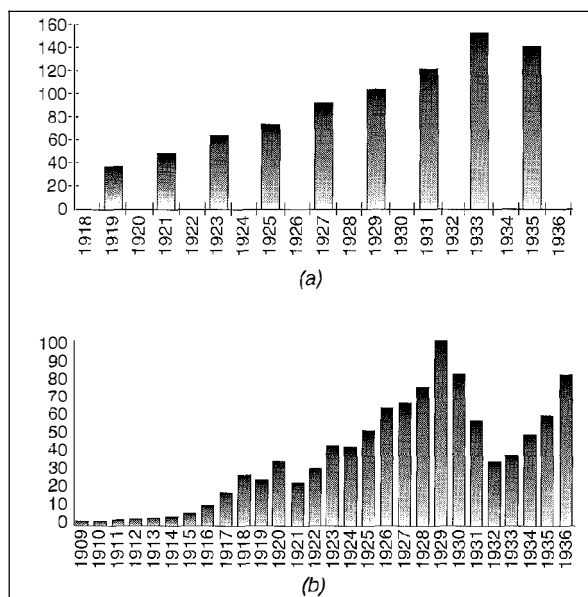


Fig. 1. (a) Ratio of instrument to machinery sales in the United States, 1918 to 1936 (1921 = 100). (b) Index of instrument sales in the United States, 1909 to 1936 (1921 = 100).

no common language in which to discuss problems, and (2) there were no simple, easily applied analysis and design methods. The only available analysis tool was the differential equation and the application of the still not widely known Routh-Hurwitz stability test. This is a laborious process, dependent on being able to obtain values for the parameters, and one that gives no guidance to the designer on the degree of stability, or what to do to make the system stable.

As applications multiplied, engineers became puzzled and confused: controllers that worked satisfactorily for one application, or for one set of conditions, were unsatisfactory when applied to different systems or different conditions: problems arose when a change in one part of the system (process, controller, measuring system, or actuator) resulted in a change in the major time constant of that part. This frequently caused instability in what had previously been, or seemed to have been, a stable system. Some acute observers, for example Elmer Sperry and Morris E. Leeds, noted that the best human operators did not use an on-off approach to control but used both anticipation, backing off the power as the controlled variable approached the set-point, and small, slow adjustments when the error persisted. Sperry tried to incorporate these ideas into his devices, and for many years Leeds resisted attaching simple on-off control outputs to his recorders because he realized that this would not provide good control.

In 1922, Nicholas Minorsky (1885-1970) presented a clear analysis of the control involved in position control systems and formulated a control law that we now refer to as three-term or PID control. He arrived at his law by observing the way in which a helmsman steered a ship. This work did not become widely known until the late 1930s, after Minorsky had contributed a series of articles to *The Engineer*. But even if designers had been aware of Minorsky's work they would still have lacked suitable linear, stable, amplification devices to convert the low power signals obtained from measuring instruments to a power level suitable to operate a control actuator. Slide and spool valves developed during the early part of the 20th century were beginning to provide the solution for hydro-mechanical systems, although valve overlap that resulted in dead space and stiction were problems that had to be overcome. However, there was an impasse with respect to amplifiers for electronic and pneumatic systems. As early as 1920 the amplification problem was proving a serious obstacle to the further development of long-distance telephony. Improvements in cable design and the use of impedance loading had extended the distance over which telephone transmissions could take place without amplification, yet the transcontinental service in the U.S. was dependent on amplification. Telephone repeaters based on electronic amplification of the signal were used around 1920, but the distortion they introduced limited the number that could be used in series. Expansion of traffic on the network was also causing problems since it necessitated an increase in bandwidth of the lines with the consequent increase in transmission loss. Harold Stephen Black (1898-1983) began work on this problem in the early 1920s. He realized that if some of the amplification of a high-gain amplifier were sacrificed by feeding back part of the output signal, the distortion due to noise and component drift could be reduced. On August 2, 1927, he sketched a circuit for a negative feedback amplifier. Following extensive development work, full-scale practical trials were carried out in 1930, and the amplifier began

to be used within AT&T in 1931. Information about the amplifier was not published in the open literature until 1934. In developing the practical amplifier and in understanding its behavior, Black was assisted by Harry Nyquist (1889-1976), whose paper "Regeneration Theory" laid down the foundations of the so-called Nyquist analysis and was published in 1932.

This work provided a practical device—the negative feedback amplifier—and led to a deeper understanding of the benefits of negative feedback in systems. It also, eventually, led to a method of analyzing and designing control systems which did not require the derivation and manipulation of differential equations, and for which experimental data—the measured frequency response—could be combined with calculated data; from the combined response the degree of stability of the system could be estimated and a picture of changes necessary to improve the performance could be deduced.

Contemporaneously with Black's work, Clesson E. Mason of the Foxboro Company developed a pneumatic negative feedback amplifier. Edgar H. Bristol, one of the founders of the Foxboro Company, had invented the flapper-nozzle amplifier in 1914. The early versions of the flapper-nozzle amplifier were highly non-linear (effectively on-off behavior), and during the 1920s extensive modifications had only succeeded in increasing its linear range to about 7% of full range. In 1928, Mason began experimenting with feeding back part of the output movement of the amplifier, and in 1930 produced a feedback circuit that linearized the valve operation. This circuit enabled integral (or reset) action to be easily introduced into the behavior of the system. In 1931, the Foxboro Company began selling the Stabilog pneumatic controller which incorporated both linear amplification (based on the negative feedback principle) and integral (reset) action (Fig. 2). There was some initial market resistance to this device, on the grounds of cost and because its behavior was not understood. Foxboro responded by producing, in 1932, a bulletin explaining the principles of the system in clear and simple terms and stressing how the behavior was different from what it termed "narrow-band" controllers, that is, those with limited linear range.

The electronic negative feedback amplifier and the pneumatic controller were the outcomes of work on industrial problems. During the same period, extensive work was being carried out on analog calculating machines under the direction of Vannevar Bush at the Massachusetts Institute of Technology. This work resulted in the differential analyzer, which provided a means of simulating the behavior of dynamic systems and of obtaining numerical solutions to differential equations. It also led to the study and design of a high-performance servomechanism by Harold Locke Hazen (1901-1980) and his students. In addition to designing a servo system, Hazen also undertook the first major theoretical study of servomechanisms. His papers, published in 1934, provided the starting point for the next generation of control system specialists.

The Classical Period: 1935-1950

During the period 1935-1940, advances in understanding of control system analysis and design were made independently by several groups in several countries. The best known and most influential work came from three groups working in the U.S. The development in Europe and in Russia during this period followed a somewhat different path deriving from Vyschnegradsky's work in

Russia and then Barkhausen's work in Germany, followed by developments due to Cremer, Leonhard, and Mikhailov.

AT&T continued with its attempts to find ways of extending the bandwidth of its communication systems, and upon obtaining good frequency response characteristics. The ideal which they were seeking was a constant gain over a wide bandwidth with a sharp cut-off and with a small phase lag. Engineers in the Bell Telephone Laboratories worked extensively on this problem, but found that if they achieved the desired gain characteristic then the phase lag was too large. In 1940, Hendrik Bode, who had been studying extensions to the frequency-domain design method, showed that no definite and universal attenuation and phase shift relationship for a physical structure exists, but that there is a relationship between a given attenuation characteristic and the minimum phase shift that can be associated with it. In the same paper he adopted the point $(-1,0)$ as the critical point rather than the point $(+1,0)$ used by Nyquist, and he introduced the concept of gain and phase margins, and the gain-bandwidth limitation. Full details of Bode's work appeared in 1945 in his book *Network Analysis and Feedback Amplifier Design*.

The second important group, mechanical engineers and physicists working in the process industries in the U.S., encouraged by Ed S. Smith of the Builders Iron Foundry Company, began systematically developing a theoretical understanding of the control systems they used. They sought to establish a common terminology and tried to develop design methods. They persuaded the American Society of Mechanical Engineers to form an Industrial Instruments and Regulators Committee in 1936, thus becoming the first major professional body to form a section specifically to deal with automatic control. Several members of this loose grouping were aware of developments in Germany and in England. During this period the manufacturers of pneumatic controllers continued to improve and develop their instruments, and by 1940 field-adjustable instruments with PID control were available—for example, an improved version of the Stabilog and the Taylor Fulscope. In 1942, J.G. Ziegler and N.B. Nichols of the Taylor Instrument Companies published papers describing how to find the optimum settings for PI and PID control—the so called Ziegler-Nichols tuning rules. These were extended in the mid-1950s by Geraldine Coon (Taylor Instrument).

The third group was located in the Electrical Engineering Department of MIT and was led by Harold L. Hazen and Gordon S. Brown. They used time-domain methods based on operator techniques, began to develop the use of block diagrams, and used the differential analyzer to simulate control systems. Scholarly interchanges between MIT and the University of Manchester led to a differential analyzer being built at Manchester University and, in 1936, Douglas Hartree and Arthur Porter assisted A. Callender of ICI to use the machine to simulate an industrial control system and to derive design charts for the system.

The advent of the second world war concentrated control system work on a few specific problems. The most important of these was the aiming of anti-aircraft guns. This is a complex problem that involves the detection of the position of the airplane, calculation of its future position, and the precise control of the movement of a heavy gun. The operation required up to 14 people to carry out complicated observation and tracking tasks in a coordinated way. The design of an adequate servomechanism to control the gun position was a difficult task. It also

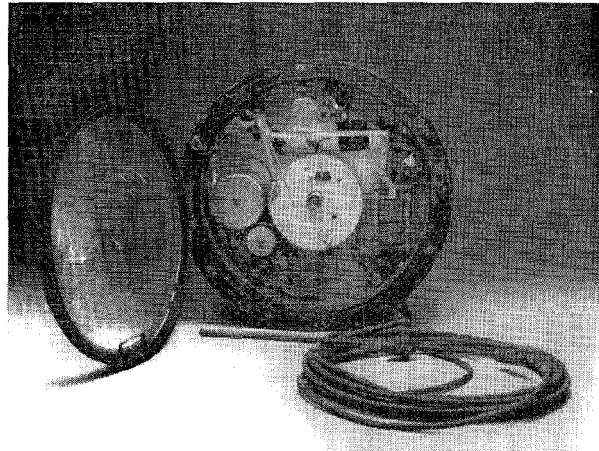


Fig. 2. Internal view of the Foxboro Stabilog circa 1936.

became clear during 1941 that the cumbersome system of relaying manually the information obtained from radar devices to the gun controllers was not adequate to combat the threat of fast aircraft and that there was a need to develop a system in which an automatic tracking radar system was directly linked to the gun director, which was in turn linked to the gun position controller.

Work on this "systems" problem brought together mechanical, electrical, and electronic engineers, and an outcome of this cross-fertilization of ideas was a recognition that neither the frequency response approach used by the communication engineers nor the time domain approach favored by the mechanical engineers were, separately, effective design approaches for servomechanisms. What was required was an approach that used the best features of each.

Work by Gordon S. Brown and his students at MIT showed how many mechanical and electrical systems could be represented and manipulated using block diagrams. Albert C. Hall showed, in 1943, that by treating the blocks as transfer functions (he used the Laplace transform approach) the system transfer locus could be drawn, and hence the Nyquist test for stability could be used. More importantly the gain and phase margin could be determined, and he introduced the use of M and N circles which enable estimates of the closed loop time domain behavior to be made. Another group working the so called Radiation Laboratory at MIT (this laboratory was concerned with developing radar systems for the detection and tracking of aircraft) designed the SCR-584 radar system, which, linked with the M9 director, was deployed in southeast England and had a high success rate against V1 rockets. The M9 director was designed by a group led by Bode and including Blackman, C.A. Lovell, and Claude Shannon, working in the Bell Telephone Laboratory. Out of the work on the SCR-584 came the Nichols chart design method, work by R.S. Phillips on noise in servomechanisms, and W. Hurewicz's work on sampled data systems. After the war, details of the work were published in the seminal book *Theory of Servomechanisms*.

The Radiation Laboratory group used phase advance circuits in the forward loop to modify the performance of their control system. Several other workers, particularly in the U.K., used minor loop feedback to modify system response and hence found

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.