4. A History of Automatic Control

C.C. Bissell

Automatic control, particularly the application of feedback, has been fundamental to the development of automation. Its origins lie in the level control, water clocks, and pneumatics/hydraulics of the ancient world. From the 17th century onwards, systems were designed for temperature control, the mechanical control of mills, and the regulation of steam engines. During the 19th century it became increasingly clear that feedback systems were prone to instability. A stability criterion was derived independently towards the end of the century by Routh in England and Hurwitz in Switzerland. The 19th century, too, saw the development of servomechanisms, first for ship steering and later for stabilization and autopilots. The invention of aircraft added (literally) a new dimension to the problem. Minorsky's theoretical analysis of ship control in the 1920s clarified the nature of three-term control, also being used for process applications by the 1930s. Based on servo and communications engineering developments of the 1930s, and driven by the need for high-performance gun control systems, the coherent body of theory known as *classical control* emerged during and just after WWII in the US, UK and elsewhere, as did cybernetics ideas. Meanwhile, an alternative approach to dynamic modelling had been developed in the USSR based on the approaches of Poincaré and Lyapunov.

Information was gradually disseminated, and *state-space* or *modern control* techniques, fuelled by Cold War demands for missile control systems, rapidly developed in both East and West. The immediate post-war period was marked by great claims for automation, but also great fears, while the digital computer opened new possibilities for automatic control.

4.1 Antiquity and the Early Modern Period

Feedback control can be said to have originated with the float valve regulators of the Hellenic and Arab worlds [4.1]. They were used by the Greeks and Arabs to control such devices as water clocks, oil lamps and wine dispensers, as well as the level of water in tanks. The precise construction of such systems is still not entirely clear, since the descriptions in the original Greek or Arabic are often vague, and lack illustrations. The best known Greek names are Ktsebios and Philon (third century BC) and Heron (first century AD) who were active in the eastern Mediterranean (Alexandria, Byzantium). The water clock tradition was continued in the Arab world as described in books by writers such as Al-Jazari (1203) and Ibn al-Sa-ati (1206), greatly influenced by the anonymous Arab author known as Pseudo-Archimedes of the ninth–tenth century AD, who makes specific reference to the Greek work of Heron and Philon. Float regulators in the tradition of Heron were also constructed by the three brothers Banu Musa in Baghdad in the ninth century AD.

The float valve level regulator does not appear to have spread to medieval Europe, even though translations existed of some of the classical texts by the above writers. It seems rather to have been reinvented during the industrial revolution, appearing in England, for example, in the 18th century. The first independent European feedback system was the temperature regulator of Cornelius Drebbel (1572–1633). Drebbel spent most of his professional career at the courts of James I and Charles I of England and Rudolf II in Prague. Drebbel himself left no written records, but a number of contemporary descriptions survive of his invention. Essentially an alcohol (or other) thermometer was used to operate a valve controlling a furnace flue, and hence the temperature of an enclosure [4.2]. The device included screws to alter what we would now call the set point.

If level and temperature regulation were two of the major precursors of modern control systems, then

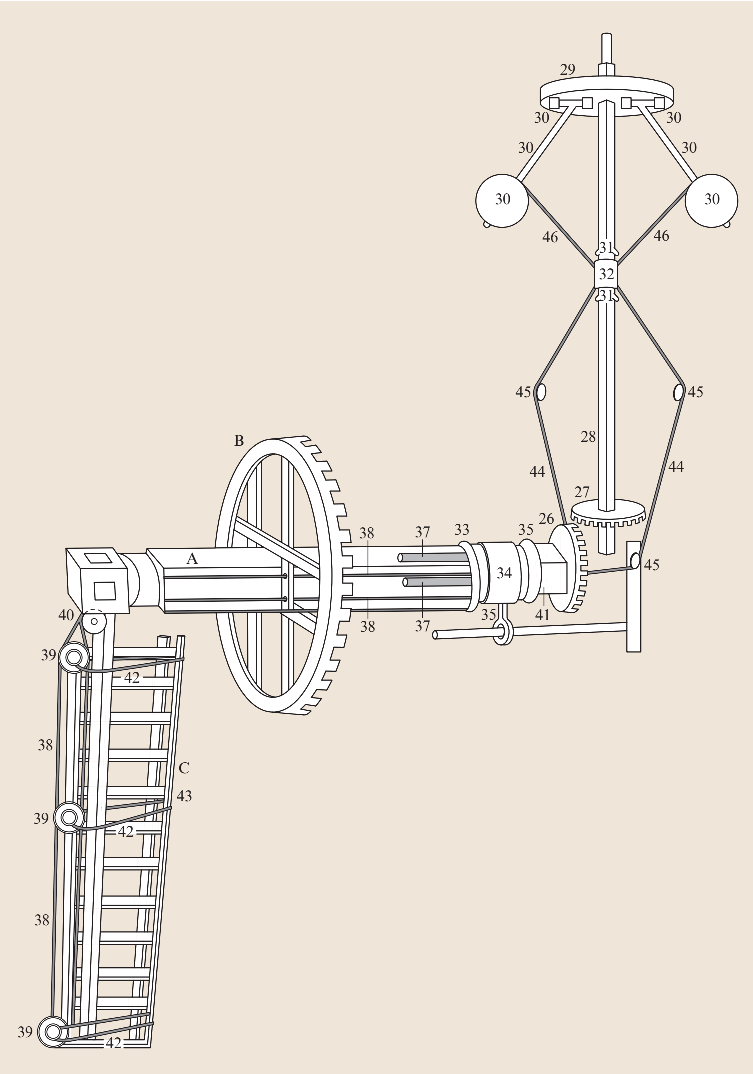

Fig. 4.1 Mead's speed regulator (after [4.1])

a number of devices designed for use with windmills pointed the way towards more sophisticated devices. During the 18th century the mill fantail was developed both to keep the mill sails directed into the wind and to automatically vary the angle of attack, so as to avoid excessive speeds in high winds. Another important device was the lift-tenter. Millstones have a tendency to separate as the speed of rotation increases, thus impairing the quality of flour. A number of techniques were developed to sense the speed and hence produce a restoring force to press the millstones closer together. Of these, perhaps the most important were *Thomas Mead*'s devices [4.3], which used a centrifugal pendulum to sense the speed and – in some applications – also to provide feedback, hence pointing the way to the centrifugal governor.

The first steam engines were the reciprocating engines developed for driving water pumps; James Watt's rotary engines were sold only from the early 1780s. But it took until the end of the decade for the centrifugal governor to be applied to the machine, following a visit by Watt's collaborator, Matthew Boulton, to the

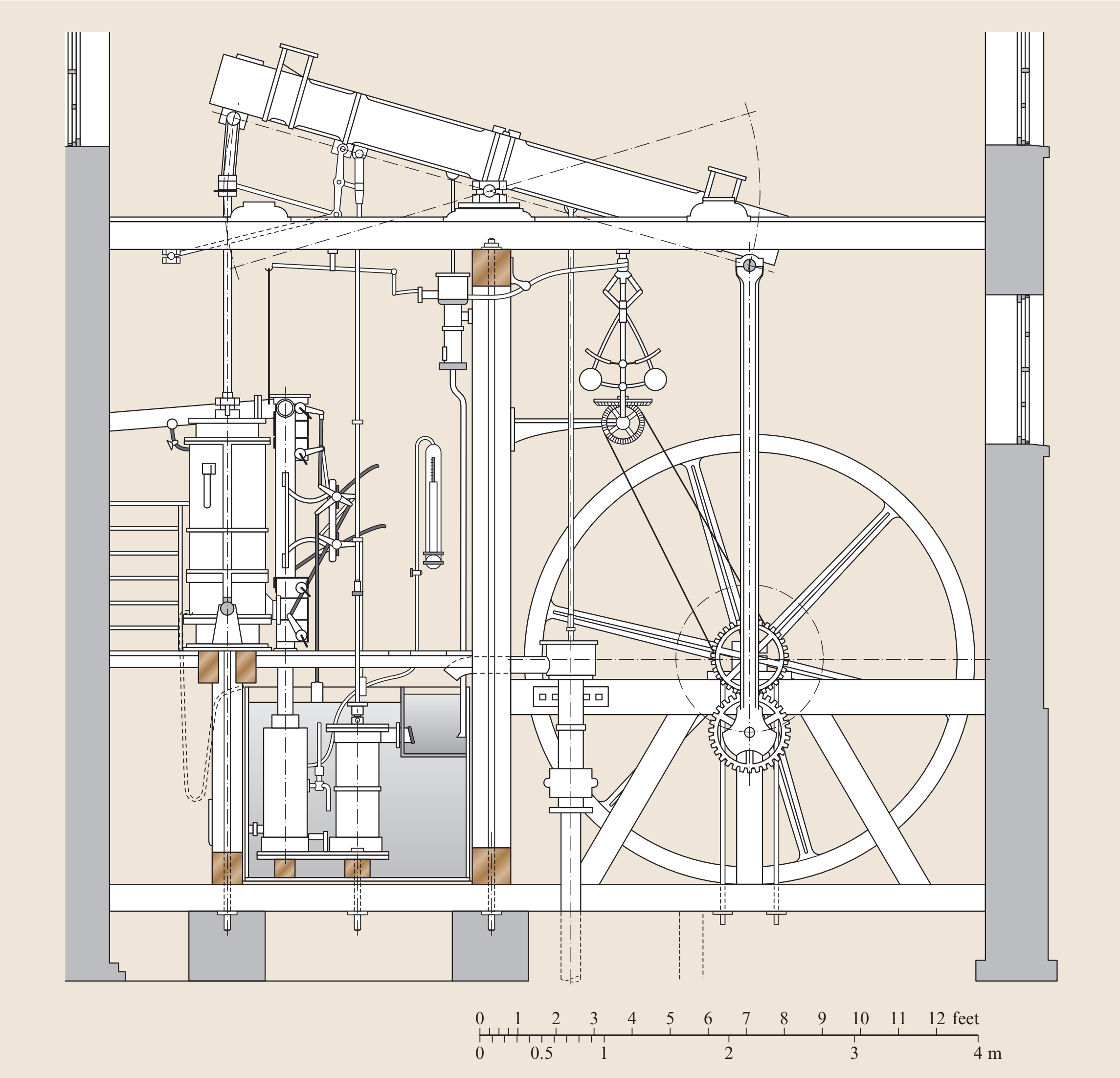

Fig. 4.2 Boulton & Watt steam engine with centrifugal governor (after [4.1])

Albion Mill in London where he saw a lift-tenter in action under the control of a centrifugal pendulum. Boulton and Watt did not attempt to patent the device (which, as noted above, had essentially already been patented by Mead) but they did try unsuccessfully to keep it secret. It was first copied in 1793 and spread throughout England over the next ten years [4.4].

4.2 Stability Analysis in the 19th Century

With the spread of the centrifugal governor in the early 19th century a number of major problems became apparent. First, because of the absence of integral action, the governor could not remove offset: in the terminology of the time it could not *regulate* but only *moderate*. Second, its response to a change in load was slow. And thirdly, (nonlinear) frictional forces in the mechanism could lead to *hunting* (limit cycling). A number of attempts were made to overcome these problems: for example, the Siemens chronometric governor effectively introduced integral action through differential gearing, as well as mechanical amplification. Other approaches to the design of an *isochronous* governor (one with no offset) were based on ingenious mechanical constructions, but often encountered problems of stability.

Nevertheless the 19th century saw steady progress in the development of practical governors for steam engines and hydraulic turbines, including spring-loaded designs (which could be made much smaller, and operate at higher speeds) and relay (indirect-acting) governors [4.6]. By the end of the century governors of various sizes and designs were available for effective regulation in a range of applications, and a number of graphical techniques existed for steady-state design. Few engineers were concerned with the analysis of the dynamics of a feedback system.

In parallel with the developments in the engineering sector a number of eminent British scientists became interested in governors in order to keep a telescope directed at a particular star as the Earth rotated. A formal analysis of the dynamics of such a system by *George Bidell Airy*, Astronomer Royal, in 1840 [4.7] clearly demonstrated the propensity of such a feedback system to become unstable. In 1868 James Clerk Maxwell analyzed governor dynamics, prompted by an electrical experiment in which the speed of rotation of a coil had to be held constant. His resulting classic paper *On governors* [4.8] was received by the Royal Society on 20 February. Maxwell derived a third-order linear model and the correct conditions for stability in terms of the coefficients of the characteristic equation. Unable to derive a solution for higher-order models, he expressed the hope that the question would gain the attention of mathematicians. In 1875 the subject for the Cambridge University Adams Prize in mathematics was set as *The criterion of dynamical stability*. One of the examiners was Maxwell himself (prizewinner in 1857) and the 1875 prize (awarded in 1877) was won by Edward James Routh. *Routh* had been interested in dynamical stability for several years, and had already obtained a solution for a fifth-order system. In the published paper [4.9] we find derived the Routh version of the renowned Routh–Hurwitz stability criterion.

Related, independent work was being carried out in continental Europe at about the same time [4.5]. A summary of the work of I.A. Vyshnegradskii in St. Petersburg appeared in the French *Comptes Rendus de l'Academie des Sciences* in 1876, with the full version appearing in Russian and German in 1877, and in French in 1878/79. Vyshnegradskii (generally transliterated at the time as Wischnegradski) transformed a third-order differential equation model of a steam en-

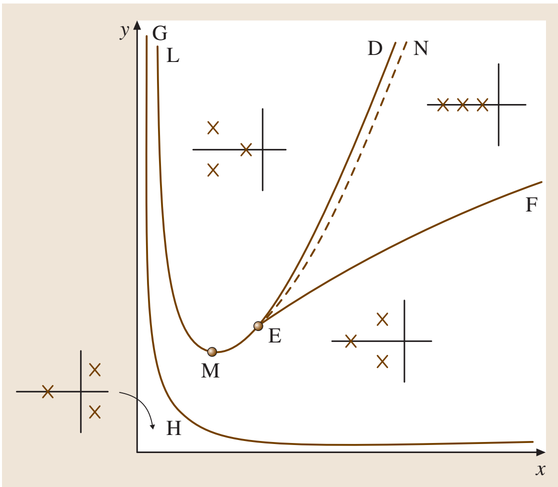

Fig. 4.3 Vyshnegradskii's stability diagram with modern pole positions (after [4.5])

gine with governor into a standard form

$$\varphi^3 + x\varphi^2 + y\varphi + 1 = 0\,,$$

where x and y became known as the Vyshnegradskii parameters. He then showed that a point in the x–y plane defined the nature of the system transient response. Figure 4.3 shows the diagram drawn by Vyshnegradskii, to which typical pole constellations for various regions in the plane have been added.

In 1893 Aurel Boreslav Stodola at the Federal Polytechnic, Zurich, studied the dynamics of a high-pressure hydraulic turbine, and used Vyshnegradskii's method to assess the stability of a third-order model. A more realistic model, however, was seventh-order, and Stodola posed the general problem to a mathematician colleague *Adolf Hurwitz*, who very soon came up with his version of the Routh–Hurwitz criterion [4.10]. The two versions were shown to be identical by *Enrico Bompiani* in 1911 [4.11].

At the beginning of the 20th century the first general textbooks on the regulation of prime movers appeared in a number of European languages [4.12, 13]. One of the most influential was *Tolle*'s *Regelung der Kraftmaschine*, which went through three editions between 1905 and 1922 [4.14]. The later editions included the Hurwitz stability criterion.

4.3 Ship, Aircraft and Industrial Control Before WWII

The first ship steering engines incorporating feedback appeared in the middle of the 19th century. In 1873 Jean Joseph Léon Farcot published a book on servomotors in which he not only described the various designs developed in the family firm, but also gave an account of the general principles of position control. Another important maritime application of feedback control was in gun turret operation, and hydraulics were also extensively developed for transmission systems. Torpedoes, too, used increasingly sophisticated feedback systems for depth control – including, by the end of the century – gyroscopic action.

During the first decades of the 20th century gyroscopes were increasingly used for ship stabilization and autopilots. Elmer Sperry pioneered the *active stabilizer*, the gyrocompass, and the gyroscope autopilot, filing various patents over the period 1907–1914. Sperry's autopilot was a sophisticated device: an inner loop controlled an electric motor which operated the steering engine, while an outer loop used a gyrocompass to sense the heading. Sperry also designed an *anticipator* to replicate the way in which an experienced helmsman would *meet* the helm (to prevent oversteering); the anticipator was, in fact, a type of adaptive control [4.16].

Sperry and his son Lawrence also designed aircraft autostabilizers over the same period, with the added complexity of three-dimensional control. *Bennett* de-

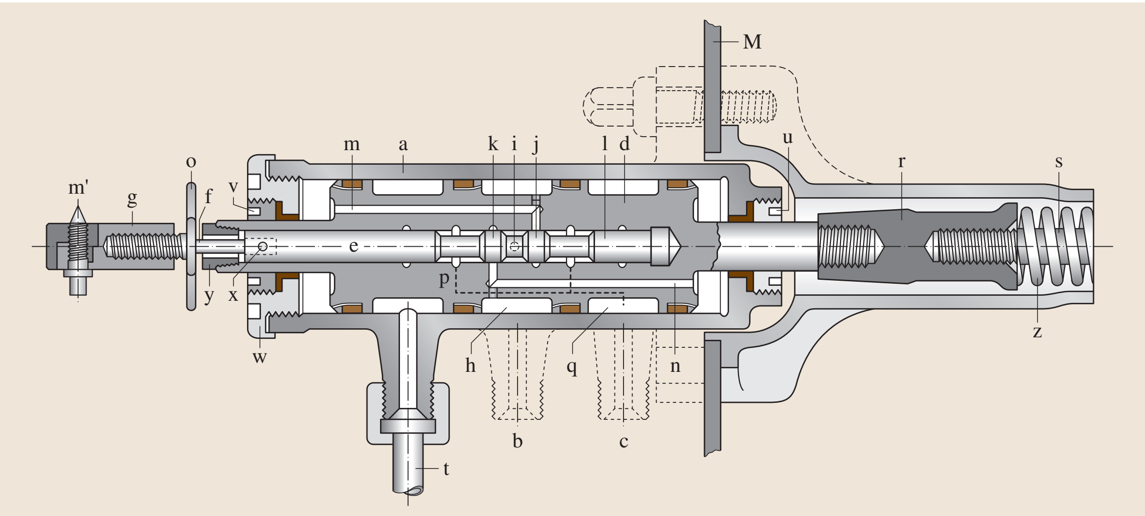

Fig. 4.4 Torpedo servomotor as fitted to Whitehead torpedoes around 1900 (after [4.15])

scribes the system used in an acclaimed demonstration in Paris in 1914 [4.17]

> *For this system the Sperrys used four gyroscopes mounted to form a stabilized reference platform; a train of electrical, mechanical and pneumatic components detected the position of the aircraft relative to the platform and applied correction signals to the aircraft control surfaces. The stabilizer operated for both pitch and roll [. . .] The system was normally adjusted to give an approximately deadbeat response to a step disturbance. The incorporation of derivative action [. . .] was based on Sperry's intuitive understanding of the behaviour of the system, not on any theoretical foundations. The system was also adaptive [. . .] adjusting the gain to match the speed of the aircraft.*

Significant technological advances in both ship and aircraft stabilization took place over the next two decades, and by the mid 1930s a number of airlines were using Sperry autopilots for long-distance flights. However, apart from the stability analyses discussed in Sect. 4.2 above, which were not widely known at this time, there was little theoretical investigation of such feedback control systems. One of the earliest

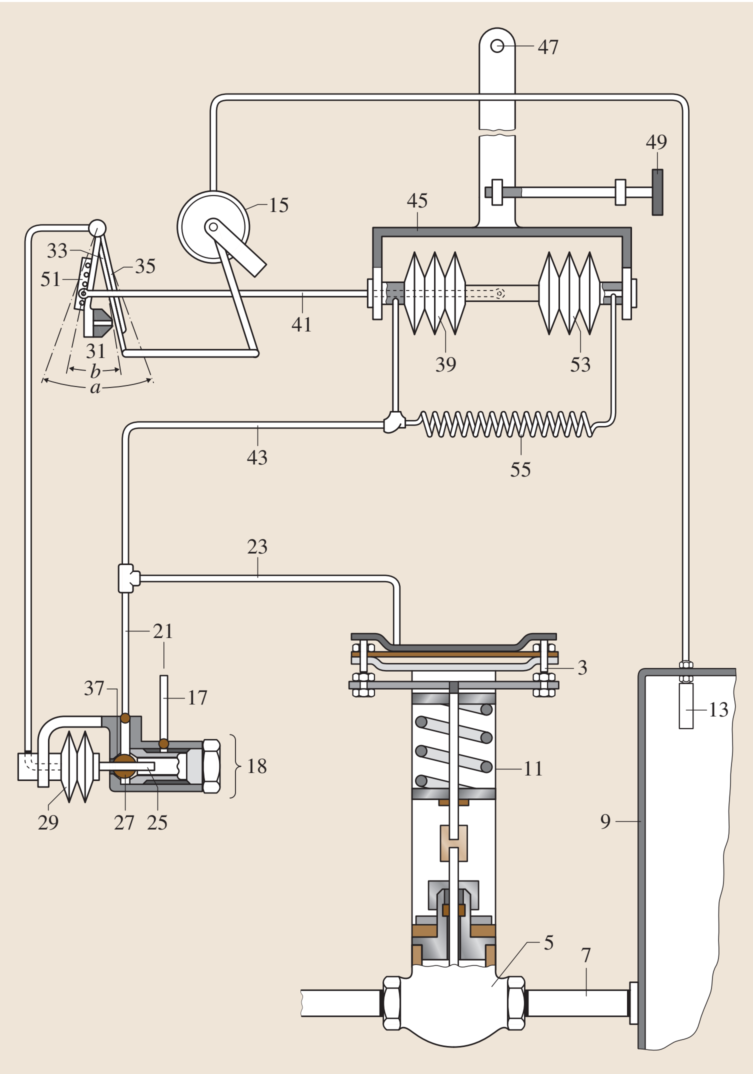

Fig. 4.5 The *Stabilog*, a pneumatic controller providing proportional and integral action [4.18]

significant studies was carried out by *Nicholas Minorsky*, published in 1922 [4.19]. Minorsky was born in Russia in 1885 (his knowledge of Russian proved to be important to the West much later). During service with the Russian Navy he studied the ship steering problem and, following his emigration to the USA in 1918, he made the first theoretical analysis of automatic ship steering. This study clearly identified the way that control action should be employed: although Minorsky did not use the terms in the modern sense, he recommended an appropriate combination of proportional, derivative and integral action. Minorsky's work was not widely disseminated, however. Although he gave a good theoretical basis for closed loop control, he was writing in an age of heroic invention, when intuition and practical experience were much more important for engineering practice than theoretical analysis.

Important technological developments were also being made in other sectors during the first few decades of the 20th century, although again there was little theoretical underpinning. The electric power industry brought demands for voltage and frequency regulation; many processes using driven rollers required accurate speed control; and considerable work was carried out in a number of countries on systems for the accurate pointing of guns for naval and anti-aircraft gunnery. In the process industries, measuring instruments and pneumatic controllers of increasing sophistication were developed. Mason's *Stabilog*, patented in 1933, included integral as well as proportional action, and by the end of the decade three-term controllers were available that also included *preact* or derivative control. Theoretical progress was slow, however, until the advances made in electronics and telecommunications in the 1920s and 30s were translated into the control field during WWII.

4.4 Electronics, Feedback and Mathematical Analysis

The rapid spread of telegraphy and then telephony from the mid 19th century onwards prompted a great deal of theoretical investigation into the behaviour of electric circuits. *Oliver Heaviside* published papers on his operational calculus over a number of years from 1888 onwards [4.20], but although his techniques produced valid results for the transient response of electrical networks, he was fiercely criticized by contemporary mathematicians for his lack of rigour, and ultimately he was blackballed by the establishment. It was not until the second decade of the 20th century that Bromwich, Carson and others made the link between Heaviside's operational calculus and Fourier methods, and thus proved the validity of Heaviside's techniques [4.21].

The first three decades of the 20th century saw important analyses of circuit and filter design, particularly in the USA and Germany. *Harry Nyquist* and Karl Küpfmüller were two of the first to consider the problem of the maximum transmission rate of telegraph signals, as well as the notion of *information* in telecommunications, and both went on to analyze the general stability problem of a feedback circuit [4.22]. In 1928 Küpfmüller analyzed the dynamics of an automatic gain control electronic circuit using feedback. He appreciated the dynamics of the feedback system, but his integral equation approach resulted only in a approximations and design diagrams, rather than a rigorous stability criterion. At about the same time in the USA, Harold Black was designing feedback amplifiers for transcontinental telephony. In a famous epiphany on the Hudson River ferry in August 1927 he realized that negative feedback could reduce distortion at the cost of reducing overall gain. Black passed on the problem of the stability of such a feedback loop to his Bell Labs colleague *Harry Nyquist*, who published his celebrated frequency-domain encirclement criterion in 1932 [4.23]. Nyquist demonstrated, using results derived by Cauchy, that the key to stability is whether or not the open loop frequency response locus in the complex plane encircles (in Nyquist's original convention) the point $1+\mathrm{i}0$. One of the great advantages of this approach is that no analytical form of the open loop frequency response is required: a set of measured data points can be plotted without the need for a mathematical model. Another advantage is that, unlike the Routh–Hurwitz criterion, an assessment of the transient response can be made directly from the Nyquist plot in terms of gain and phase margins (how close the locus approaches the critical point).

Black's 1934 paper reporting his contribution to the development of the negative feedback amplifier included what was to become the standard closed-loop analysis in the frequency domain [4.24].

The third key contributor to the analysis of feedback in electronic systems at Bell Labs was *Hendrik*

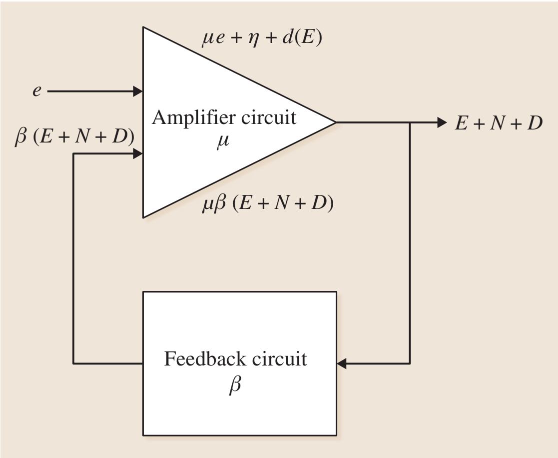

Fig. 4.6 Black's feedback amplifier (after [4.24])

Bode who worked on equalizers from the mid 1930s, and who demonstrated that attenuation and phase shift were related in any realizable circuit [4.25]. The dream of telephone engineers to build circuits with fast cut-off and low phase shift was indeed only a dream. It was Bode who introduced the notions of gain and phase margins, and redrew the Nyquist plot in its now conventional form with the critical point at $-1+\mathrm{i}0$. He also introduced the famous straight-line approximations to frequency response curves of linear systems plotted on log–log axes. *Bode* presented his methods in a classic text published immediately after the war [4.26].

If the work of the communications engineers was one major precursor of classical control, then the other was the development of high-performance servos in the 1930s. The need for such servos was generated by the increasing use of analogue simulators, such as network analysers for the electrical power industry and differential analysers for a wide range of problems. By the early 1930s six-integrator differential analysers were in operation at various locations in the USA and the UK. A major centre of innovation was MIT, where Vannevar Bush, Norbert Wiener and Harold Hazen had all contributed to design. In 1934 *Hazen* summarized the developments of the previous years in *The theory of servomechanisms* [4.27]. He adopted normalized curves, and parameters such as time constant and damping factor, to characterize servo-response, but he did not given any stability analysis: although he appears to have been aware of Nyquists's work, he (like almost all his contemporaries) does not appear to have appreciated the close relationship between a feedback servomechanism and a feedback amplifier.

The 1930s American work gradually became known elsewhere. There is ample evidence from prewar USSR, Germany and France that, for example, Nyquist's results were known – if not widely disseminated. In 1940, for example, *Leonhard* published a book on automatic control in which he introduced the inverse Nyquist plot [4.28], and in the same year a conference was held in Moscow during which a number of Western results in automatic control were presented and discussed [4.29]. Also in Russia, a great deal of work was being carried out on nonlinear dynamics, using an approach developed from the methods of Poincaré and Lyapunov at the turn of the century [4.30]. Such approaches, however, were not widely known outside Russia until after the war.

4.5 WWII and *Classical* Control: Infrastructure

Notwithstanding the major strides identified in the previous subsections, it was during WWII that a discipline of feedback control began to emerge, using a range of design and analysis techniques to implement high-performance systems, especially those for the control of anti-aircraft weapons. In particular, WWII saw the coming together of engineers from a range of disciplines – electrical and electronic engineering, mechanical engineering, mathematics – and the subsequent realisation that a common framework could be applied to all the various elements of a complex control system in order to achieve the desired result [4.18, 31].

The so-called fire control problem was one of the major issues in military research and development at the end of the 1930s. While not a new problem, the increasing importance of aerial warfare meant that the control of anti-aircraft weapons took on a new significance. Under manual control, aircraft were detected by radar, range was measured, prediction of the aircraft position at the arrival of the shell was computed, guns were aimed and fired. A typical system could involve up to 14 operators. Clearly, automation of the process was highly desirable, and achieving this was to require detailed research into such matters as the dynamics of the servomechanisms driving the gun aiming, the design of controllers, and the statistics of tracking aircraft possibly taking evasive action.

Government, industry and academia collaborated closely in the US, and three research laboratories were of prime importance. The Servomechanisms Laboratory at MIT brought together Brown, Hall, Forrester and others in projects that developed frequency-domain methods for control loop design for high-performance servos. Particularly close links were maintained with Sperry, a company with a strong track record in guidance systems, as indicated above. Meanwhile, at MIT's Radiation Laboratory – best known, perhaps, for its work on radar and long-distance navigation – researchers such as James, Nichols and Phillips worked on the further development of design techniques for auto-track radar for AA gun control. And the third institution of seminal importance for fire-control development was Bell Labs, where great names such as Bode, Shannon and Weaver – in collaboration with Wiener and Bigelow at MIT – attacked a number of outstanding problems, including the theory of smoothing and prediction for gun aiming. By the end of the war, most of the techniques of what came to be called classical control had been elaborated in these laboratories, and a whole series of papers and textbooks appeared in the late 1940s presenting this new discipline to the wider engineering community [4.32].

Support for control systems development in the United States has been well documented [4.18, 31]. The National Defence Research Committee (NDRC) was established in 1940 and incorporated into the Office of Scientific Research and Development (OSRD) the following year. Under the directorship of Vannevar Bush the new bodies tackled anti-aircraft measures, and thus the servo problem, as a major priority. Section D of the NDRC, devoted to Detection, Controls and Instruments was the most important for the development of feedback control. Following the establishment of the OSRD the NDRC was reorganised into divisions, and Division 7, Fire Control, under the overall direction of Harold Hazen, covered the subdivisions: ground-based anti-aircraft fire control; airborne fire control systems; servomechanisms and data transmission; optical rangefinders; fire control analysis; and navy fire control with radar.

Turning to the United Kingdom, by the outbreak of WWII various military research stations were highly active in such areas as radar and gun laying, and there were also close links between government bodies and industrial companies such as Metropolitan–Vickers, British Thomson–Houston, and others. Nevertheless, it is true to say that overall coordination was not as effective as in the USA. A body that contributed significantly to the dissemination of theoretical developments and other research into feedback control systems in the UK was the so called Servo-Panel. Originally established informally in 1942 as the result of an initiative of Solomon (head of a special radar group at Malvern), it acted rather as a *learned society* with approximately monthly meetings from May 1942 to August 1945. Towards the end of the war meetings included contributions from the US.

Germany developed successful control systems for civil and military applications both before and during the war (torpedo and flight control, for example). The period 1938–1941 was particularly important for the development of missile guidance systems. The test and development centre at Peenemünde on the Baltic coast had been set up in early 1936, and work on guidance and control saw the involvement of industry, the government and universities. However, there does not appear to have been any significant national coordination of R&D in the control field in Germany, and little development of high-performance servos as there was in the US and the UK. When we turn to the German situation outside the military context, however, we find a rather remarkable awareness of control and even cybernetics. In 1939 the *Verein Deutscher Ingenieure*, one of the two major German engineers' associations, set up a specialist committee on control engineering. As early as October 1940 the chair of this body *Herman Schmidt* gave a talk covering control engineering and its relationship with economics, social sciences and cultural aspects [4.33]. Rather remarkably, this committee continued to meet during the war years, and issued a report in 1944 concerning primarily control concepts and terminology, but also considering many of the fundamental issues of the emerging discipline.

The Soviet Union saw a great deal of prewar interest in control, mainly for industrial applications in the context of five-year plans for the Soviet command economy. Developments in the USSR have received little attention in English-language accounts of the history of the discipline apart from a few isolated papers. It is noteworthy that the *Kommissiya Telemekhaniki i Avtomatiki* (KTA) was founded in 1934, and the *Institut Avtomatiki i Telemekhaniki* (IAT) in 1939 (both under the auspices of the Soviet Academy of Sciences, which controlled scientific research through its network of institutes). The KTA corresponded with numerous western manufacturers of control equipment in the mid 1930s and translated a number articles from western journals. The early days of the IAT were marred, however, by the *Shchipanov affair*, a classic Soviet attack on a researcher for *pseudo-science*, which detracted from technical work for a considerable period of time [4.34]. The other major Russian centre of research related to control theory in the 1930s and 1940s (if not for practical applications) was the University of Gorkii (now Nizhnii Novgorod), where Aleksandr Andronov and colleagues had established a centre for the study of nonlinear dynamics during the 1930s [4.35]. Andronov was in regular contact with Moscow during the 1940s, and presented the emerging control theory there – both the nonlinear research at Gorkii and developments in the UK and USA. Nevertheless, there appears to have been no co-ordinated wartime work on control engineering in the USSR, and the IAT in Moscow was evacuated when the capital came under threat. However, there does seem to have been an emerging control community in Moscow, Nizhnii Novgorod and Leningrad, and Russian workers were extremely well-informed about the open literature in the West.

4.6 WWII and *Classical* Control: Theory

Design techniques for servomechanisms began to be developed in the USA from the late 1930s onwards. In 1940 Gordon S. Brown and colleagues at MIT analyzed the transient response of a closed loop system in detail, introducing the *system operator* $1/(1+\text{open loop})$ as functions of the Heaviside differential operator p. By the end of 1940 contracts were being drawn up between the NDRC and MIT for a range of servo projects. One of the most significant contributors was *Albert Hall*, who developed classic frequency-response methods as part of his doctoral thesis, presented in 1943 and published initially as a confidential document [4.37] and then in the open literature after the war [4.36]. Hall derived the frequency response of a unity feedback servo as $KG(\mathrm{i}\omega)/[1+KG(\mathrm{i}\omega)]$, applied the Nyquist criterion, and introduced a new way of plotting system response that he called M-circles, which were later to inspire the Nichols Chart. As *Bennett* describes it [4.38]

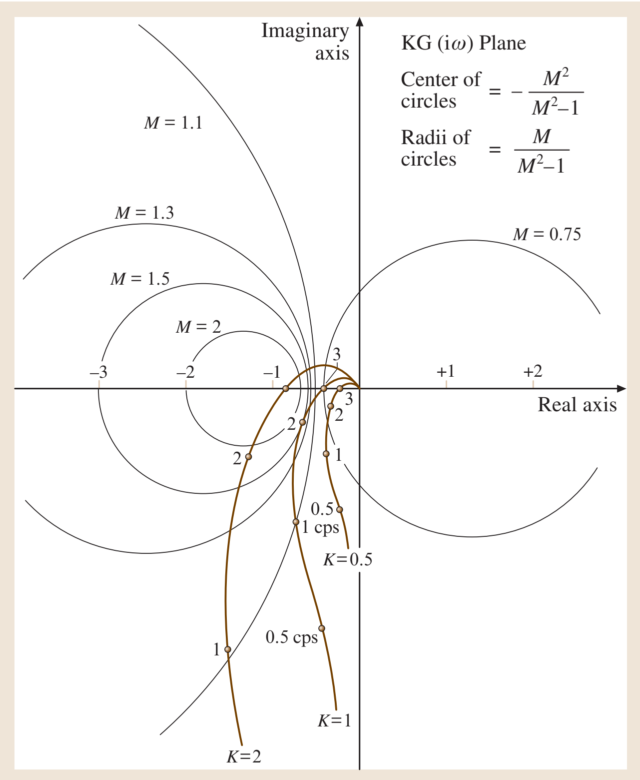

Fig. 4.7 Hall's M-circles (after [4.36])

> *Hall was trying to design servosystems which were stable, had a high natural frequency, and high damping. [...] He needed a method of determining, from the transfer locus, the value of K that would give the desired amplitude ratio. As an aid to finding the value of K he superimposed on the polar plot curves of constant magnitude of the amplitude ratio. These curves turned out to be circles... By plotting the response locus on transparent paper, or by using an overlay of M-circles printed on transparent paper, the need to draw M-circles was obviated...*

A second MIT group, known as the Radiation Laboratory (or RadLab) was working on auto-track radar systems. Work in this group was described after the war in [4.39]; one of the major innovations was the introduction of the Nichols chart, similar to Hall's M-circles, but using the more convenient decibel measure of amplitude ratio that turned the circles into a rather different geometrical form.

The third US group consisted of those looking at smoothing and prediction for anti-aircraft weapons – most notably Wiener and Bigelow at MIT together with others, including Bode and Shannon, at Bell Labs. This work involved the application of correlation techniques to the statistics of aircraft motion. Although the prototype Wiener predictor was unsuccessful in attempts at practical application in the early 1940s, the general approach proved to be seminal for later developments.

Formal techniques in the United Kingdom were not so advanced. Arnold Tustin at Metropolitan–Vickers (Metro–Vick) worked on gun control from the late 1930s, but engineers had little appreciation of dynamics. Although they used harmonic response plots they appeared to have been unaware of the Nyquist criterion until well into the 1940s [4.40]. Other key researchers in the UK included *Whitely*, who proposed using the inverse Nyquist diagram as early as 1942, and introduced his *standard forms* for the design of various categories of servosystem [4.41]. In Germany, Winfried Oppelt, Hans Sartorius and Rudolf Oldenbourg were also coming to related conclusions about closed-loop design independently of allied research [4.42, 43].

The basics of sampled-data control were also developed independently during the war in several countries. The z-transform in all but name was described in a chapter by *Hurewizc* in [4.39]. Tustin in the UK developed the bilinear transformation for time series models, while Oldenbourg and Sartorius also used difference equations to model such systems.

From 1944 onwards the design techniques developed during the hostilities were made widely available in an explosion of research papers and text books – not only from the USA and the UK, but also from Germany and the USSR. Towards the end of the decade perhaps the final element in the classical control toolbox was added – *Evans*' root locus technique, which enabled plots of changing pole position as a function of loop gain to be easily sketched [4.44]. But a radically different approach was already waiting in the wings.

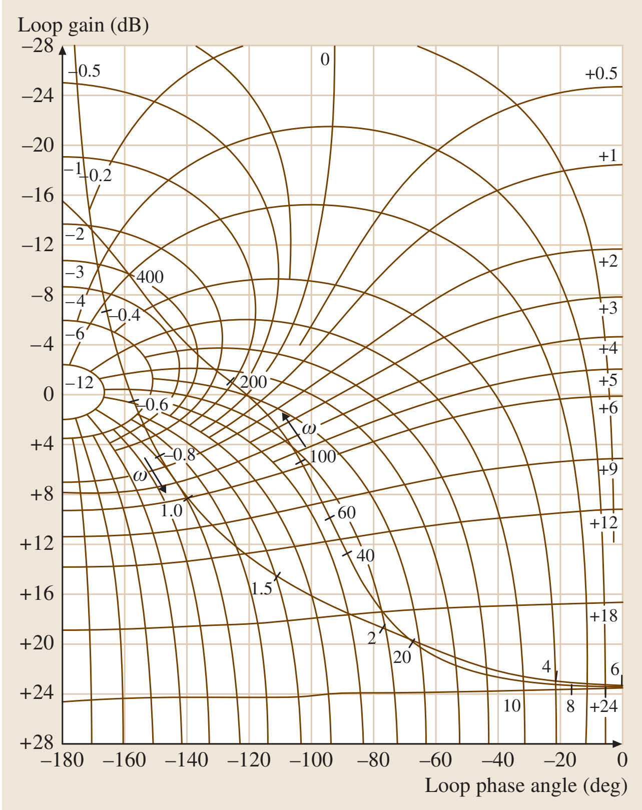

Fig. 4.8 Nichols Chart (after [4.38])

4.7 The Emergence of *Modern* Control Theory

The *modern* or *state space* approach to control was ultimately derived from original work by Poincaré and Lyapunov at the end of the 19th century. As noted above, Russians had continued developments along these lines, particularly during the 1920s and 1930s in centres of excellence in Moscow and Gorkii (now Nizhnii Novgorod). Russian work of the 1930s filtered slowly through to the West [4.45], but it was only in the post war period, and particularly with the introduction of cover-to-cover translations of the major Soviet journals, that researchers in the USA and elsewhere became familiar with Soviet work. But phase plane approaches had already been adopted by Western control engineers. One of the first was *Leroy MacColl* in his early textbook [4.46].

The cold war requirements of control engineering centred on the control of ballistic objects for aerospace applications. Detailed and accurate mathematical models, both linear and nonlinear, could be obtained, and the classical techniques of frequency response and root locus – essentially approximations – were increasingly replaced by methods designed to optimize some measure of performance such as minimizing trajectory time or fuel consumption. Higher-order models were expressed as a set of first order equations in terms of the state variables. The state variables allowed for a more sophisticated representation of dynamic behaviour than the classical single-input, single-output system modelled by a differential equation, and were suitable for multi-variable problems. In general, we have in matrix form

$$\boldsymbol{x} = \mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{u} ,$$

$$\boldsymbol{y} = \mathbf{C}\boldsymbol{x} ,$$

where $\boldsymbol{x}$ are the state variables, $\boldsymbol{u}$ the inputs and $\boldsymbol{y}$ the outputs.

Automatic control developments in the late 1940s and 1950s were greatly assisted by changes in the engineering professional bodies and a series of international conferences [4.47]. In the USA both the American Society of Mechanical Engineers and the American Institute of Electrical Engineers made various changes to their structure to reflect the growing importance of servomechanisms and feedback control. In the UK similar changes took place in the British professional bodies, most notably the Institution of Electrical Engineers, but also the Institute of Measurement and Control and the mechanical and chemical engineering bodies. The first conferences on the subject appeared in the late 1940s in London and New York, but the first truly international conference was held in Cranfield, UK in 1951. This was followed by a number of others, the most influential of which was the Heidelberg event of September 1956, organized by the joint control committee of the two major German engineering bodies, the VDE and VDI. The establishment of the International Federation of Automatic Control followed in 1957 with its first conference in Moscow in 1960 [4.48]. The Moscow conference was perhaps most remarkable for Kalman's paper *On the general theory of control systems* which identified the duality between multivariable feedback control and multivariable feedback filtering and which was seminal for the development of optimal control.

The late 1950s and early 1960s saw the publication of a number of other important works on dynamic programming and optimal control, of which can be singled out those by *Bellman* [4.49], *Kalman* [4.50–52] and *Pontryagin* and colleagues [4.53].

4.8 The Digital Computer

The introduction of digital technologies in the late 1950s brought enormous changes to automatic control. Control engineering had long been associated with computing devices – as noted above, a driving force for the development of servos was for applications in analogue computing. But the great change with the introduction of digital computers was that ultimately the approximate methods of frequency response or root locus design, developed explicitly to avoid computation, could be replaced by techniques in which accurate computation played a vital role.

There is some debate about the first application of digital computers to process control, but certainly the introduction of computer control at the Texaco Port Arthur (Texas) refinery in 1959 and the Monsanto ammonia plant at Luling (Louisiana) the following year are two of the earliest [4.54]. The earliest systems were supervisory systems, in which individual loops were controlled by conventional electrical, pneumatic or hydraulic controllers, but monitored and optimized by computer. Specialized process control computers followed in the second half of the 1960s, offering direct digital control (DDC) as well as supervisory control. In DDC the computer itself implements a discrete form of a control algorithm such as three-term control or other procedure. Such systems were expensive, however, and also suffered many problems with programming, and were soon superseded by the much cheaper minicomputers of the early 1970s, most notably the DEC PDP-11. But, as in so many other areas, it was the microprocessor that had the greatest effect. Microprocessor-based digital controllers were soon developed that were compact, reliable, included a wide selection of control algorithms, had good communications with supervisory computers, and comparatively easy to use programming and diagnostic tools via an effective operator interface. Microprocessors could also easily be built into specific pieces of equipment, such as robot arms, to provide dedicated position control, for example.

A development often neglected in the history of automatic control is the programmable logic controller (PLC). PLCs were developed to replace individual relays used for sequential (and combinational) logic control in various industrial sectors. Early plugboard devices appeared in the mid 1960s, but the first PLC proper was probably the Modicon, developed for General Motors to replace electromechanical relays in

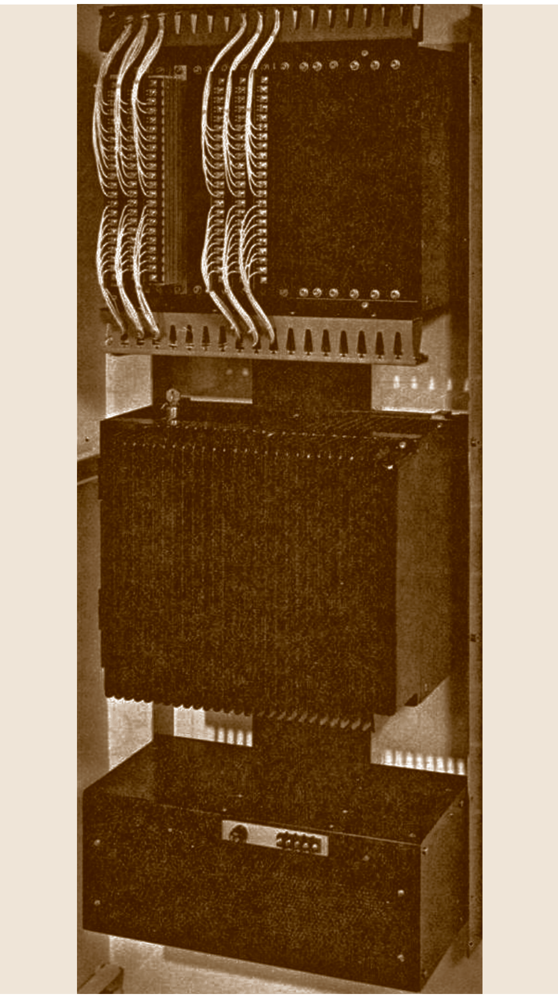

Fig. 4.9 The Modicon 084 PLC

automotive component production. Modern PLCs offer a wide range of control options, including conventional closed loop control algorithms such as PID as well as the logic functions. In spite of the rise of the ruggedized PCs in many industrial applications, PLCs are still widely used owing to their reliability and familiarity.

Digital computers also made it possible to implement the more advanced control techniques that were being developed in the 1960s and 1970s [4.55]. In adaptive control the algorithm is modified according to circumstances. Adaptive control has a long history: so called gain scheduling, for example, when the gain of a controller is varied according to some measured parameter, was used well before the digital computer. (The classic example is in flight control, where the altitude affects aircraft dynamics, and needs therefore to be taken into account when setting gain.) Digital adaptive control, however, offers much greater possibilities for:

1. Identification of relevant system parameters
2. Making decisions about the required modifications to the control algorithm
3. Implementing the changes

Optimal and robust techniques too, were developed, the most celebrated perhaps being the linear-quadratic-Gaussian (LQG) and H_∞ approaches from the 1960s onwards. Without digital computers these techniques, that attempt to optimize system rejection of disturbances (according to some measure of behaviour) while at the same time being resistant to errors in the model, would simply be mathematical curiosities [4.56].

A very different approach to control rendered possible by modern computers is to move away from purely mathematic models of system behaviour and controller algorithms. In fuzzy control, for example, control action is based on a set of rules expressed in terms of *fuzzy* variables. For example

IF the speed is "high"
AND the distance to final stop is "short"
THEN apply brakes "firmly".

The fuzzy variables *high*, *short* and *firmly* can be translated by means of an appropriate computer program into effective control for, in this case, a train. Related techniques include *learning control* and *knowledge-based control*. In the former, the control system can *learn* about its environment using artificial intelligence techniques (AI) and modify its behaviour accordingly. In the latter, a range of AI techniques are applied to reasoning about the situation so as to provide appropriate control action.

4.9 The Socio-Technological Context Since 1945

This short survey of the history of automatic control has concentrated on technological and, to some extent, institutional developments. A full social history of automatic control has yet to be written, although there are detailed studies of certain aspects. Here I shall merely indicate some major trends since WWII.

The wartime developments, both in engineering and in areas such as operations research, pointed the way towards the design and management af large-scale, complex, projects. Some of those involved in the wartime research were already thinking on a much larger scale. As early as 1949, in some rather prescient remarks at an ASME meeting in the fall of that year, *Gordon Brown* and *Duncan Campbell* said [4.57–59]

> *We have in mind more a philosophic evaluation of systems which might lead to the improvement of product quality, to better coordination of plant operation, to a clarification of the economics related to new plant design, and to the safe operation of plants in our composite social-industrial community. [...] The conservation of raw materials used in a process often prompts reconsideration of control. The expenditure of power or energy in product manufacture is another important factor related to control. The protection of health of the population adjacent to large industrial areas against atmospheric poisoning and water-stream pollution is a sufficiently serious problem to keep us constantly alert for advances in the study and technique of automatic control, not only because of the human aspect, but because of the economy aspect.*

Many saw the new technologies, and the prospects of automation, as bringing great benefits to society; others were more negative. *Wiener*, for example, wrote [4.60]

> *the modern industrial revolution is [...] bound to devalue the human brain at least in its simpler and more routine decisions. Of course, just as the skilled carpenter, the skilled mechanic, the skilled dressmaker have in some degree survived the first industrial revolution, so the skilled scientist and the skilled administrator may survive the second. However, taking the second revolution as accomplished, the average human of mediocre attainments or less has nothing to sell that it is worth anyone's money to buy.*

It is remarkable how many of the wartime engineers involved in control systems development went on to look at social, economic or biological systems. In addition to Wiener's work on cybernetics, Arnold Tustin wrote a book on the application to economics of control ideas, and both Winfried Oppelt and Karl Küpfmüller investigated biological systems in the postwar period.

One of the more controversial applications of control and automation was the introduction of the computer numerical control (CNC) of machine tools from the late 1950s onwards. Arguments about increased productivity were contested by those who feared widespread unemployment. We still debate such issues today, and will continue to do so. *David Noble*, in his critique of automation, particularly CNC, remarks [4.61]

> *[...] when technological development is seen as politics, as it should be, then the very notion of progress becomes ambiguous: what kind of progress? progress for whom? progress for what? And the awareness of this ambiguity, this indeterminacy, reduces the powerful hold that technology has had upon our consciousness and imagination [...] Such awareness awakens us not only to the full range of technological possibilities and political potential but also to a broader and older notion of progress, in which a struggle for human fulfillment and social equality replaces a simple faith in technological deliverance....*

4.10 Conclusion and Emerging Trends

Technology is part of human activity, and cannot be divorced from politics, economics and society. There is no doubt that automatic control, at the core of automation, has brought enormous benefits, enabling modern production techniques, power and water supply, environmental control, information and communication technologies, and so on. At the same time automatic control has called into question the way we organize our societies, and how we run modern technological enterprises. Automated processes require much less human intervention, and there have been periods in the recent past when automation has been problematic in those parts of industrialized society that have traditionally relied on a large workforce for carrying out tasks that were subsequently automated. It seems unlikely that these socio-technological questions will be settled as we move towards the next generation of automatic control systems, such as the transformation of work through the use of information and communication technology ICT and the application of control ideas to this emerging field [4.62].

Future developments in automatic control are likely to exploit ever more sophisticated mathematical models for those applications amenable to exact technological modeling, plus a greater emphasis on human-machine systems, and further development of human behaviour modelling, including decision support and cognitive engineering systems [4.63]. As safety aspects of large-scale automated systems become ever more important, large scale integration, and novel ways of communicating between humans and machines, are likely to take on even greater significance.

4.11 Further Reading

- R. Bellman (Ed.): *Selected Papers on Mathematical Trends in Control Engineering* (Dover, New York 1964)
- C.C. Bissell: ict.open.ac.uk/classics (electronic resource)
- M.S. Fagen (Ed.): *A History of Engineering and Science in the Bell System: The Early Years (1875–1925)* (Bell Telephone Laboratories, Murray Hill 1975)
- M.S. Fagen (Ed.): *A History of Engineering and Science in the Bell System: National Service in War and Peace (1925–1975)* (Bell Telephone Laboratories, Murray Hill 1979)
- A.T. Fuller: *Stability of Motion*, ed. by E.J. Routh, reprinted with additional material (Taylor Frances, London 1975)
- A.T. Fuller: The early development of control theory, Trans. ASME J. Dyn. Syst. Meas. Control **98**, 109–118 (1976)
- A.T. Fuller: Lyapunov centenary issue, Int. J. Control **55**, 521–527 (1992)
- L.E. Harris: *The Two Netherlanders, Humphrey Bradley and Cornelis Drebbel* (Cambridge Univ. Press, Cambridge 1961)
- B. Marsden: *Watt's Perfect Engine* (Columbia Univ. Press, New York 2002)
- O. Mayr: *Authority, Liberty and Automatic Machinery in Early Modern Europe* (Johns Hopkins Univ. Press, Baltimore 1986)
- W. Oppelt: A historical review of autopilot development, research and theory in Germany, Trans ASME J. Dyn. Syst. Meas. Control **98**, 213–23 (1976)
- W. Oppelt: On the early growth of conceptual thinking in control theory – the German role up to 1945, IEEE Control Syst. Mag. **4**, 16–22 (1984)
- B. Porter: *Stability Criteria for Linear Dynamical Systems* (Oliver Boyd, Edinburgh/London 1967)
- P. Remaud: *Histoire de l'automatique en France 1850–1950* (Hermes Lavoisier, Paris 2007)
- K. Rörentrop: *Entwicklung der modernen Regelungstechnik* (Oldenbourg, Munich 1971), in German
- Scientific American: *Automatic Control* (Simon Shuster, New York 1955)
- J.S. Small: *The Analogue Alternative* (Routledge, London/New York 2001)
- G.J. Thaler (Ed.): *Automatic Control: Classical Linear Theory* (Dowden, Stroudsburg 1974)

References

4.1 O. Mayr: *The Origins of Feedback Control* (MIT, Cambridge 1970)

4.2 F.W. Gibbs: The furnaces and thermometers of Cornelius Drebbel, Ann. Sci. **6**, 32–43 (1948)

4.3 T. Mead: Regulators for wind and other mills, British Patent (Old Series) 1628 (1787)

4.4 H.W. Dickinson, R. Jenkins: *James Watt and the Steam Engine* (Clarendon Press, Oxford 1927)

4.5 C.C. Bissell: Stodola, Hurwitz and the genesis of the stability criterion, Int. J. Control **54**, 2313–2332 (1989),

4.6 S. Bennett: *A History of Control Engineering 1800–1930* (Peregrinus, Stevenage 1979)

4.7 G.B. Airy: On the regulator of the clock-work for effecting uniform movement of equatorials, Mem. R. Astron. Soc. **11**, 249–267 (1840)

4.8 J.C. Maxwell: On governors, Proc. R. Soc. **16**, 270–283 (1867)

4.9 E.J. Routh: *A Treatise on the Stability of a Given State of Motion* (Macmillan, London, 1877)

4.10 A. Hurwitz: Über die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Teilen besitzt, Math. Ann. **46**, 273–280 (1895), in German

4.11 E. Bompiani: Sulle condizione sotto le quali un equazione a coefficienti reale ammette solo radici

con parte reale negative, G. Mat. **49**, 33–39 (1911), in Italian
4.12 C.C. Bissell: The classics revisited – Part I, Meas. Control **32**, 139–144 (1999)
4.13 C.C. Bissell: The classics revisited – Part II, Meas. Control **32**, 169–173 (1999)
4.14 M. Tolle: *Die Regelung der Kraftmaschinen*, 3rd edn. (Springer, Berlin 1922), in German
4.15 O. Mayr: *Feedback Mechanisms* (Smithsonian Institution Press, Washington 1971)
4.16 T.P. Hughes: *Elmer Sperry: Inventor and Engineer* (Johns Hopkins Univ. Press, Baltimore 1971)
4.17 S. Bennett: *A History of Control Engineering 1800–1930* (Peregrinus, Stevenage 1979) pp. 137–TSnotePlease provide page range.
4.18 S. Bennett: *A History of Control Engineering 1930–1955* (Peregrinus, Stevenage 1993)
4.19 N. Minorsky: Directional stability of automatically steered bodies, Trans. Inst. Nav. Archit. **87**, 123–159 (1922)
4.20 O. Heaviside: *Electrical Papers* (Chelsea, New York 1970), reprint of the 2nd edn.
4.21 S. Bennett: *A History of Control Engineering 1800–1930* (Peregrinus, Stevenage 1979), Chap. 6
4.22 C.C. Bissell: Karl Küpfmüller: a German contributor to the early development of linear systems theory, Int. J. Control **44**, 977–89 (1986)
4.23 H. Nyquist: Regeneration theory, Bell Syst. Tech. J. **11**, 126–47 (1932)
4.24 H.S. Black: Stabilized feedback amplifiers, Bell Syst. Tech. J. **13**, 1–18 (1934)
4.25 H.W. Bode: Relations between amplitude and phase in feedback amplifier design, Bell Syst. Tech. J. **19**, 421–54 (1940)
4.26 H.W. Bode: *Network Analysis and Feedback Amplifier Design* (Van Nostrand, Princeton 1945)
4.27 H.L. Hazen: Theory of servomechanisms, J. Franklin Inst. **218**, 283–331 (1934)
4.28 A. Leonhard: *Die Selbsttätige Regelung in der Elektrotechnik* (Springer, Berlin 1940), in German
4.29 C.C. Bissell: The *First All-Union Conference on Automatic Control*, Moscow, 1940, IEEE Control Syst. Mag. **22**, 15–21 (2002)
4.30 C.C. Bissell: A.A. Andronov and the development of Soviet control engineering, IEEE Control Syst. Mag. **18**, 56–62 (1998)
4.31 D. Mindell: *Between Human and Machine* (Johns Hopkins Univ. Press, Baltimore 2002)
4.32 C.C. Bissell: Textbooks and subtexts, IEEE Control Syst. Mag. **16**, 71–78 (1996)
4.33 H. Schmidt: Regelungstechnik – die technische Aufgabe und ihre wissenschaftliche, sozialpolitische und kulturpolitische Auswirkung, Z. VDI **4**, 81–88 (1941), in German
4.34 C.C. Bissell: Control Engineering in the former USSR: some ideological aspects of the early years, IEEE Control Syst. Mag. **19**, 111–117 (1999)
4.35 A.D. Dalmedico: Early developments of nonlinear science in Soviet Russia: the Andronov school at Gorky, Sci. Context **1/2**, 235–265 (2004)
4.36 A.C. Hall: Application of circuit theory to the design of servomechanisms, J. Franklin Inst. **242**, 279–307 (1946)
4.37 A.C. Hall: *The analysis and synthesis of linear servomechanisms (restricted circulation)* (The Technology Press, Cambridge 1943)
4.38 S. Bennett: *A History of Control Engineering 1930–1955* (Peregrinus, Stevenage 1993) pp. 142–
4.39 H.J. James, N.B. Nichols, R.S. Phillips: *Theory of Servomechanisms*, Radiation Laboratory, Vol. 25 (McGraw-Hill, New York 1947)
4.40 C.C. Bissell: Pioneers of control: an interview with Arnold Tustin, IEE Rev. **volume]Please provide volume number.**, 223–226 (1992)
4.41 A.L. Whiteley: Theory of servo systems with particular reference to stabilization, J. Inst. Electr. Eng. **93**, 353–372 (1946)
4.42 C.C. Bissell: Six decades in control: an interview with Winfried Oppelt, IEE Rev. , 17–21 (1992)
4.43 C.C. Bissell: An interview with Hans Sartorius, IEEE Control Syst. Mag. **27**, 110–112 (2007)
4.44 W.R. Evans: Control system synthesis by root locus method, Trans. AIEE **69**, 1–4 (1950)
4.45 A.A. Andronov, S.E. Khaikin: *Theory of Oscillators* (Princeton Univ. Press, Princeton 1949), translated and adapted by S. Lefschetz from Russian 1937 publication
4.46 L.A. MacColl: *Fundamental Theory of Servomechanisms* (Van Nostrand, Princeton 1945)
4.47 S. Bennett: The emergence of a discipline: automatic control 1940–1960, Automatica **12**, 113–121 (1976)
4.48 E.A. Feigenbaum: Soviet cybernetics and computer sciences, 1960, Commun. ACM **4**(12), 556–579 (1961),
4.49 R. Bellman: *Dynamic Programming* (Princeton Univ. Press, Princeton 1957)
4.50 R.E. Kalman: Contributions to the theory of optimal control, Bol. Soc. Mat. Mex. **5**, 102–119 (1960)
4.51 R.E. Kalman: A new approach to linear filtering and prediction problems, Trans. ASME J. Basic Eng. **82**, 34–45 (1960)
4.52 R.E. Kalman, R.S. Bucy: New results in linear filtering and prediction theory, Trans. ASME J. Basic Eng. **83**, 95–108 (1961)
4.53 L.S. Pontryagin, V.G. Boltyansky, R.V. Gamkrelidze, E.F. Mishchenko: *The Mathematical Theory of Optimal Processes* (Wiley, New York 1962)
4.54 T.J. Williams: Computer control technology – past, present, and probable future, Trans. Inst. Meas. Control **5**, 7–19 (1983)
4.55 T. Williams, S.Y. Nof: Control Models. In: *Handbook of Industrial Engineering*, ed. by G. Salvendy (Wiley, New York 1992) pp. 211–238, 2nd edn.

4.56 J.C. Willems: In control, almost from the beginning until the day after tomorrow, Eur. J. Control **13**, 71–81 (2007)
4.57 G.S. Brown, D.P. Campbell: Instrument engineering: its growth and promise in process-control problems, Mech. Eng. **72**, 124–127 (1950)
4.58 G.S. Brown, D.P. Campbell: Instrument engineering: its growth and promise in process-control problems, Mech. Eng. **72**, 136 (1950),
4.59 G.S. Brown, D.P. Campbell: Instrument engineering: its growth and promise in process-control problems, Mech. Eng. **72**, 587–9 (1950)
4.60 N. Wiener: *Cybernetics: Or Control and Communication in the Animal and the Machine* (Wiley, New York 1948)
4.61 D.F. Noble: *Forces of Production. A Social History of Industrial Automation* (Knopf, New York 1984)
4.62 S.Y. Nof: Collaborative control theory for e-work, e-production and e-service, Ann. Rev. Control **31**, 281–292 (2007)
4.63 G. Johannesen: From control to cognition: historical views on human engineering, Stud. Inf. Control **16**(4), 379–392 (2007)

http://www.springer.com/978-3-540-78830-0

Springer Handbook of Automation
Nof, S.Y. (Ed.)
2009, LXXVI, 1812 p. 1005 illus. in color. With DVD.,
Hardcover
ISBN: 978-3-540-78830-0