1 Historical Introduction

> Any sufficiently advanced technology is indistinguishable from magic.
>
> *Arthur C. Clarke*[1]

This is a book about magic. Not the magic of wizardry and sorcerers, nor the magic of fairy tales or fables, but *real* magic: the magic that powers planes and runs computers, that keeps our investments high and our blood pressure low, our beer cold and our bodies warm. It is the magic behind physics experiments of extraordinary precision, from gravity-wave detectors that probe the cosmos to scanning probe microscopes that image atoms. It is the magic that regulates biological processes from the pupil size in our eyes to the gene expression in our cells. It is the magic made possible by control theory.

The study of control theory can lead to something of a culture shock for physicists. Of course, jargon, technical methods, and applications may all be new. But something more fundamental is at play: As physicists, we study the world as it is. We look for the fundamental laws that govern time and energy, fields and forces, matter and motion, at the level of individual particles and collective phenomena. We do this in settings that range from the very large scales of the cosmos to the very small scales of fundamental particles to the very complex systems that rule the human scale. But we do all of this on Nature's terms, content to describe the actual dynamics of real "physical" systems.

Control theorists ask, instead, what might be. They seek to alter the states and dynamics of a system to make it *better*. The word "better" already implies a human element, or at least an active agent that can influence its environment. The Ancient Greeks coined the notion of *teleology* to denote the purpose or end (*telos*) of an object. While science has moved away from endowing objects in themselves with purpose, engineers design machines or systems to accomplish predefined tasks. Control theory tells, in a precise way, how to accomplish these tasks and indicates what is possible or not. *Uncertainty* – about initial conditions, external disturbances, dynamical rules, etc. – can limit possibilities.

Since all systems are physical ones, ruled by the laws of physics, physics will play a role in our story. But in many ways, it will have a supporting role, as we seek to create "augmented" systems that perform in ways that seemingly ignore the laws of physics. Of course they do not. Even so, we will see that a larger, open, physical "supersystem" can give a subsystem effective dynamics with new laws and properties.

[1] *Profiles of the Future: An Inquiry Into the Limits of the Possible*, New York, Harper and Row, Rev. ed., 1973.

https://doi.org/10.1017/9780511734809.002 Published online by Cambridge University Press

In this book, we will take a broad look at control, from both the fundamental point of view that seeks to understand what it can accomplish and what not, and how control in general meshes with other topics in physics such as thermodynamics and statistical physics. At the same time, we will also be interested in control for its practical applications. Just as control is fundamental to the technological devices of modern life, so too does it play a key role in the techniques an experimental physicist should know.

Sometimes called the "hidden technology," control is often invisible, despite its omnipresence in modern technology. We do not notice it until something fails. Planes are very safe, but occasionally they fall from the sky. Our bodies also depend on many control loops. To name one: to survive, we must maintain a core temperature within 27–44 °C, implying the need to keep maximum deviations to $< \pm 3\%$ and to regulate typical fluctuations to be $< \pm 0.3\%$. Again, we pay little attention to our body's temperature – except when it begins to deviate when we get sick or cold or hot. Our ability to ignore control under normal circumstances is a testament to its *robustness* to specific types of situations; our need to confront the often-drastic consequences of its failure is a consequence of its *fragility* to unforeseen circumstances. As we will see, the two aspects are linked.

In this introductory chapter, we present briefly the historical development of control and its theory, which gives some insight as to what "better" dynamics might mean. We then list some of these *goals* for control. Then we introduce, in an intuitive way, some of the principal *methods* of control, notably *feedback* and *feedforward*. We conclude with a discussion of the *types* of control systems.

1.1 Historical Overview

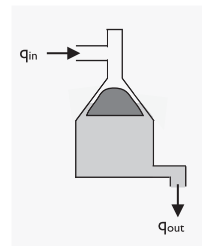
We can divide the development of control techniques and theory into five periods:

- Early control (before 1900)
- Preclassical period (1900–1940)
- Classical period (1930–1960)
- Modern control (1945–2000)
- Contemporary control (after 2000)

The overlaps are deliberate, as actual developments are not as well ordered chronologically as the classifications would imply. Although it seems logical to "begin at the beginning," this summary may be easier to follow after you have learned some of the material from later chapters. Partly for this reason, the discussion is relatively brief, with some aspects deferred to the relevant later chapters. Of course, a short exposition inevitably simplifies a complex story. The notes and references give pointers to more extensive presentations.

1.1.1 Early Control (before 1900)

The word *feedback* is of relatively recent origin, with the Oxford English Dictionary reporting its first use in 1920, in connection with an electrical circuit.[2] However, uses of feedback and the broader notion of control are far more ancient. Ktesibios (285–222 BC), a Greek working in Alexandria, Egypt, used feedback to improve the stability of water clocks, vessels that measure time by the outflow of water. However, as the fluid level in a vessel decreases, so too will its outflow rate. Keeping the level constant, or *regulating* it, stabilizes the rate of outflow. There are no original records of the device, but reconstructions based on Vitruvius's *De architectura* (~ 30–15 BC) and later Arab water clocks indicate that the mechanism was the same as that used in the modern flush toilet: a ball floating in the tank follows the water level. When the level is low, a float lets in more water, raising the level and increasing q_{out}; when high, the float shuts off the valve, decreasing q_{out} (see right).



In the Middle Ages, *mechanical clocks* powered by falling weights or springs were developed, with various ratchets ("escapements") that translate oscillating into rotational motion. These clocks also have feedback mechanisms to ensure constant rotation rates.

Because fluid density depends on temperature, the level of a fluid can be used to regulate temperature. René-Antoine Ferchault de Réamur (1683–1757) invented such a device, based on the temperature sensor of Cornelius Drebbel (1572–1663), a Dutch engineer working in England. In France, Jean-Simon Bonnemain (1743–1830) patented in 1783 an improved temperature controller based on a bimetallic rod that flexed when the temperature changed. He used it to make practical hot-water central heating for buildings.

The beginning of the Industrial Revolution, centered on England in the second half of the eighteenth century, led to the first important applications of feedback. The most prominent was the *governor*, which was developed to keep windmills turning at a constant rate and then adapted to the steam engine for more general purposes by James Watt in the late 1780s.[3] The issue was that variable loads would alter the rotation rate of the engine. To keep it constant, Watt and his partner Matthew Boulton adapted a *flyball* sensor for rotation rates that had been patented by Thomas Mead in 1787. As illustrated in Figure 1.1, the sensor has two heavy balls that rotate with the engine shaft. If the engine rotates too quickly, centrifugal force pushes the balls out, pulling down a lever and shutting off the throttle valve that lets steam in, thus slowing the motor. If it rotates too slowly, the balls fall in, pushing up the lever, opening the value, letting more steam in, and speeding up the motor. If all goes well, the steam-engine rotation rate settles at a desired value.

The nineteenth century saw a steady improvement in the technology of governors. The 1868 paper *On governors* by James Clerk Maxwell gave the first theoretical analysis. A flaw of governors was their tendency to make the engine "hunt" for the right

[2] The related term *feedforward* was first used even more recently, in 1952 (also according to the OED).
[3] By the 1670s, Christiaan Huygens had invented a governor to regulate pendulum clocks (Bateman, 1945).

https://doi.org/10.1017/9780511734809.002 Published online by Cambridge University Press

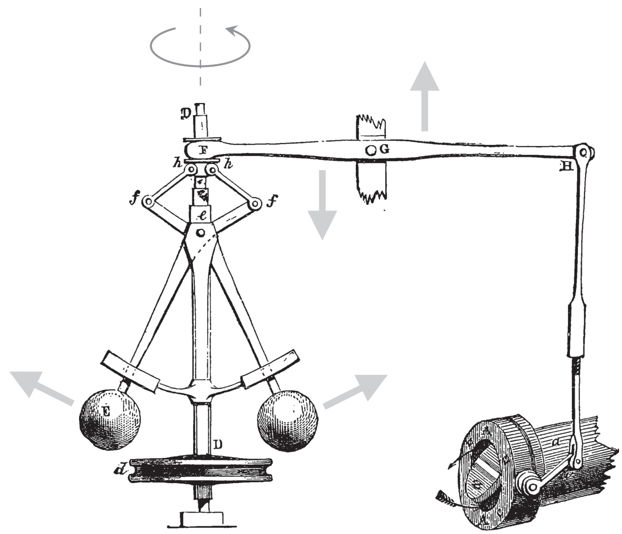


Fig. 1.1 Flyball governor and throttle valve, with rotation around the indicated axis. Flyballs move out, lever at F is pulled down, pivots about G, and pushes up at H, closing the throttle valve at lower right. Adapted from Routledge (1900).

rotation speed. In more modern language, there could be long-lived oscillatory transients before settling to a steady state. Even worse, the engine could become unstable and show erratic motion. Maxwell analyzed the conditions for stability of regulation against small perturbations using *linear stability analysis*. His stability conditions were generalized by Edward J. Routh and Adolf Hurwitz later in the nineteenth century. Although these early analyses of control systems eventually became part of the techniques of control theory in the mid twentieth century, they had little immediate impact on practical realizations, which was driven by the innovations of "tinkerers." Another emerging class of control applications concerned the position of a moving object. Thus, ships needed steering and missiles guiding to their target. In England, J. McFarlane Gray patented in 1866 a steering engine using feedback. In France, Jean Joseph Farcot introduced a range of position-control devices that he called *servomotors*. More generally, servomechanisms were used to *track* desired time-dependent trajectories, a generalization of the simpler goal of *regulation*, where the desired trajectory is simply a constant.

1.1.2 Preclassical Period (1900–1940)

Pre-1900 regulators were all *direct acting*: the elements that measured the quantity being regulated also had to change the system. The lever in a fluid-level regulator that moves in response to a change in level also opens the value that lets in more water. Of course, there is a "power source" (a high-pressure supply of water) that makes the response possible, but one "gadget" must still carry out two actions. Around 1900 a long process of abstraction began that led to distinct notions of *sensors*, *controllers*,

https://doi.org/10.1017/9780511734809.002 Published online by Cambridge University Press

and *actuators*. The sensor *measures* a quantity of concern, the controller *decides* how to respond, and the actuator *executes* the response. Each element can have its own, independent source of power. Such ideas, however, took several decades to become clear.

Meanwhile, the first decades of the twentieth century saw the beginnings of industrial process control. Applications included boiler control for steam generation, electric motor speed regulation, steering for ships and airplanes, temperature and pressure control, and more. A key development was of stand-alone controllers that could be added on to existing equipment. For example, around 1910, Elmer Sperry greatly improved the gyrocompass and designed a gyroscope autopilot to steer ships. The Sperry Gyroscope Company supplied the US Navy with navigational aids, as well as bomb sights and fire-control systems.

In 1922, Nicholas Minorsky gave a detailed analysis of such mechanisms, introducing the notion of three types of control. The first is *proportional* to the error between *set point* and actual signal, the second to the *integral* of that error, and the third to its *derivative*. Together, they form the three-term regulator, or proportional-integral-derivative (PID) control, which is discussed in Chapter 3. Although these ideas now seem very general, they were at first encountered separately in each domain of application. Thus, Minorsky's analyses were little known in the broader technical community for a number of years.

Another important development was the first airplane flight by Orville and Wilbur Wright. Others had built (and sometimes died testing) unsuccessful flying machines. The Wright brothers' success was based on their mastery of control, using flaps to alter yaw, pitch, and roll (three axes). Moreover, they recognized the advantages of an inherently *unstable* design stabilized by control (e.g., a human pilot). Unstable systems are more maneuverable than stable ones. They need active feedback to produce stable motion but can respond to disturbances (gusts of wind, abrupt change in terrain, etc.) much more quickly. The concept should be familiar: when we stand upright, we are unstable and must use (unconscious) small muscle movements to prevent ourselves from falling over. Indeed, the ability to walk on two legs is what distinguished the first hominids from other apes.[4]

Along with developments in mechanical control systems came parallel ones in electrical circuits. By the end of the nineteenth century, there was already a division between the power and signal applications of electricity. In both, the *amplifier* was a key element, allowing separation of the functions of sensor and actuator. Early high-power amplifiers took the form of relays and spring-based solenoids, which became the basis of many kinds of actuators.

For low-power electrical signals and their circuits, a key development was Lee de Forest's 1906 *grid audion*, a vacuum-tube amplifier that could boost the voltage level of a weak signal, compensating for signal losses in transmission and making possible

[4] What were the evolutionary advantages of walking upright? Darwin thought that it improved our ability to fight. But walking is also more efficient for traveling large distances on the ground (e.g., over grasslands). As with many evolutionary developments, the "why" is elusive (Wayman, 2012).

long-distance telephone networks. But the amplifiers had serious flaws: the signal gain was both nonlinear and prone to drifts, which led to distortion and volume variations.

Finally, there was a transformation in our view of living beings. Life in the nineteenth century was fixed on Newtonian, mechanical motion. Things alive *moved*, powered perhaps by the electric spark that jolted Frankenstein's monster to life or by some other unknown vital force. In the 1920s, the physiologist Walter Cannon introduced the term *homeostasis*, the ability to maintain conditions in the face of external perturbations. These conditions include the core temperature of the body and the concentrations of glucose, iron, oxygen, calcium, sodium, potassium, and other chemicals or ions. All these quantities are closely regulated, even when external conditions change dramatically: through hot or cold, our core temperatures are close to 37°C, our sodium levels stay between 135 and 145 milliequivalents per liter, and so on. The ability to regulate so many quantities in the body using multiple, hierarchical systems is one of the defining features of the modern view of life. Conversely, death is associated with a *failure cascade* that shuts down the essential functions of the body with its nested control loops, often one after the other. Understanding homeostasis was a goal of Wiener's influential book *Cybernetics*, a founding text of control theory, discussed below.

1.1.3 Classical Period (1930–1960)

At Bell Telephone Laboratories, a group of engineers was set up to address quality problems in the growing telephone network. Initial progress was slow, but on Tuesday morning, August 2, 1927, Harold Black had an epiphany while riding the Lackawanna Ferry across the Hudson to Manhattan to get to work. His idea, sketched out on a blank page of the *New York Times*, was that by taking a portion of the amplifier output signal and subtracting it from the input, one could reduce distortion, at the cost of a reduced gain. Thus was born the *negative-feedback amplifier*, which had the immediate effect of improving long-distance telephone calls and was a key development in the history of control. Its descendant, the *operational amplifier*, is described in Chapter 3. More broadly, efforts to understand what Black had created led to a "classical" formulation of control theory.

In the 1930s, Black's colleagues at Bell Labs, Harold Nyquist and Hendrik Bode, contributed theoretical analyses that put negative feedback and other ideas of *classical control* on a firmer footing. Work by Harold Hazen and Gordon Brown at MIT also was influential. In contrast to earlier studies based on solving ordinary differential equations in the time domain, they used frequency-domain methods based on the Laplace transform to derive a set of heuristic rules (often expressed in graphical form) for controllers of reasonable performance that work well for a relatively large class of systems. Bode's 1945 book *Network Analysis and Feedback Amplifier Design*, delayed because of the war, is perhaps the apotheosis of classical control. It considered *robustness* in depth, pointing out the fundamental compromises inherent in control: feedback that suppresses the response to disturbances at some frequencies will inevitably boost that response at other frequencies.

https://doi.org/10.1017/9780511734809.002 Published online by Cambridge University Press

The next great impetus to the fledgling field of control engineering (and its counterpart, control theory) came with World War II. Engineers worked on a variety of control problems, notably the aiming of antiaircraft guns and automatic radar tracking. The *Radiation Lab* at MIT was a particularly important center for such research. To the scientists and engineers working at such centers, the war made particularly clear the need for unified, abstracted treatments of control based on concepts that were independent of specific applications. The classified results released en masse at the end of the war spurred rapid progress afterwards.

1.1.4 Modern Control (1945–2000)

After the end of World War II, control emerged as a distinct technical discipline. Engineering societies such as the American Society for Mechanical Engineers (AMSE), the Instrument Society of America (ISA), and the Institute of Radio Engineers (IRE, later the IEEE) all launched subgroups, and new professional societies such as the American Automatic Control Council (AACC) and International Federation of Automatic Control (IFAC) were created. Where MIT had stood almost alone as an academic center, many universities around the world added groups focusing on automatic control. The military-industrial complex took shape: think tanks such as the RAND corporation in Santa Monica, California and the Research Institute for Advanced Study in Baltimore and companies blurred military and industrial roles on scales larger than had been known before the War.[5] Prominent companies included IBM, General Electric, Hughes Aircraft, Bell Labs, Honeywell, Westinghouse, Leeds and Northrup in the United States, and Siemens (Germany), Schneider (France), ASEA (Sweden), and Yokogawa and Mitsubishi (Japan). Regular national and international conferences began: The first IFAC World Conference, in 1960 in Moscow at the height of the Cold War, marked the emergence of *modern control*.[6]

Modern control introduced state-space methods that marked a return to analysis in the time domain, in contrast to the frequency-domain methods characterizing classical control. The latter is fine for time-invariant, linear systems but cannot describe easily time-varying, nonlinear dynamics, which is omnipresent in applications. Although "modern," the state-space approach reaches back to the late nineteenth and early twentieth centuries. and includes figures such as Aleksandr Lyapunov in Russia and Henri Poincaré in France. A key insight was that knowing the system dynamics could improve performance spectacularly relative to the classical methods, which were developed assuming much less about the system under control. The resulting *optimal control* gave a systematic way to generate "the best" controller for a given task. With key contributions from Richard Bellman and Rudolph Kalman in the US and Lev Pontryagin in the Soviet Union, optimal control had spectacularly successful applications in the space program, particularly the Apollo moon-landing project.

[5] The RIAS was absorbed into the Martin Marietta Corporation, which survives as Lockheed Martin.

[6] Obviously, this use of "modern" is dated, as is "modern physics" (relativity and quantum mechanics), "modern art" (Impressionism, Dada, etc.), and "modern architecture" (Bauhaus, International Style, etc.).

The *digital computer* had a long gestation that was greatly advanced by war efforts – e.g., to formulate tables to aid in fire control. The history of computers is a separate story; in control, there was a gradual shift from *analog controllers* to *digital controllers*. The former had been implemented by external electrical, hydraulic, or mechanical circuits. Then came a long evolution to digital mainframes, minicomputers, microcomputers, laptops, and microcontrollers. In parallel came a shift from analog control methods for continuous-time dynamical systems to digital control methods for discrete dynamical systems.

At MIT in World War II, Norbert Wiener, introduced the stochastic analysis of control problems at roughly the same time and independently of efforts in the Soviet Union led by Andrei Kolmogorov. Wiener's primary technical publication, *The Extrapolation, Interpolation, and Smoothing of Time Series with Engineering Applications*, was circulated as a classified report in 1942 and eventually published in 1949.[7] His famous 1948 contribution, *Cybernetics: or Control and Communication in the Animal and the Machine*, showed that control theory applied not only to engineering systems but also to human, biological, and social systems. The book was inspired by the notion of homeostasis in organisms and by similar issues in controlling complex systems. Coined by Wiener from the Greek word for "governance," the word "cybernetics" was a tribute to Maxwell's 1868 governor paper.

The successes of optimal- and stochastic-control methods during the 1960s soon led to overconfidence, as it was forgotten how much the optimized performance depends on knowing system dynamics. In the 1970s and 1980s, control theory underwent something of an identity crisis. While state-space methods work well in the aerospace industry, where dynamics can be known accurately, they do poorly in industrial settings where the dynamics are more complex and harder to characterize (paper mills, chemical plants, etc.). This disenchantment led many practical engineers (and physicists) to avoid advanced techniques in favor of the tried-and-true PID controller. Indeed, academic research on control theory from 1960 through at least the 1970s had "negligible" impact on industrial applications. In response came a new subfield, *robust control*, to optimize the performance of systems whose underlying dynamics had at least moderate uncertainty. Its goal was to merge the robustness of classical control methods with the performance of modern control. In parallel came the subfields of *system identification* and *adaptive control*, where the goal was to *learn* better the system dynamics, either through independent or online measurements. Common applications of adaptive algorithms include noise-cancelling headsets, automobile cruise control, and thermostats.

1.1.5 Contemporary Control (after 2000)

Beginning around 2000, control theorists began to tackle increasingly complex systems. One notable example is the attempt to understand biological systems from

[7] Its nickname "Yellow Peril" came from the color of its cover, difficulty of its contents, and racism of its times.

an engineering perspective emphasizing control especially. The resulting field of *systems biology* contrasts, occasionally sharply, with a parallel effort in physics known as *biological physics*. Perhaps the most striking conclusion concerning control is strong empirical evidence that organisms have *internal models* of the world that allow them to anticipate and plan ahead. Recognizing the role of such planning and anticipation has been key to understanding how humans move and act in the world. See Example 3.4 for an application in connection with human balance – how we manage to stand upright without falling down. Later, in Chapter 10 we introduce *reinforcement learning*, a technique for learning how to plan and anticipate from repeated supervised trials.

Another example is the development of *autonomous*, self-driving vehicles. Indeed, the development of the automobile recapitulates the entire history of control. The early twentieth-century automobile was a mechanical device, like the governor and steam engine. In the 1970s, a variety of electrical control systems appeared, many based on microcontrollers that implemented feedback loops. By 2007, the typical automobile had 20–80 microprocessors, dealing with powertrain control to reduce emissions (e.g., by controlling the air-fuel ratio), performance optimization (e.g., variable cam timing), and driver assistance (e.g., cruise control and antilock brakes), and more. And, while the driver – the "human in the loop" – remains the ultimate controller for an automobile, many responsibilities are off-loaded (e.g., GPS and its associated navigational aids). At present, many companies seek to eliminate the driver from these control loops, a goal that must integrate many subproblems and use techniques from fields such as machine learning, big data, and wireless communications. Finally, control is expanding beyond the scale of single vehicles. Highways and smart phones already give real-time information on traffic for more efficient routing. In the future, platoons of trucks may travel in closely spaced groups that reduce traffic and increase fuel efficiency by controlling the collective air flow around the group (*drafting*).

Another application is to *climate science* and models of climate change, where it is crucial to understand feedbacks on both fast and slow timescales. On the one hand, water vapor is an effective greenhouse gas that is a "fast feedback" because the amount of water vapor in the air adjusts within days to changes in temperature. On the other hand, the area of land covered by glaciers and ice sheets adjusts much more slowly. (Glaciers melt, exposing darker surfaces, which absorb more sunlight.) Such *positive feedback* can lead to instability that will drive a dynamical system to another attractor (a new steady state or, sometimes, an oscillatory one). Negative feedbacks occur in climate models, too. As warmer temperatures lead to greater cloud cover, more light will be reflected away by the clouds, lessening absorption. Unfortunately, positive feedbacks seem likely to dominate.

The desire to understand complex systems has led to a discipline of *network science*. In the context of control, the goal is to understand collective network dynamics rather than individual dynamical systems. One focus has been to understand how the structure of a network affects one's ability to control it. Applications are widespread, as networks are everywhere, from the world-wide web to the proteins that control cellular processes, to control of the brain and neurological disorders, to ecological species interactions, to sets of friends and other social networks. See Chapter 14 for background on the principles.

In parallel with distributed systems came distributed control. The world is moving rapidly to an *Internet of things*, where everyday objects are networked together but also have autonomous controllers. The cities of today and homes of tomorrow are or will be filled with intelligent devices that communicate with and among each other.

These developments have been accompanied by an increasing awareness of the role of *information* in control. Information theory is an older topic, born in the work of Claude Shannon at Bell Labs, whose 1949 book *The Mathematical Theory of Communication* had an enormous impact, spread over many fields. Curiously, it did not have much initial impact on control theory. But by the 2000s, it was clear that many of the laws due to Bode, Nyquist, and others that limit what is possible in control have connections to information theory. Their complicated frequency-domain expressions turn out to have simpler time-domain versions when expressed in information-theoretic terms. See Chapter 15.

In parallel, there has been a renewed interest from physicists in control. In addition to his paper on governors, Maxwell actually made a second, nearly simultaneous contribution to control in 1867.[8] Asked to review a draft of an early treatise on thermodynamics by Peter Guthrie Tait, Maxwell responded with a letter outlining the thought experiment now known as *Maxwell's demon*. In Chapter 15, we discuss how this thought experiment led eventually to an understanding of some of the fundamental limits of control and to deep connections among thermodynamics, information theory, and feedback.

Finally, there has been great theoretical interest and practical success in controlling quantum systems. We give a brief introduction in Chapter 13.

1.2 Lessons from History

History can be very interesting on its own and can illuminate the sociology of a profession – where engineers are "coming from." But we can also learn lessons from history. From control applications, we can extract the types of goals or objectives that control can have. We can also attempt to define terms such as feedback and feedback loop that can be slippery when tested against the full range of situations where the terms are used. Finally, we can classify control – open and closed loop, autonomous and nonautonomous, simple and complex controller – to help make sense of the various historical examples.

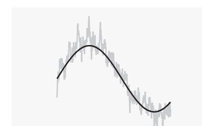
[8] Maxwell's two contributions to control theory are considered to be among his "minor work," but only compared to those on kinetic theory and, especially, on the unified description of electricity and magnetism.

https://doi.org/10.1017/9780511734809.002 Published online by Cambridge University Press

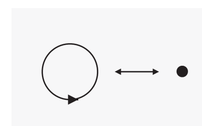
1.2.1 Goals of Control

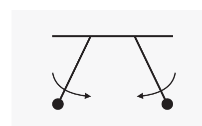
From the historical development, we can abstract several goals for control systems. To best express these goals, we define an internal state $\mathbf{x}$, a real-valued vector that evolves continuously in time t and obeys an ordinary differential equation, $\dot{x}(t) = \boldsymbol{f}(\boldsymbol{x})$, a case that we shall study a lot in this book.[9] The internal state might correspond to a point in phase space, as in classical mechanics, but the notion is more general in that any dynamical rule will do, including ones with phenomenological damping terms, and the like. We will discuss this notion of internal state in more detail beginning in Chapter 2. Here, we just list some *goals for control* and illustrate them in sketches at right:

1 *Regulation*. The internal state should be kept constant (e.g., $\mathbf{x}(t) = \mathbf{x}^*$). This was historically perhaps the first goal of control. The idea is to fix a system's state in the face of perturbations of various sorts. Thus, at home, we regulate temperature against cold and hot swings of the environment. Our body does much the same for our internal body temperature. And temperature control is often an important requirement in an experiment. Of course, many other quantities such as pressure, position, velocity, concentration – any kind of internal state or quantity related to such a state – can be regulated.

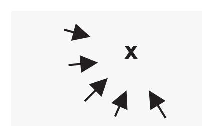


2 *Tracking*. The system state vector $\mathbf{x}(t)$ should approximate a desired time-dependent trajectory $\mathbf{x}^*(t)$ as closely as possible. That is, we generalize regulation about a constant to regulation about a curve in state space.





3 *Changing the attractor*. Rather than trying to generate a particular trajectory, we can change the *type* of motion that a dynamical system can generate. For example, a system that would normally oscillate might be altered to have a steady state. More complicated motion such as *chaos* can also be suppressed. Conversely, you might want to induce any of these types of motion.



4 *Collective dynamics*. The goal is to induce multiple dynamical systems to work together to accomplish some task. An example is *synchronization*, where two or more systems evolve identically in time, even in the absence of "target" system dynamics. Temporal behavior can also be correlated according to more complex rules, as found in *musical ensembles* and *dancing*. By contrast, *swarming*, illustrated at right, refers to correlated spatial motion.

While not exhaustive, the above list gives a sense of the broad range of possibilities. Mostly, we will focus on regulation and tracking, as they involve most of the basic ideas and issues in control theory. In Chapter 11, we discuss some of the other types of goals.

1.2.2 What is a Feedback Loop?

Examining the historical development shows an evolution in the different types of control and controllers. Figure 1.2 shows four examples. In (a), we show the governor

[9] We shall also study many other types of dynamical systems, such as extended systems, discrete-time evolution, and discrete states.

https://doi.org/10.1017/9780511734809.002 Published online by Cambridge University Press

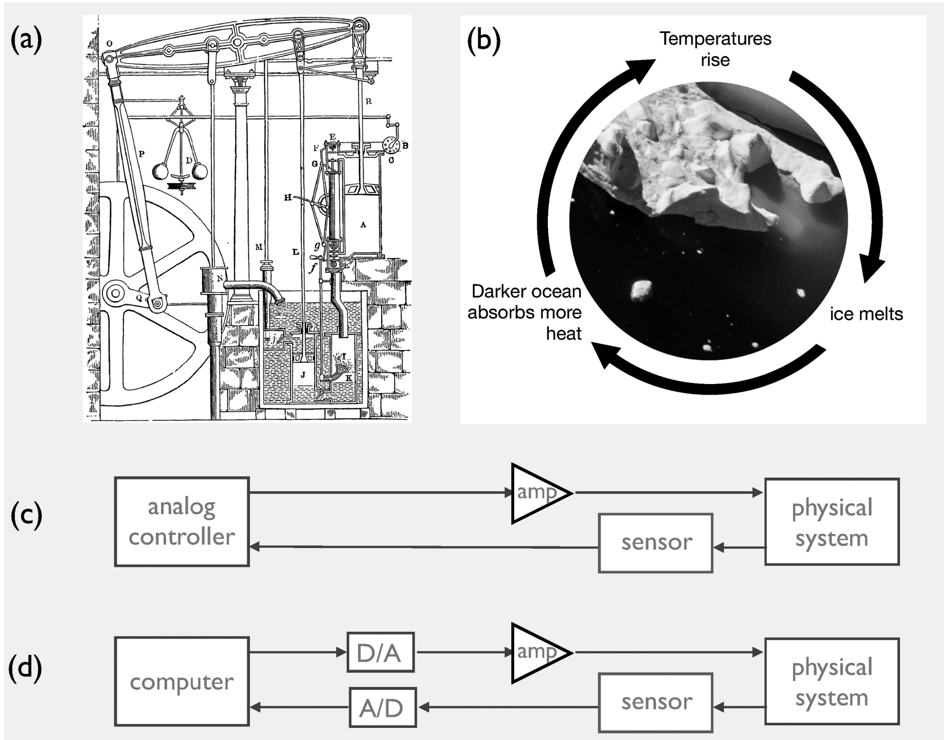


Fig. 1.2 Four examples of feedback systems. (a) Watt's steam engine with governor (labeled D). Adapted from Routledge (1900). (b) Positive feedback loop can reinforce climate change. Author's adaptation of photograph supplied by Monica Bertolazzi/Getty Images. (c) Analog feedback controller with accompanying sensor and power amplifier. (d) Digital feedback controller with digital-analog and analog-digital converters.

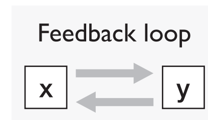
depicted previously in Figure 1.1 integrated into the full steam engine designed by Watt. The governor is indicated by "D" in the drawing. If you had not already seen a governor (e.g., in Figure 1.1), you might have difficulty identifying it amid all the clutter of parts. The whole "contraption" is one physical system consisting of many wheels, rods, and the like joined together in complicated ways, and nothing in particular distinguishes the pieces that form part of the engine from those that form part of the controller. Both are obviously physical subsystems.

In (b), we set history aside to consider a "natural system," where it is difficult to distinguish the controller from the system. The system is the earth, or at least its climate system, and control (in this case, an undesirable destabilizing control) is given by sea ice and ocean.

In (c), we show the kind of analog controller that evolved in the first half of the twentieth century. It consists of the physical system of interest, plus some add-ons: controller, sensor, power amplifier, and power supply (not shown). In contrast to the cases of (a) and (b), it is easy to identify all elements – they are connected by wires – and each is obviously a physical (sub)system on its own. The add-ons are analog circuits. The system of interest can range widely (e.g., a block of metal to keep at constant temperature). In Chapter 3, we will see how to make a controller circuit using operational amplifiers.

In (d), we have transitioned from analog to digital control. The controller is now implemented as a program on a computer (or microcontroller or other similar digital device). The computer communicates in 1s and 0s to the analog world via converters that take a number and output a corresponding voltage, and vice versa. After conversion to analog signals, the other elements are the same as in (c). See Chapter 5.

One of the most important concepts expressed in Figure 1.2 is the notion of a *feedback loop*, where the output is "fed back" to the input. The system thus affects the controller (via a measurement) and, conversely, the controller affects the system (by sending out a response). In the sketch at right, we show the structure of this basic, stripped-down notion of a feedback loop. Its two distinct parts, the system of interest $\boldsymbol{x}$ and controller $\boldsymbol{y}$, typically obey coupled dynamical equations, such as



$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})\,, \qquad \dot{\boldsymbol{y}} = \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})\,. \tag{1.1}$$

The sketch illustrates two essential features of a feedback loop: *separability* and *causality*.

Separability implies that the dynamics can be decoupled into two coupled (vector) equations (here into $\boldsymbol{x}$ and $\boldsymbol{y}$) in a way that makes sense physically. That is, $\boldsymbol{x}$ and $\boldsymbol{y}$ should each describe physically distinct systems with a meaningful identity when considered in isolation. Mathematically, in the equation $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})$, we could freeze $\boldsymbol{y}$ at some constant value (e.g., $\boldsymbol{0}$) and then look at the "pure" $\boldsymbol{x}$ dynamics. Or we could do the same with the $\boldsymbol{y}$ dynamics. But physically, the separation should make sense. For example, the N identical molecules in a gas are typically described as a single N-particle system of interacting particles, whereas two large particles in a gas of N molecules are usefully described separately, and one can imagine the motion of one "controlling" the motion of the other.

Causality is linked with notions of time: The controller *first* measures and *then* responds. Or, $\boldsymbol{x}$ affects $\boldsymbol{y}$, which then affects $\boldsymbol{x}$. In Eq. (1.1), imagine that the coupled system settles to a fixed point $\{\boldsymbol{x}^*, \boldsymbol{y}^*\}$. If the response to fast variations goes to zero as the frequency characterizing the variation increases, then some kind of notion of causality is present. In Chapter 15, we will develop the formal links between causality and response functions, beginning with the classic Kramers–Kronig relations.

Not all coupled systems form feedback loops. Imagine a simpler set of equations, $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0}$, $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0}$, which also have solution $\{\boldsymbol{x}^*, \boldsymbol{y}^*\}$. But now if we perturb one of these equations, the equilibrium solution will be modified instantaneously, because we are just solving an algebraic set of equations. Even though it is still true that $\boldsymbol{x}$ affects $\boldsymbol{y}$ and vice versa, we would probably not call this a case of feedback – it would just be coupling.

We might try to formalize this notion by comparing the interaction time scales, the time it takes $\boldsymbol{x}$ to affect $\boldsymbol{y}$ or vice versa, with time scales within each system. Typically, the interaction times will be faster than most time scales in the system. If not, as in

https://doi.org/10.1017/9780511734809.002 Published online by Cambridge University Press

the "instantaneous" case from the previous paragraph, we would hesitate to call the interaction a feedback loop. I am sure that one can find violations of any specific requirement, making these criteria heuristic, not rigorous. But I think they can be useful in understanding the different types of systems that different communities call "feedback loops."

"Feedback loop" is thus a theoretical construct imposed on a system. In a modern version of a controlled system where each element (controller, amplifier, sensor, system, etc.) is housed in its own box and connected to other elements by wires, it is easy to see the loop. The governor and similar devices lack the wires, but we can still easily distinguish system from controller. In the climate-change example, still more imagination is required. Indeed, in a number of fields such as biophysics, identifying feedback loops is an important research topic. The cell is a complicated pile of proteins, but the proteins function as machines – of a different sort, since they live in a strongly fluctuating environment, but machines nonetheless. They are every bit as complex as the mechanical factories of the machine age, and there has been a long effort to identify functions such as gene regulation, where feedback plays a key role.

Finally, one aspect that is implicit in the notion of feedback loops but not readily apparent visually, is the notion of correcting against *uncertainty* of various types. We will see that the major – some would argue the only – reason to use feedback is to compensate for unknown disturbances and system dynamics. If there is no uncertainty, then other types of control such as feedforward (see below) will be more appropriate.

1.2.3 Progression of Control Types

In Figure 1.2, we see a progression in abstraction of control systems, from "contraptions" – machines that are constructed to improve a particular performance characteristic such as engine rotation rate – to an ensemble of physical subsystems to digital systems that seem quite divorced from the original physical phenomena, even though computers are, of course, physical systems in themselves. This increasing abstraction has been useful in understanding key issues: the need for feedback loops, the role of time delays and noise, and so on. It is at the cost of an analysis that strays far from any given particular system, and one must be careful in applications that too much has not been lost. For example, we will see that all controller design implies a model of the system, whether explicitly expressed or not. The quality of control will then be linked to the quality of the model.

How does feedback in natural systems fit into the above picture? For example, we have seen that in climate science, positive and negative feedback is an important part of understanding the earth's climate and its response to man-made changes. On the one hand, you might take the view that, since these are entirely natural systems that have not been engineered, feedback is not an appropriate concept. On the other hand, the terms are in common use and well accepted. They are best applied to situations where one can distinguish subsystems that operate independently, at least to some extent. In the example shown in Figure 1.2b, the different parts (air temperature, sea ice, and so on) each have distinct identities and can be measured independently. Thus, it acts very similarly to engineered systems such as the steam-engine governor, and using the term "feedback" seems reasonable. But even in climate science, identifying all such relevant feedbacks is far from clear, never mind quantifying their potential impacts.

1.2.4 Classifying Control and Controllers

We introduce a few distinctions that will be useful in organizing our understanding of control. In particular, we distinguish between open-loop and closed-loop designs, between feedforward and feedback control algorithms, and between autonomous and non-autonomous systems. We will not try for rigorous definitions because these terms are used inconsistently in different communities. Moreover, without learning more about the subject, it is difficult to appreciate fine points. Nonetheless, there are general ideas and useful distinctions that are worth making.

Open and Closed Loops

We have already met the concept of feedback and the feedback loop in Section 1.2.2. We view feedback as a technique for *closed-loop* control, which is a design where the control is based on the system's current (and past) states. So pervasive is closed-loop control that its alternative is known as *open-loop* control, which is a strange name when you think about it. The distinction is shown graphically in Figure 1.3. Open-loop designs use *feedforward*, where an input is "fed forward" to the system, and the control acts *independently* of the system state.[10] Although feedforward might seem just a limited version of feedback, its reliance on anticipated events can speed response relative to feedback, which reacts to perturbations after the fact. Feedforward also will not destabilize a system, a pitfall of feedback. We will see in Chapter 3 that the best control designs often incorporate elements of both feedforward and feedback.

The distinction between open- and closed-loop control makes sense – and indeed resulted from, human-engineered systems, where subsystems and wiring connections are often known. It is less clearly useful when applied to the feedbacks of natural systems such as the climate example given above.

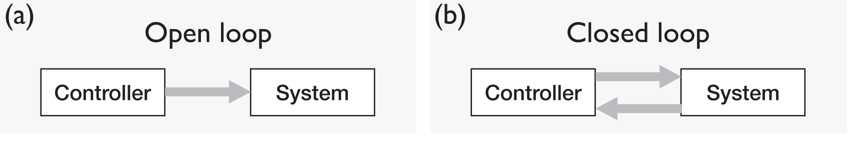


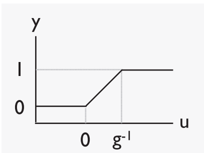
Fig. 1.3 Types of control. (a) open loop; (b) closed loop.

[10] Some people reserve the term "feedforward" for the special case where measurements of incoming disturbances allow a controller to act before the disturbance arrives. We will refer to that case as *disturbance feedforward* (Section 3.4.3) and here use "feedforward" more broadly.

https://doi.org/10.1017/9780511734809.002 Published online by Cambridge University Press

Negative and Positive Feedback

Another basic distinction in control theory is between *negative* and *positive feedback*. To develop our intuition, consider a system with a static, *open-loop response*, where the output y is a static function of the input u. Since the response of a physical system is limited in range, let



$$y = \mathrm{sat}_g(u)\,. \tag{1.2}$$

where the *saturation* function is defined at left: the output y goes from 0 to 1 as the input u goes from 0 to g^{-1}. The constant g is the linear *gain* of the response function.

Using the open-loop response as a reference (b), we can explore the impact of negative (a) and positive (c) feedbacks in Figure 1.4. In (a), we show a *block diagram* illustrating the signal flow for negative feedback, with $k < 0$. The output y is fed back and subtracted from the input, using a *feedback gain*, k. There is now a *closed-loop response*

$$y = \mathrm{sat}_g(u - |k|y) \equiv \mathrm{sat}_{g'}(u)\,, \tag{1.3}$$

where the new saturation function has gain

$$g' = \frac{g}{1 + |k|g}\,. \tag{1.4}$$

The linear-response range of the input has increased from g^{-1} to $g^{-1} + |k|$, as illustrated by the dotted lines in the response curve at lower left in Figure 1.4a. Notice that for large gain ($|k| \gg g^{-1}$), the linear range is approximately independent of the linear range of the original system. If the goal is to regulate the output y, we see that we can keep y in a given range over a wider range of inputs, which might come from a noisy environment. Thus, negative feedback can make the output less sensitive (more robust) to variations of the input.

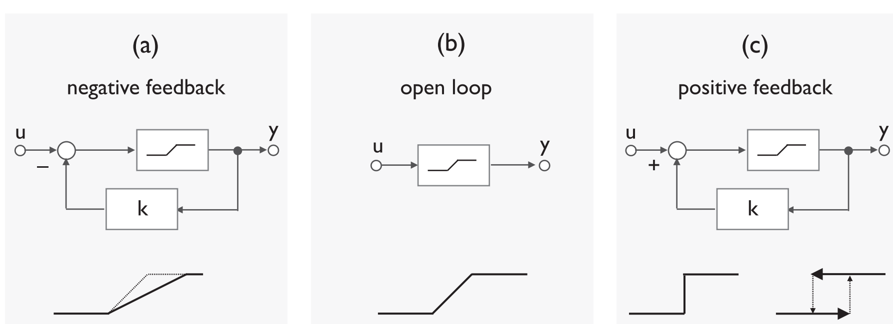


Fig. 1.4 Feedback affects response. (a) Negative feedback increases the linear range. (b) Open-loop response. (c) Positive feedback decreases the linear range, until it produces a switch (left) or memory (right). Adapted from Sepulchre et al. (2019).

https://doi.org/10.1017/9780511734809.002 Published online by Cambridge University Press

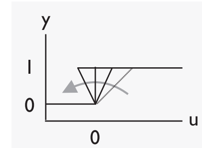


Figure 1.4c and the sketch at right show the effects of positive feedback, $k > 0$. For $k < g^{-1}$ the response is steepened, obeying

$$g' = \frac{g}{1 - kg}\,, \tag{1.5}$$

until at $k = g^{-1}$, the gain diverges and the response becomes a step (part c, lower left sketch, and vertical line at left). Such a response can be viewed as a *switch* or threshold detector. It implies extreme sensitivity: a finite response due to an infinitesimal input variation. For larger gains $k > g^{-1}$, the response has a backwards slope, and it is easy to show that there is *hysteresis* and a binary output. The output is $y = 0$ until it suddenly jumps to $y = 1$ when u is increased from negative values through 0. But if the output is $y = 1$ for positive u, it stays at that value until $u = g^{-1} - k < 0$, when it transitions back to $y = 0$. See the lower right sketch in Figure 1.4c. We have created a *memory*, which can exist in two states over some range of input values.

The system described here is very simple, as it is just a static response. Yet already, we see that negative feedback can regulate an output, making it less sensitive to variations in its input and that positive feedback can create a switch or memory. In Chapter 3, we shall explore how dynamics generalizes this picture, leading to gains that can be negative at some time scales (frequencies) but positive at others. And in Chapter 11, we will understand better how the positive-feedback response is shaped by nonlinearities in the dynamics. Here, the nonlinearity corresponds to the saturation of the output at 0 and 1.

Autonomous and Nonautonomous

To understand the various types of feedback systems, physicists interested in control have classified systems into *autonomous* and *nonautonomous* cases. The terminology comes from the analogous classification of ordinary differential equations:[11]

$$\underbrace{\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})}_{\text{autonomous}}\,, \qquad \underbrace{\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, t)}_{\text{nonautonomous}}\,, \tag{1.6}$$

where the distinction is that the dynamics of nonautonomous systems have explicit time dependence, whereas autonomous systems do not. Thus, Watt's combination of steam engine and governor would be an autonomous feedback system, as the ultimate power sources are simply two heat reservoirs at different temperatures. All the time dependence of motion and its regulation is generated through the internal dynamics. By contrast, digital control systems with their explicit measurements and decisions taken internally in a computer program would be the archetype of nonautonomous control.

One subtlety is that the distinction between autonomous and nonautonomous depends on the level of description. This is true for the theory of ordinary differential equations, where one can always rewrite an n-dimensional system with

[11] We define and review such equations in more detail in Chapter 2.

https://doi.org/10.1017/9780511734809.002 Published online by Cambridge University Press

nonautonomous dynamics as the projection of a system that lives in an $n+1$-dimensional state space. That is,

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, t) \quad \rightarrow \quad \begin{cases} \dot{\boldsymbol{x}} &= \boldsymbol{f}(\boldsymbol{x}, x_t) \\ \dot{x}_t &= 1\,, \qquad\qquad y(0) = 0\,. \end{cases} \tag{1.7}$$

In Eq. (1.7), the $\boldsymbol{x}$ subsystem behaves nonautonomously in both cases. Because the second system has $x_t = t$, we would not be able to distinguish the variable x_t from time t.

Physically, the dynamics describing one system may be nonautonomous, but they can be part of a larger autonomous system. For example, a digital control loop (Figure 1.2d) implies that the system under control will seem nonautonomous, because the digital controller can drive the system in a time-dependent way. But the entire system – with computer, amplifiers, system, and the like – is autonomous because it is powered by fixed power supplies delivering a constant power level.

Simple and Complex Controllers

Let us define complex controllers as ones that are (equivalent to) *universal Turing machines*. In particular, they are capable of implementing any algorithm. Simple ones are all the others. This distinction will allow us to distinguish controllers depicted in Figure 1.2c from d. The analog controller in (c) can be very simple. We will see in Chapter 3 how to make a basic controller from an electrical circuit with just a small number of components. The controller in (d) is a digital computer, and we know that such machines can be programmed to give almost any output.

The distinction has the advantage of allowing us to include in the set of complex control systems *cybernetical* cases, where a human being is part of the control loop. (If you drive a car or ride a bike, you are the human in the control loop.)

As with the other distinctions, this definition is meant to be more instructive than rigorous. In particular, we can imagine a family of machines progressing in complexity from the simple circuits of basic analog controllers to the universal computers in digital controllers. Just where to draw the line between simple and complex may not be immediately obvious.

Other ambiguous cases would arise for natural systems such as the climate-change example shown in Figure 1.2b. Consider, too, a biophysical example, the various functions of a cell. Some actions – sensing the environment, deciding which gene-expression pathway to activate, translating the appropriate proteins – might end up be classified as simple. And the brain, as we have already noted, is complex in our terminology. But what about the metabolic control system, which is fiendishly complex? And immune response? Where to draw the line? Conceptually, the requirement to be a Turing machine seems reasonable, but understanding whether a given system is complex may not be an easy task.

1.3 Control and Information

We already mentioned briefly the recent interest in links between control theory and *information theory*. Historically, the two subjects grew up together. Wiener's book on cybernetics and Shannon's major papers on information theory both appeared in 1948, and the two authors knew each other well. Yet, until recently, there were very few explicit connections between the two topics. We will develop some of these connections in Chapter 15, but it is useful to summarize even now the heuristic perspective that comes with this view. In the information-theoretic version of control, the quality of control is limited by the information available about the system of interest. Control then consists of gathering information, making a decision about what to do, and then acting. How to act? A basic principle is

Exploit what you know; learn what you can.

These eight words summarize most of the field. If this book is long, it is because what you know or can learn and what is worth doing with that knowledge differ greatly from situation to situation. This book is a kind of catalog of situational responses incorporating these guiding ideas.[12]

In the next chapter, we begin with a discussion of continuous-time dynamical systems, which will be the focus of much of our discussion. In part a review, it introduces some differences in the way engineers, as opposed to physicists, view such systems. Read it even if you know the general topic. Then, in Chapters 3 and 4, we present core elements of classical and modern control theory. These chapters form the core of a brief course on control. Later chapters will develop more sophisticated and more specialized points.

1.4 Notes and References

The first historical book on control was published by Otto Mayr in German in 1969 and in English translation a year later (Mayr, 1970). At the time, feedback was viewed as largely a twentieth-century. development, with a few antecedents such as Watt's governor; Mayr showed that it had ancient historical roots. Mayr's work was summarized by Bennett (2002), who "[took] up the story where Mayr left off" and wrote a two-volume history of control covering 1800–1930 (Bennett, 1979) and 1930–1955 (Bennett, 1993). The water-clock description comes from Mayr, but the sketch is based on Lepschy (1992), who also gives a quantitative analysis of the clock dynamics and a description of medieval mechanical clocks, too.

[12] You might argue that it is better to learn before exploiting. But learning has a cost: Typically, we start with prior knowledge and see how far we get. Only when simple, generic solutions fail do we need to learn more.

Unfortunately, the three books by Mayr and Bennett are all out of print and relatively hard to find. Two useful and brief historical summaries are by Bennett (1996) and Bissell (2009). The presentation I give here also draws from a review by two influential control theorists (Åström and Kumar, 2014) that includes more modern developments. (The term "hidden technology" is due to Åström.) The review on quantum feedback by Zhang et al. (2017) has many good insights in its brief sections on history and on classical control.

The classification into periods given is a mixed scheme taken from Bennett (first four periods) and Åström and Kumar for the "contemporary" period, although the name is my own coinage. I also increased the overlap in dates and reclassified some material. For example, the work of Wiener and Kolmogorov on stochastic systems occurred prior to and during World War II but did not have an impact until well after.

The description of homeostasis and the view of life as a hierarchy of regulated systems and death as a failure cascade is taken from a beautiful, poignant essay by Siddhartha Mukherjee (2018) on the death of his father. Although the term "homeostasis" was coined by Walter Cannon in the 1920s, the notion itself was discussed in the 1860s by Claude Bernard, in work that had no influence until the early twentieth century (Gross, 1998).

The Harold Black epiphany that led to the negative-feedback amplifier is recounted in Black (1977), but David Mindell notes how, like most "origin myths," it gives short shrift to its precursors and the many contributions required to abstract the problem of "signals" well enough to formulate such ideas (Mindell, 2002). Mindell also presents a useful "prehistory" to *Cybernetics* (Wiener, 1961, originally published in 1948), drawing attention to the "human in the loop" in instrumentation developed in American industry (Ford Instrument Company, Sperry Gyroscope, Bell Telephone Laboratory) and how these control systems fed into US military efforts prior to and during World War II (Mindell, 2002).

For a review of control applications in the automotive industry, see Cook et al. (2007). The discussion of climate change is from Hansen et al. (2017). A key paper in the development of systems biology and control is Hartwell et al. (1999). Ingalls (2013) discusses how to model control of the metabolic system. And Germain (2001) reviews how similar ideas connect to the immune system. McNamee and Wolpert (2019) review, from the perspective of control, the hypothesis that animals (especially humans) use internal models of the world to make rational inferences about sensory data and plans for action. Tang and Bassett (2018) discuss efforts to control dynamics in brain networks.

The note at the end of Section 1.1.4 on the "negligible" impact of academic control methods prior to the development of robust control is from Morari and Zafiriou (1989). For more discussion, see Chapter 9 and also Åström and Kumar (2014). The discussion of goals is adapted (and then slightly altered) from a discussion by Fradkov (2007), Chapter 2. The notion of feedback without loops has been emphasized (from a slightly different point of view) by Jacobs (2014). Finally, the discussion of negative and positive feedback is from Sepulchre et al. (2019).

O'Keeffe et al. (2017) contrast the notions of synchronization (temporal correlation) with swarming (spatial correlation) and go on to introduce dynamical systems with spatiotemporal correlations, which they dub *swarmalators*.