

Applications of Kalman Filtering in Aerospace 1960 to the Present

MOHINDER S. GREWAL and ANGUS P. ANDREWS

Nothing speaks more eloquently of the impact of Kalman filtering on society than the 2008 Draper Prize awarded by the National Academy of Engineering to Rudolf Emil Kalman “for the development and dissemination of the optimal digital technique (known as the Kalman filter) that is pervasively used to control a vast array of consumer, health, commercial, and defense products.” To quote further from the announcement of the Awards Committee, “The Kalman filter revolutionized the field of control theory and has become pervasive in engineering systems.” The latter statement is especially pertinent within the aerospace community.

In the 1960s, the Kalman filter was applied to navigation for the Apollo Project, which required estimates of the trajectories of manned spacecraft going to the Moon and back. With the lives of the astronauts at stake, it was essential that the Kalman filter be proven effective and reliable before it could be used. This article is about the lead up to Kalman’s work, key discoveries in the development and maturation of the filter, a sampling of its many applications in aerospace, and recognition of some who played key roles in that history.

KALMAN FILTER DEVELOPMENT

The development and early applications of the Kalman filter occurred during the Cold War between the Soviet Block and the North American Treaty Organization. Early Soviet triumphs in aerospace technology leading to the 1957 launch of the first artificial satellite and manned space launches starting in 1959 convinced U.S. leadership of the need for improving aerospace technology in the United States. The federal government increased funding for research and encouraged aerospace companies to devote more of their efforts to advanced research.

The Research Institute for Advanced Studies

In 1955, Glen L. Martin Company Vice President George Trimble sought to establish a center for advanced research comparable to what was then being done in the Soviet Union [1]. G. Trimble appointed electronics and propulsion control engineer Welcome Bender as director [2]. In the next few years, W. Bender established the Research Institute for

Advanced Studies (RIAS) in a Baltimore suburb and recruited its senior management and staff. He recruited Lou Whitten from Johns Hopkins University, who hired Robert W. Bass in June of 1956 to help in selecting and recruiting technical staff. R.W. Bass had completed the Ph.D. at John Hopkins in 1955 under Fields Medalist Solomon Lefschetz and published seminal contributions in control theory in 1956 [2]. R.W. Bass first met R.E. Kalman in March of 1956 at an ASME meeting hosted by Johns Hopkins University [1].

S. Lefschetz came to the United States after graduating from École Centrale in Paris in 1905. He taught mathematics to junior staff at Westinghouse Electric Company after he had lost both hands and forearms in a transformer explosion in the laboratory in 1907. In 1910 he enrolled in the graduate mathematics program at Clark University, where he was a classmate of rocket pioneer Robert H. Goddard. S. Lefschetz was a visiting professor at Princeton in 1924 and became a professor of mathematics there in 1925. He was thesis advisor to Richard E. Bellman, among others. L. Whitten approached S. Lefschetz to join RIAS in 1956, but S. Lefschetz declined. In 1957, when S. Lefschetz was a Princeton emeritus professor, R.W. Bass and L. Whitten finally persuaded him to head up the mathematics and control groups at RIAS.

The Work of R.E. Kalman and R.S. Bucy

On the recommendation of R.W. Bass, S. Lefschetz hired R.E. Kalman in 1957 [1]. R.E. Kalman in turn recommended Richard S. Bucy, who joined him at RIAS in 1958 [1]. Soon after that, they were funded by the Air Force Office of Scientific Research (AFOSR) to perform basic research for estimation and control of aerospace systems [3].

In November of 1958 R.E. Kalman conceived the idea of recasting in state-space form the optimal estimation methods developed by Norbert Wiener and Andrei N. Kolmogorov in the 1940s. These earlier methods had been derived in the frequency domain, using power spectral densities to characterize statistical properties of the random processes involved. R.E. Kalman and R.S. Bucy reformulated the problem in the time domain, using a generalized state-space form of a linear differential equation introduced in 1908 by Paul Langevin to model Brownian motion, the erratic motion of minute particles in fluids reported by Robert Brown in 1827. Mathematical models for Brownian motion

had been studied by many, including Louis Bachelier in 1900, Albert Einstein in 1905, and Norbert Wiener between 1920 and 1924. P. Langevin's differential equation included as a nonhomogeneous term a zero-mean random process uncorrelated in time. A compatible integral for white noise processes would not emerge until Kiyoshi Itô presented a version of the stochastic calculus in 1961, but R.S. Bucy and R.E. Kalman were able to proceed without it.

The Riccati Equation

According to R.W. Bass, who was at RIAS at the time, it was R.S. Bucy who recognized that the Wiener-Hopf equation used to derive the Wiener filter is equivalent, under the assumption that the state-space model is finite dimensional, to a nonlinear ordinary differential equation studied by Italian mathematician Jacopo F. Riccati around 1720. J.F. Riccati had shown how a certain nonlinear differential equation (now called the *Riccati equation*) could be transformed into a system of linear equations.

The dependent variable in the Riccati differential equation of the Kalman-Bucy filter is the covariance matrix of the *estimation error*, defined as the difference between the estimated state vector \hat{x} and the true state vector x .

R.E. Kalman had studied the behavior of linear dynamic systems in discrete time for his 1954 master's thesis at MIT. In 1959 he recast the linear stochastic dynamic model of the Riccati equation from continuous time to discrete time and derived the optimal linear feedback gain for an estimator. The derivation requires only the vector mean and covariance matrix of the underlying probability distribution, and the optimal gain computation requires only the solution to the matrix Riccati equation. The resulting estimator is now called the *Kalman filter*. The analogous filter in continuous time is called the *Kalman-Bucy filter*.

R.E. Kalman published the discrete-time filter in a mechanical engineering journal in 1960 [5] and (with R.S. Bucy) the continuous-time filter in 1961 [6]. In the meantime, physicist Peter Swerling had derived an equivalent formulation of the Kalman filter and applied it to the problem of estimating the trajectories of satellites using ground-based sensors [7]. His results were published in an astronomy journal the year before [5] appeared. P. Swerling had derived the equivalent Riccati equation from the model for a recursive least-mean-squares estimator. The main differences are that R.E. Kalman's approach makes fewer assumptions about the underlying stochastic process model, and the resulting formula for the observational update of the Riccati equation has better numerical stability in computer implementations.

APPLICATION TO THE APOLLO MOON PROJECT

The Role of Stanley F. Schmidt

When [5] was published, S.F. Schmidt, who was chief of the Dynamic Analysis Branch at the National Aeronautics

and Space Administration (NASA) Ames Research Center (ARC) in Mountain View, California [8], invited R.E. Kalman to present his results at ARC and subsequently visited R.E. Kalman at RIAS. S.F. Schmidt felt that the Kalman filter might provide the solution to a problem his organization was tasked to solve, namely, the trajectory estimation and control problem for sending astronauts to the Moon and back. The effort at ARC became part of the Apollo Project after its announcement by President John F. Kennedy on May 25, 1961.

By early 1961, S.F. Schmidt's group had implemented the Kalman filter in digital computer simulations of the circumlunar navigation problem. S.F. Schmidt realized that the filter could be divided into two distinct parts, with one part for time periods between sensor outputs and another part for incorporating measurements. This partitioning of the estimation problem was advantageous for the Apollo mission because hours could elapse between sensor outputs.

The effort at ARC accomplished several key steps in the development of the Kalman filter as a practical method for real-time onboard navigation in the Apollo mission. It proved that the Kalman filter worked as predicted and that it was capable of solving the Apollo guidance and navigation problem on a mainframe computer with 36-bit floating-point arithmetic. The ARC group used Monte Carlo analysis of the problem to prove that estimation uncertainty was modeled with sufficient accuracy by the matrix Riccati equation. Additional key steps were the development of what is now called the *extended Kalman filter* (EKF, described below) and the use of Monte Carlo analysis to show that nonlinearities of the trajectory model did not compromise the accuracy of the EKF for the ranges of errors expected during the Apollo missions.

The solution of the matrix Riccati equation was found to provide a quantitative measure of how well the state variables can be estimated in terms of mean-squared estimation errors. This result proved to be much more practical and useful than the concept of observability from least-squares methods, which provided only a qualitative indication of whether or not the unknown variables are uniquely determinable from the data. As a consequence, the matrix Riccati equation from the Kalman filter was soon recognized as a practical model for predicting the performance of sensor systems, and it became the standard model for designing aerospace sensor systems to meet specified performance requirements.

The work at ARC uncovered numerical stability problems in solving the matrix Riccati equation in finite precision arithmetic and showed that this behavior could be influenced by changing the order of computing associative matrix products such as $A(BC) = (AB)C$. This insight eventually led to major improvements in solution methods for matrix Riccati equations.

S.F. Schmidt was quick to share his results with others, including contacts at Lockheed Missiles and Space

Company in nearby Sunnyvale, California, and Richard H. Battin at the Instrumentation Laboratory (now the Charles Stark Draper Laboratory) of MIT. In 1961, MIT won the NASA contract for developing an onboard guidance and navigation system for the Apollo spacecraft. Kalman filtering became an integral part of the Apollo navigation system from that point forward [9].

Extended Kalman Filter

Much of the earlier work on the estimation and control of space trajectories used linearized approximations for small perturbations from nominal trajectories to linearize the guidance and navigation problem, an approach that worked well in most applications. In S.F. Schmidt's early evaluations of the Kalman filter, his group evaluated the effects of nonlinearities by using Monte Carlo simulations of the true trajectory, which is not precisely known during the actual missions, as well as the estimated trajectory. These studies demonstrated that the filter could achieve excellent performance by linearizing the problem about the estimated trajectory, an approach now called *extended Kalman filtering*. Using extended Kalman filtering, the estimated trajectories could be shown to converge to the true trajectories, even with unusually large initial trajectory estimation errors.

Apollo Onboard Trajectory Estimation

The navigation and guidance system developed at MIT included an onboard inertial navigator for measuring accelerations during thrusting periods as well as an onboard optical space sextant for measuring angles between stars and points on the Earth and Moon.

Trajectory estimation and control was critical for the Apollo Project. As illustrated in Figure 1, the outbound trajectory to the Moon must have its closest approach on the far side of the Moon at a height of about 97 km (60 mi) above the surface so that the spacecraft can efficiently transfer to a circular orbit about the Moon at that altitude. However, the Moon is a moving target, traveling at about 1 km/s.

The return trajectory back to Earth is even more critical. The Apollo Command Module with its crew of three must enter the atmosphere at a carefully controlled descent angle for atmospheric braking. If the descent angle were too steep, the Command Module would burn up; too shallow, and it would skip back out into space.

The EKF was used at ARC and MIT in ground-based simulations to determine the required accuracies of the Apollo space sextant and inertial navigator and to design observation schedules for using the onboard space sextant to satisfy mission navigation requirements without overtaxing the three-man crew. Meanwhile, Joseph P. O'Malley at the Space and Information Systems Division of North American Aviation, the prime contractor for the Apollo Command and Service Modules containing the astronauts

and supporting systems, had independently derived and used the Riccati equation implementation of P. Swerling [7] for estimating propellant requirements for midcourse trajectory corrections on the Apollo missions.

James E. Potter and Square Root Filtering

Apollo trajectory simulations at ARC and MIT were mostly done on IBM 704 mainframe computers and later models with similar arithmetic processing. These simulations used 36-bit floating point arithmetic, which was adequate for trajectory simulations but marginal for implementing the Riccati equation solution for the measurement updates in the Kalman filter. Performance was not reliable in 36-bit floating point arithmetic, and the Apollo flight computer would have to implement the Kalman filter in 15-bit fixed-point arithmetic. Microprocessors were still a long way off.

J.E. Potter was at that time a graduate student in mathematics at MIT, working part time at the Instrumentation Laboratory on the Apollo Project. He took the problem home with him one Friday afternoon and arrived back on Monday with a solution. The trick was to use a Cholesky factor of the covariance matrix as the dependent variable in an equivalent Riccati equation. The solution was brilliant, and the improvement was profound. The approach came to be called *square-root filtering*, and alternative implementations of square-root filters with better numerical stability were soon discovered. The more computationally efficient implementations by Neil A. Carlson, Gerald J. Bierman, and Catherine Thornton use triangular Cholesky factors of the covariance matrix of estimation uncertainty [4]. G.J. Bierman once suggested a rule of thumb for the improvement of square root filtering over conventional Kalman filtering as "the same accuracy with half as many bits" of precision.

J.E. Potter's implementation of the Kalman filter was programmed into the Apollo Guidance Computer, designed by an MIT team led by Eldon Hall, and executed in 15-bit arithmetic to navigate to the Moon and back [10].

J.E. Potter's method solves only one part of the Riccati equation, namely, modeling the effects of measurements on trajectory estimation uncertainty. The rest of the Riccati equation models the effect of random disturbances, which were insignificant for the Apollo navigation problem.

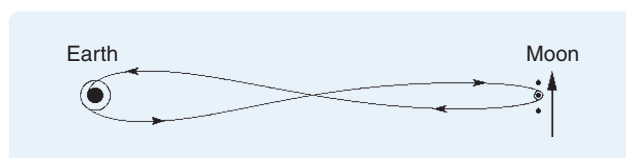


FIGURE 1 Nominal Apollo trajectories, drawn approximately to scale in nonrotating coordinates. The trajectory control problem is further complicated by the orbital velocity of the Moon around the Earth at about 1 km/s. The Moon completes nearly a quarter of its orbit around the Earth during the mission to the Moon and back.

Extensions of the square-root solution to the Riccati equation to account for the effects of disturbances evolved in the 1960s and 1970s [4].

Kalman Filter Tuning

Applications of the Kalman filter for the Apollo Project required initial values for the covariance matrix of estimation uncertainty and the estimated system state vector. These implementations also required covariance matrices for sensor output noise and for random dynamic disturbances of the system state vector. In many of these applications, the values for these parameters had to be determined from test results, error analysis, and engineering judgment. The process of adjusting these parameter values, based on analysis of filter performance, is called *filter tuning*. Engineering judgment, for example, often assigns conservative values to the initial covariance matrix of estimation uncertainty, with variances somewhat larger than those derived from analysis. This tuning has been shown to improve the filter response time [11]. Although sensor noise covariance matrices are often provided by the sensor manufacturers, obtaining suitable values for the covariance matrix of dynamic disturbance noise generally requires some iterative tuning based on observed filter performance.

APPLICATIONS TO INERTIAL NAVIGATION SYSTEMS

Inertial Navigation

Although rocket pioneer R.H. Goddard and the Peenemunde rocket scientists used inertial sensors for navigation and control of missiles, a complete navigation system using inertial sensors did not emerge until the 1940s under Charles Stark Draper, considered to be “the father of inertial navigation.” C.S. Draper established the Instrumentation

Laboratory at MIT as a major player in the early development of inertial navigation. In the 1960s, engineers at MIT designed the inertial navigation system (INS) for sensing and controlling rocket thrusting during trajectory changes of the Apollo spacecraft [12]. The dominant inertial sensor errors for the Moon missions were unpredictable shifts in output biases of the gyroscopes and accelerometers. These shifts could be corrected by observing accelerometer outputs during 0-g coasting periods and observing the buildup of attitude errors using the space sextant, which was mounted to a common base with INS, to observe the known directions to reference stars. These corrections did not require a Kalman filter. However, once the Kalman filter was introduced in the Instrumentation Laboratory, it was used to calibrate inertial sensors (gyroscopes and accelerometers) for all applications of inertial navigation, and thus the Riccati equation became a standard model for relating inertial sensor performance to inertial navigation performance. In the 1960s, Thomas L. Gunckel II brought Kalman filtering from Stanford University to Autonetics, another major developer of high-precision inertial navigation. The Kalman filter soon became the dominant tool for the design and implementation of inertial navigation systems throughout the industry and for the integration of inertial navigation systems with other sensors. S.F. Schmidt played a major role in the integration of an INS and radar system on the C-5A aircraft [8], the first airborne implementation of a Kalman filter for sensor integration.

Modeling INS Error Dynamics

INS errors behave differently when navigation is implemented in Earth-fixed coordinates, due principally to the curvature and vertical gradient of the gravitational field in the near-Earth environment. Vertical navigation errors in position and velocity tend to grow exponentially, and horizontal errors tend to induce oscillations with a period equal to that of a satellite in a circular orbit at the same altitude as the INS. When coupled with Earth rotation, the horizontal navigation errors behave like a Foucault pendulum with that period, as illustrated in Figure 2. These error dynamics can be represented by a linear state-space model. The pendulum period, which is about 84.4 min at sea level, was discovered in 1923 by Maximilian Schuler in his analysis of gyrocompassing errors [13]. This period is called the *Schuler period* [12].

Some sort of barometric or radar altimeter is needed to stabilize vertical errors. The accuracy of inertial navigation can also be improved significantly if the Schuler oscillations

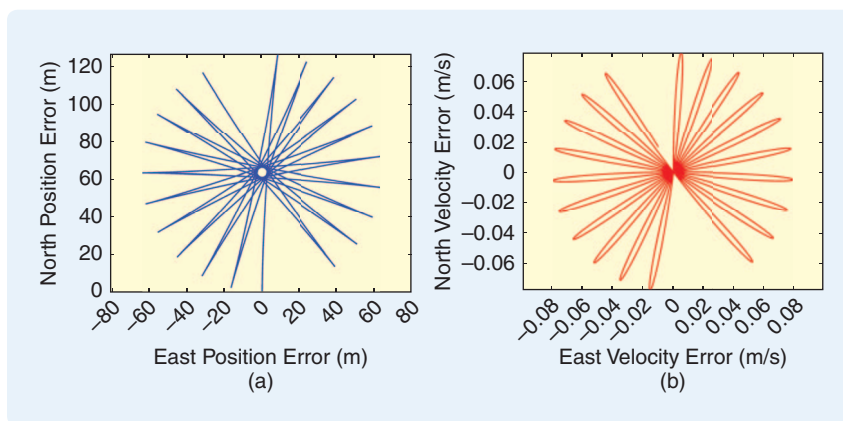


FIGURE 2 Inertial navigation error Schuler oscillations in (a) position and (b) velocity caused by 10^{-5} g output offset error from the north-pointing accelerometer. During automatic attitude alignment of the inertial navigation systems (INSs), the underlying linear dynamic model is used in a Kalman filter to detect and correct for such errors. To illustrate the Schuler oscillation patterns, the simulated data represent about 10 h of INS operation without any other error sources. The time required for alignment is typically on the order of 10 min.

can be detected and corrected using auxiliary position, velocity, or acceleration sensors. For this purpose, early inertial systems used star trackers to measure tilt errors and velocity sensors to detect Schuler oscillations. Velocity sensors include water speed sensors for ships, and airspeed sensors or Doppler radar for aircraft. Except for applications where Global Navigation Satellite System (GNSS) signals are inaccessible (underground or in submarines, for example), most present-day INS applications use GNSS receivers as low-cost auxiliary sensors.

Applications of Kalman filtering in inertial navigation include sensor calibration, INS alignment, and the detection and compensation of INS navigation errors.

In sensor calibration, inertial sensor errors are estimated and compensated prior to navigation. This task requires control of the sensor inputs, such as rotation rates or acceleration components, while the sensor outputs are measured [14].

In INS alignment, attitude is estimated from auxiliary information. INS alignment aboard parked aircraft or docked ships, for example, uses stochastic models for the random dynamic disturbances of the host vehicle. In similar fashion, alignment of an INS in a carrier-based aircraft or airborne missile can use the outputs from the INS in the host vehicle during maneuvers prior to launch in a process called transfer alignment.

Detection and compensation of INS navigation errors uses auxiliary sensors such as INS-based star trackers for attitude estimation or radio-navigation aids for position estimation.

Integrated Calibration Implementations

The simultaneous calibration, alignment, and error correction of complex inertial navigation systems has been implemented using Kalman filters. By augmenting the state vector of the navigation error model with the unknown calibration parameters of the gyroscopes and accelerometers, the parameter estimation problem becomes part of the state-estimation problem. A schematic of this implementation for a gimballed INS is shown in Figure 3. To reduce the computational load, prefiltering in the form of data compression by measurement averaging can be implemented with minimal degradation in the performance of the filter [14]. This approach can be made to work with fairly complex models for the gyroscope and accelerometer errors, provided that the system excitation trajectories of attitude and acceleration during calibration can be designed to provide

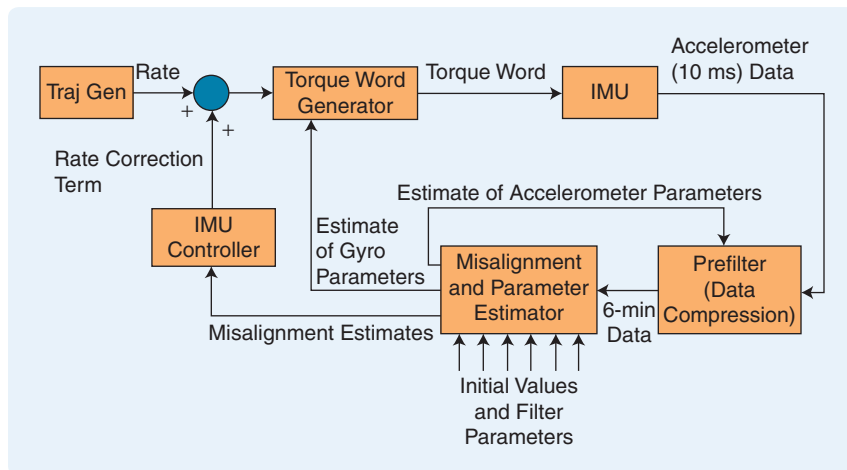


FIGURE 3 Integrated control, calibration, and alignment implementation for a gimballed inertial navigator. This diagram shows that data from the accelerometers at 10-ms intervals are prefiltered and compressed so that each Kalman filter update can be completed in 6 min.

observability of the sensor parameters [14]. This approach was used for INS calibration on the F-111, Minuteman (I, II and III), and Peacekeeper (Missile X) programs.

RADAR TRACKING

Funding for the development of the Wiener filter in the 1940s was directed toward radar tracking of aircraft for automatic control of antiaircraft guns. These tracking filters were used when computers were integrated with radar systems in the 1950s, but more sophisticated Kalman filter models for aircraft tracking came into use in the 1960s. As computer technology matured, EKFs for radar tracking would migrate from ground-based radars, to airborne radars, and eventually to radar tracking systems on missiles.

NASA was formed in 1958, the same year the United States launched its first satellite, Explorer I. This satellite was designed by the U.S. Army's Jet Propulsion Laboratory (JPL), managed by the California Institute of Technology. The army had contracted with JPL to set up portable radar installations around the world to track Explorer I. JPL and the radar project were transferred to NASA when it was formed later in 1958. The radar tracking network would evolve into the NASA Deep Space Network, which was used to track the Apollo spacecraft and many generations of space exploration missions. Extended Kalman filtering became an integral part of this spacecraft tracking system, as well as most radar tracking systems in aerospace applications.

Real-time, online applications in missile defense include estimating and predicting reentry vehicle (RV) position from ground-based radar data. Early testing at White Sands Missile Test Range led to better methods for detection and tracking of RVs designed with low observability to deter acquisition and tracking. Similar radar tracking methods

were developed for Shuttle Orbiter Operations to estimate vehicle trajectories as an aid to ground-based personnel in establishing and maintaining mission control.

Offline applications include estimation and correction of radar errors such as azimuth bias, elevation bias, and survey (base location) errors. This approach is used in the United States on the Eastern Test Range, White Sands Missile Range, and Western Test Range. These techniques have also been used to estimate the trajectories of thousands of Earth satellites and space debris as well as for verifying the accuracy of onboard aircraft navigation systems by comparison with Kalman filtered data from ground-based sensors.

GLOBAL NAVIGATION SATELLITE SYSTEMS

The first operational GNSS, the Global Positioning System (GPS), was put into service in 1993 by the U.S. Department of Defense. Kalman filtering played a critical role in the development of the GPS, which has been described as “one enormous Kalman filter” [15]. This Kalman filter has a large system state vector, including the trajectories of the 24+ satellites, the drift rates and phases of all system clocks, and hundreds of parameters related to atmospheric propagation delay as a function of time and location. The GPS uses precision receivers all over the world as sensors for estimating these variables. In addition, every GNSS receiver uses an EKF to estimate its own position and velocity, and to synchronize the receiver clock with GPS time.

GNSS/INS Integration

There are several methods for integrating GNSS receivers with INS. Some of the simplest are called *loosely coupled* integration methods, which use GNSSs as an auxiliary position sensor to detect and correct INS navigation errors, the same way an altimeter is used to stabilize vertical errors. On the other end of the spectrum, some of the more tightly coupled implementations use the GNSS-determined *pseudoranges* from the satellite antennas to receiver antennas as sensor measurements to estimate receiver clock errors and time-varying parameters of the inertial sensors.

GNSS Augmentation

Currently, four space-based GNSS augmentation systems are under development worldwide. These systems include the Wide Area Augmentation System (WAAS) in the United States, the European Geostationary Navigation Overlay System (EGNOS) in the European Union, the Multifunctional Transport Satellite (MTSAT) Based Augmentation System (MSAS) in Japan, and Geostationary (GEO) and GPS Augmented Navigation (GAGAN) in India. These augmentation systems are intended to eventually replace current ground-based air traffic control by providing seamless navigation of civil aviation using signals from satellites for departure, en route, arrival, and approach operations. All of these systems depend heavily on Kalman filtering.

WAAS is also used to support Federal Aviation Administration (FAA) safety, capacity, and efficiency initiatives. These initiatives are designed to provide more efficient use of airspace, improved flight operation procedures, and better situational awareness during aircraft ground operations. Augmented GPS is the primary radio navigation system for the foreseeable future. Phase 1 of WAAS, approved by the FAA, has been used since 2003.

Differential corrections for improving the accuracy and integrity of GNSS signals are being developed for use in correction and verification (C&V) processors, which implement nonlinear Kalman filters. Some of these processors track GNSS and geostationary satellite orbits and clock estimation errors using the carrier phase smoothed pseudoranges. These Kalman filters use 11 state variables [15].

Another set of C&V Kalman filters uses data from the 1575.42 MHz L1 signal carrier and 1227.6 MHz L2 signal carrier from the GPS satellites to estimate the global distribution of ionospheric propagation delays and differential phase biases of the L1 and L2 carriers in receivers and satellites. The state variables in these Kalman filters include delays at ionospheric grid points on a 5×50 partitioning of the northern and southern hemisphere, estimated using pseudoranges to satellites visible at the wide reference station (WRS) ground receivers. Two independent Kalman filters are used to estimate the ionospheric delay at the grid points and to estimate the receiver and satellite differential biases. Each of these filters uses 250 state variables.

The resulting corrections and integrity estimates are uplinked through GEO satellites to the user receivers. This system, called the GEO Communication and Command Subsystem (GCCS), contains two Kalman filters, one for GEO ionospheric delay and delay rate estimation and one for GEO range, range rate, and range acceleration estimation. The onboard GNSS receivers may also be integrated with inertial navigation systems, using another Kalman filter [15].

PUSHING THE ENVELOPE FOR NONLINEAR APPLICATIONS

Limits of Extended Kalman Filtering

Extended Kalman filtering has been successful for many applications in which the errors introduced by linear approximations are insignificant compared to the modeled errors due to measurement noise and dynamic disturbance noise, a condition called *quasi-linearity*. The Apollo navigation problem and the GPS receiver filtering problem are quasi-linear in this sense. The following are some additional nonlinear problems that have been found to be quasi-linear for the expected ranges of linear approximation errors.

Trilateration

Trilateration is the problem of determining the relative geometric locations of objects, given only the distances measured

between them, including cases where only the distances to the nearest few objects are measurable. A pre-GPS application called *relative navigation* [16] used the Joint Tactical Information Distribution System (JTIDS), a secure communications system developed for the U.S. military. JTIDS implementation requires tight synchronization of the transceiver clocks, the relative phasings of which are used to determine the relative ranges between all pairs of transceivers. The solution, using extended Kalman filtering, requires additional information to orient the network relative to Earth.

Magnetic Detection and Location

A nonlinear detection and estimation problem arises in antisubmarine warfare using airborne magnetometers or magnetic gradiometers to measure detectable disturbances of the background magnetic field caused by the presence of a submarine. The objective in this case is to detect and locate submarines using the magnetic sensor signals on-board low-flying aircraft. A single sample of magnetic field components at a fixed location has insufficient information for locating the dipole, but three components of dipole location as well as the three components of its dipole moment are observable from a sequence of sensor measurements made along a search trajectory past the magnetic source, as illustrated in Figure 4. This signal filtering and detection problem was solved in the frequency domain by John E. Anderson in 1949 [17]. A time-domain solution to the detection and tracking problem [18] uses a signal detection method developed by Fred. C. Schweppe in 1965 [19], based on the Kalman filter model.

Essentially the same dipole detection and tracking problem was solved in an approach to automatic steering of vehicles developed by Robert Parsons and Wei-Bin Zhang at the University of California at Berkeley in the late 1980s [20]. In this application, magnetic “nails” (dipoles) are inserted in the intended vehicle pathway several meters apart along a line down the middle of the pathway. Three-axis magnetic sensors mounted near the front and bottom of the vehicle body are used to measure the magnetic field disturbances caused by the magnetic nails, and the sensor signals are used in an EKF to estimate the instantaneous location of the vehicle relative to the magnetic nails. The automated steering control implementation uses the estimated relative location to calculate lateral steering error. This EKF implementation is capable of estimating the lateral location of the vehicle relative to the dipoles within an RMS uncertainty of about 1 mm [18]. The associated Schweppe likelihood ratio detector [19] can be used to determine the orientations of the individual dipole moments so that dipole orientations can be used to encode information for enhancing steering control, such as curvature of the path ahead.

Adaptive Kalman Filtering

In some applications of Kalman filtering, parameters of the Kalman filter model are either unknown or time

varying. For example, the variance of a noise source in flight can vary from the value measured in the laboratory. In some cases, the variation is large enough to degrade performance but small enough that the variance can be included as a state variable in an EKF. This approach is called *adaptive Kalman filtering*, because the resulting filter adapts to small changes in model parameters, so long as the nonlinearities do not significantly corrupt filter performance [21], [22].

Sample-Based Methods for Nonlinear Applications

Failures of Extended Kalman Filtering

The original derivation of the Kalman filter [5] does not require that the underlying error distributions be Gaussian, even though linearity of the Kalman model preserves Gaussianity of distributions. There was some concern when Monte Carlo simulations of the more nonlinear applications using extended Kalman filtering produced severe distortions of initially Gaussian distributions. More importantly, however, it was found that the EKF did not always accurately propagate the true mean and covariance of the distributions of state variables and measurement variables.

When the measurement variables are nonlinear functions of the state variables, errors can arise in the predicted measurement and in the calculated covariance of its uncertainty. When the differential equation for propagating the state variables is nonlinear, errors can arise in the predicted

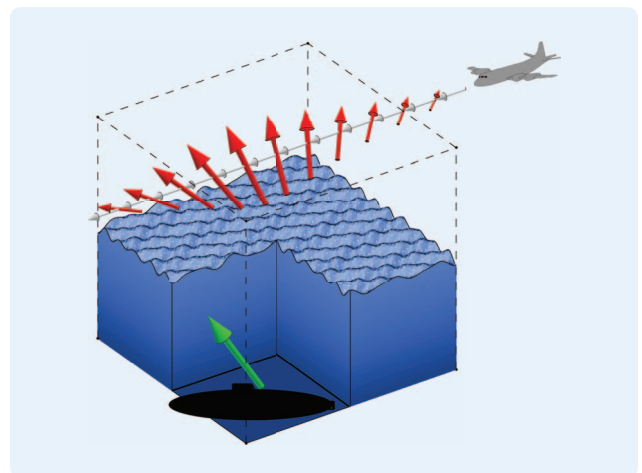


FIGURE 4 Geometry of the airborne magnetic detection and location problem. Three-axis magnetic-field sensors are housed in the aircraft tail extension to isolate them from aircraft noise. The magnetic field vector (red) is sampled at intervals, and an extended Kalman filter is used to detect field perturbations due to any submerged magnetic dipole (green) in the presence of the background geomagnetic field and harmonic noise due to sea-surface wave motion. The relative location and magnetic moment components of the submerged dipole are part of the state vector. The magnitude of the estimated dipole moment vector can be used as a threshold for detection.

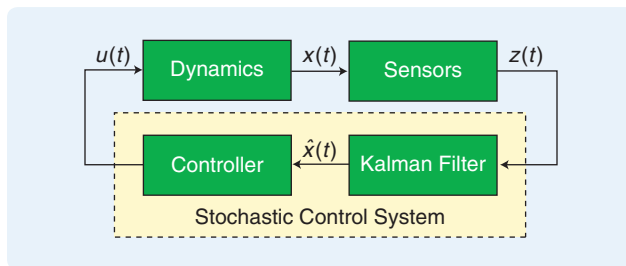


FIGURE 5 Observer-based controller. This block diagram shows the Kalman filter as an observer for the state vector $x(t)$, including the unmeasured state variables needed to generate the control $u(t)$. The measurement vector is $z(t)$.

state vector as well as in the covariance matrix of state-estimation errors prior to using the measurement.

Sample-Based Methods

More recently, significant estimator performance improvement for nonlinear applications has been realized by nonlinear propagation of a finite number of representative samples from the modeled probability distributions of state variables and measurement variables. The samples are chosen to represent the means and variances before the nonlinear transformations, and the resulting transformed samples represent the means and covariances after the nonlinear transformations.

The parts of the Kalman filter that can be implemented using these methods include the propagation over time of the estimated state variable and its associated covariance of uncertainty and computation of the expected measurement and its associated covariance matrix of uncertainty, all the variables needed for computing and applying the Kalman gain. As a result, sample-based methods eliminate the need for a Riccati equation to calculate the Kalman gain.

Sampling Strategies

Unlike the original Monte Carlo methods, which used large numbers of pseudorandom samples to approximate a probability distribution, newer methods use much smaller sample sizes for representing only the means and covariances of the distribution. As a consequence, some sample-based methods have about the same computational complexity as the EKF but perform significantly better on nonlinear problems.

Various sampling strategies have been developed for nonlinear filtering. The resulting filter implementations include *sigma-point filters* [24], *unscented Kalman filters* [25], and *particle filters* [23]. The naming of sigma-point filters refers to the distribution standard deviation (represented by the symbol σ) used to calculate sample values. Unscented filters use similar strategies depending on Cholesky factors of covariance matrices. The naming of particle filtering refers to the sampled values as

“particles,” which are propagated and used to compute the means and variances.

Distributed Implementations

Particle filtering can also be used as part of *ensemble Kalman filtering*, a distributed implementation of a nonlinear estimator for applications with large numbers of state variables [26]. In distributed computing, the computational load is partitioned across several computers, with the partitioning designed to keep the interprocess communications requirements within acceptable levels. Ensemble Kalman filters have been developed for weather forecasting [27], in which the dynamics are nonlinear and the number of state variables may be too large for filter implementation on a single computer. These nonlinear filtering techniques rely to some extent on earlier distributed implementations of the Kalman filter [28], including prefiltering [14].

CONTROLLERS, OBSERVERS, AND THE SEPARATION PRINCIPLE

The work of R.E. Kalman and R.S. Bucy also had a significant impact on the implementation cost and efficacy of control. Their results on observability and controllability allowed system designers more flexibility in deciding which dynamic state variables to measure directly, and this flexibility could be used to address other system-level issues such as cost, performance, stealth, and tolerance of sensor failure or jamming.

The Kalman filter allowed control engineers to estimate and control some dynamic variables without measuring them directly, as illustrated in Figure 5. The optimal control $u(t)$ can be calculated from the estimated state $\hat{x}(t)$ generated by the Kalman filter. The ability to minimize the number of sensors required for estimation and control was especially important for aerospace applications in which system weight and power must be minimized.

In [5], R.E. Kalman proved that a certain class of estimation and control problems are duals of one another, in that one problem can be transformed into the other by an appropriate exchange of model variables. These dual techniques constitute the *linear quadratic Gaussian* (LQG) control. Peter D. Joseph and Julius T. Tou [29] proved that the LQG estimation and control problems can be solved independently, by invoking the *separation principle*. Later proofs of the separation principle require less restrictive assumptions about the models [30], [31]. Separation of the estimation and control problems often simplifies system design by allowing independent development of a system *observer* for the estimation problem and a *controller* for the system control problem.

CONCLUSIONS

The Kalman filter found early acceptance in the aerospace industry as the basis for modern estimation and control theory, not only for the theoretically optimal solution but

as a practical and reliable solution method for LQG problems. For modern estimation and control problems, the matrix Riccati equation and its sample-based equivalents have become almost universal tools for the design of the sensor and control system to meet specified statistical performance requirements, a benefit often overlooked in assessments of the impact of Kalman filtering on technology.

The problems that most bedeviled early aerospace applications of Kalman filtering were computer size, weight, and power requirements; numerical stability; computational load; and computer costs. Subsequent developments in hardware and implementation methods have overcome most of those problems, and evolving improvements are even now extending the range of applications to which Kalman filtering can be applied successfully.

ACKNOWLEDGMENTS

The authors thank Robert W. Bass, Rudolf E. Kalman, Thomas L. Gunckel II, and the late James E. Potter for their personal accounts of the early development of the Kalman filter and its applications in aerospace.

AUTHOR INFORMATION

Mohinder S. Grewal (mgrewal@fullerton.edu) received the Ph.D. in electrical engineering from the University of Southern California, Los Angeles, in 1974, with a specialization in control systems and computers. He coauthored *Kalman Filtering: Theory and Practice Using MATLAB*, third edition (New York: Wiley, 2008) and *Global Positioning Systems, Inertial Navigation, and Integration*, second edition (New York: Wiley, 2007). He has published over 60 papers in IEEE and ION refereed journals and proceedings, including the *Institute of Navigation Redbook*. He has consulted with Raytheon Systems, Boeing Company, Lockheed-Martin, and Northrop on applications of Kalman filtering. Currently, he is a professor of electrical engineering at California State University, Fullerton, where he received the 2009 Outstanding Professor award. He is an architect of the GEO Uplink Subsystem (GUS) for the Wide Area Augmentation System (WAAS), including the GUS clock steering algorithms, and holds two patents in this area. His current research interest is in the area of application of integrated GPS and INS to navigation. He is a member of the Institute of Navigation, a Fellow of the Institute for the Advancement of Engineering, and a Senior Member of IEEE.

Angus P. Andrews was a sophomore at MIT on October 4, 1957, when the Soviet Union launched the world's first artificial satellite. He completed the undergraduate degree at MIT in 1960 and the Ph.D. in mathematics at the University of California, Los Angeles, in 1968. He was introduced to the Kalman filter in 1961 by Harold W. Sorenson, a colleague at General Dynamics Astronautics Division in San Diego, and first applied Kalman filtering for performance analysis on the Apollo Project in 1962 at the Space and Information Systems Division of North

American Aviation, prime contractor for the Apollo Command and Service Modules. He later used Kalman filtering in modeling, design analysis, and implementation of the N57 and N73 strapdown inertial navigation systems at the Autonetics Division of North American Aviation, now part of the Boeing Company. He continued applying Kalman filtering technology to a variety of aerospace and commercial problems, including antisubmarine warfare, until retiring as a senior scientist from Rockwell Science Center in 2000. Since then, he has consulted for Northrop Grumman Corporation and Raytheon Corporation on the performance analysis of space-based sensing systems, and coauthored five editions of two engineering textbooks.

REFERENCES

- [1] R. W. Bass. (2002). Some reminiscences of control and system theory in the period 1955-1960: Introduction of Dr. Rudolf E. Kalman. *Real Time*. (Spring/Summer Issue), pp. 2-6. [Online]. Available: www.ece.uah.edu/PDFs/news/RT-sprsum2002.pdf
- [2] R. S. Bucy and J. T. Lo, "Seminal contributions of Bob Bass to control," in *Proc. 1998 American Control Conf.*, 1998, pp. 411-414.
- [3] W. Berry and C. Loeb, "Breakthrough Air Force capabilities spawned by basic research," Center for Technol. and Natl. Security Policy, Natl. Defense Univ., Washington, D.C., Defense & Technology Paper No. 39, Apr. 2007.
- [4] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB*, 3rd ed. Hoboken, NJ: Wiley, 2008.
- [5] R. E. Kalman, "A new approach to linear filtering and prediction problems," *ASME Trans., J. Basic Eng., ser. D.*, vol. 82, pp. 35-45, 1960.
- [6] R. E. Kalman and R. S. Bucy, "New results in linear prediction and filtering theory," *ASME Trans., J. Basic Eng., ser. D.*, vol. 83, pp. 95-108, 1961.
- [7] P. Swerling, "First order error propagation in a stagewise differential smoothing procedure for satellite observations," *J. Astronaut. Sci.*, vol. 6, pp. 46-52, 1959.
- [8] S. F. Schmidt, "The Kalman filter: Its recognition and development for aerospace applications," *AIAA J. Guid. Contr.*, vol. 4, pp. 4-7, 1981.
- [9] R. H. Battin, "Space guidance evolution—A personal narrative," *AIAA J. Guid. Contr.*, vol. 5, pp. 97-110, 1982.
- [10] E. C. Hall, *Journey to the Moon: The History of the Apollo Guidance Computer*. Washington, D.C.: American Institute of Aeronautics and Astronautics, 1996.
- [11] A. S. Morris and M. J. H. Sterling, "Model tuning using the extended Kalman filter," *Electron. Lett.*, vol. 15, pp. 201-202, Mar. 1979.
- [12] D. MacKenzie, *Inventing Accuracy: A Historical Sociology of Nuclear Missile Guidance*. Cambridge, MA: MIT Press, 1993.
- [13] M. Schuler, "Die Störung von Pendul und Kreiselapparaten durch die Beschleunigung der Fahrzeuge," *Physikalische Zeitschrift*, vol. 24, pp. 344-350, 1923.
- [14] M. S. Grewal, V. Henderson, and R. Miyasako, "Application of Kalman filtering to the calibration and alignment of inertial navigation systems," *IEEE Trans. Automat. Contr.*, vol. 36, no. 1, pp. 4-14, Jan. 1991.
- [15] M. S. Grewal, L. R. Weill, and A. P. Andrews, *Global Positioning Systems, Inertial Navigation, and Integration*, 2nd ed. Hoboken, NJ: Wiley, 2007.
- [16] W. R. Fried, "Principles and simulation of JTIDS relative navigation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-14, pp. 76-84, Jan. 1978.
- [17] J. E. Anderson, "Magnetic anomaly detection frequency responses," Naval Air Development Center, Warminster, Pennsylvania, Rep. ADCEC-47-51, Oct. 1949.
- [18] A. P. Andrews, "The accuracy of navigation using magnetic dipole beacons," *Navigation*, vol. 38, pp. 367-397, Winter 1991-1992.
- [19] F. C. Schweppe, "Evaluation of likelihood functions for Gaussian signals," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 61-70, 1965.
- [20] R. E. Parsons and W.-B. Zhang. (1988 Oct.). Lateral guidance systems requirements definition, program on advanced technology for the highway, Univ. of California at Berkeley, Research Rep. UCB-ITS-PRR-88-1 [Online]. Available: www.path.berkeley.edu/path/publications/pdf/PRR/88/PRR-88-01.pdf

- [21] R. K. Mehra, "On the identification of variance and adaptive Kalman filtering," *IEEE Trans. Automat. Contr.*, vol. 15, pp. 175–184, 1970.
- [22] R. K. Mehra, "Approaches to adaptive filtering," *IEEE Trans. Automat. Contr.*, vol. 17, pp. 693–698, 1972.
- [23] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Processing*, vol. 50, pp. 174–188, Feb. 2002.
- [24] R. van der Merwe and E. Wan, "Gaussian mixture sigma-point particle filters for sequential probabilistic inference in dynamic state-space models," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Hong Kong, 2003, pp. 701–704.
- [25] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, pp. 401–422, 2004.
- [26] P. Houtekamer and H. L. Mitchell, "Data assimilation using an ensemble Kalman filter technique," *Monthly Weather Rev.*, vol. 126, pp. 796–811, 1998.
- [27] J. L. Anderson and S. L. Anderson, "A Monte Carlo implementation of the nonlinear filtering problem to produce assimilations and forecasts," *Monthly Weather Rev.*, vol. 127, pp. 2741–2759, 1999.
- [28] Y. Gao, E. J. Krakiwsky, M. A. Abousalem, and J. F. McLellan, "Comparison and analysis of centralized, decentralized, and federated filters," *Navigation*, vol. 40, pp. 69–86, Spring 1993.
- [29] P. D. Joseph and J. T. Tou, "On linear control theory," *AIEE Trans. Appl. Ind.*, vol. 80, pp. 193–196, 1961.
- [30] M. Athans, "The role and use of the stochastic linear-quadratic-Gaussian problem in control system design," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 529–552, 1971.
- [31] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*, 2nd ed. New York, NY: Chelsea Publishing, 1987.



» APPLICATIONS OF CONTROL (continued from page 23)

Despite this improvement in precision over 24 hours, the rate variation due to tilt angle remains over timescales that are shorter than the time it takes for the tourbillon to complete a full rotation. In this case, the term $\cos(\eta)$ in (13) does not average to zero, and thus the gravity dependence remains. Therefore, the tourbillon merely masks the rate variation over long observation periods.

Historically, accurate timepieces have been associated with astronomical observatories. Indeed, astronomers such as Airy have contributed to horology [7]. During the first half of the twentieth century, a series of different rate tests was conducted by many observatories using many types of timepieces, with and without tourbillons. Results are shown in Figure 5, and they display mixed results for the tourbillon. These calculations show explicitly why this is the case. Tests of tourbillon accuracy and precision produce different results depending on the test details; specifically, on how much the timepiece is moved. Test results are different because tourbillons do not correct the variation in rate that arises from timepiece orientation. Instead, they average the source of variation [9], which may or may not improve precision depending on how frequently the orientation of the watch changes.

Another reason for the disappointing performance of tourbillons is that they can remove only part of the gravitational error. It is possible that the frictional torque acting on bearings depends on timepiece spatial orientation. Thus, for example, the friction parameter b may vary with tilt angle φ ; if so, this variation is not eliminated by the tourbillon. Bearings in high-quality timepieces are usually made of jewels, such as rubies, which are hard and durable with low-friction coefficients.

CONCLUSIONS

Given the patchy performance of tourbillons, it is reasonable to ask why they are much sought after. Tourbillon wristwatches command astonishing prices; see "Two Centuries of Tourbillons." The answer lies outside the technical realm of physics and engineering. Tourbillons are fascinating to

observe, and the skill of the watchmaker is clearly on display. The elegance of tourbillon timepieces, combined with their expense and exclusivity due to the extreme difficulties of construction, have made these intriguing mechanical devices desirable since their invention 200 years ago.

AUTHOR INFORMATION

Mark Denny earned a Ph.D. in theoretical physics from Edinburgh University, Scotland, and then pursued research at Oxford University from 1981 to 1984. He was subsequently employed by industry, where he worked as a radar systems engineer. He has written 40 papers on radar signal processing and physics, plus four popular science books. He is semi-retired and lives on Vancouver Island.

REFERENCES

- [1] Single axis tourbillon movement [Online]. Available: http://www.youtube.com/watch?v=q_bQsv5l36k&feature=related
- [2] Assembly of escapement components, and their simulation [Online]. Available: http://www.youtube.com/watch?v=8RZZ7jql-A&feature=fv_w
- [3] Multiple axis tourbillon movement [Online]. Available: <http://www.youtube.com/watch?v=AWdlPMQlig4&feature=related>
- [4] E. Bruton, *The History of Clocks and Watches*. London, England: Little, Brown and Company, 2000. See also <http://www.timezone.com/> and links therein.
- [5] M. V. Headrick, "Origin and evolution of the anchor clock escapement," *IEEE Control Syst. Mag.*, vol. 22, no. 2, pp. 41–52, 2002.
- [6] A. V. Roup, D. S. Bernstein, S. G. Nersesov, W. M. Haddad, and V. Chellaboina, "Limit cycle analysis of the verge and foliot clock escapement using impulsive differential equations and Poincaré maps," *Int. J. Control*, vol. 76, no. 17, pp. 1685–1698, 2003.
- [7] M. Denny, "The pendulum clock: A venerable dynamical system," *Eur. J. Phys.*, vol. 23, no. 4, pp. 449–458, 2002.
- [8] T. W. B. Kibble and F. H. Berkshire, *Classical Mechanics*. Harlow, England: Longman, 1996.
- [9] J.-C. Nicolet. (2007) *Europa Star* [Online]. Available: http://www.europastar.com/europastar/watch_tech/tourbillon.jsp
- [10] D. W. Jordan and P. Smith, *Nonlinear Ordinary Differential Equations*. Oxford, England: Oxford Univ. Press, 1999.
- [11] M. Denny, "Watt steam governor stability," *Eur. J. Phys.*, vol. 23, pp. 339–351, 2002.
- [12] E. A. Kraut, *Fundamentals of Mathematical Physics*. New York: McGraw-Hill, 1967, p. 309.

