

## APPENDIX A

# HISTORICAL PERSPECTIVES

---

We are like dwarfs on the shoulders of giants, by whose grace we see farther than they. Our study of the works of the ancients enables us to give fresh life to their finer ideas, and rescue them from time's oblivion and man's neglect.

—Peter of Blois, late twelfth century<sup>1</sup>

The Kalman filter has its roots in the early 1700s in the least squares work of Roger Cotes, who died in 1716.<sup>2</sup> However, Cotes's research was vague, without example, and therefore had little influence on later developments in estimation [Sti86]. Least squares estimation began to be more firmly developed in the middle 1700s by Tobias Mayer (estimating the motion of the moon in 1750), Leonard Euler in 1749 and Pierre Laplace in 1787 (estimating the motion of Jupiter and Saturn), Roger Boscovich in 1755 (estimating the dimensions of Earth), and Daniel Bernoulli in 1777 [Ken61]. At the age of 77, Daniel Bernoulli developed the idea of

<sup>1</sup>This quote is usually attributed to Isaac Newton, but as seen from this quote, the idea did not originate with Newton. Peter of Blois penned this analogy in the context of theological knowledge, and Newton, who himself spent much time studying theology, may have been familiar with the idea from Peter of Blois's writings.

<sup>2</sup>Cotes died at the age of 33, having published only one paper during his entire life. Some of his work was published posthumously. Isaac Newton said of him, "had he lived we might have known something."

maximum likelihood estimation. Recursive least squares was essentially established by the early 1800s with the work of Karl Gauss (published in 1809, but claimed to have been completed in 1795), Adrien Legendre (1805), and Robert Adrain (1808). Gauss and Legendre's application was estimating the locations of planets and comets on the basis of imperfect measurements, and Adrain's application was surveying. Additional information on the early history of the development of least squares estimation can be found in [Sea67, Sor80, Sor85, Sti86].

In 1880 the Danish astronomer Thorvald Nicolai Thiele extended earlier least squares work and developed a recursive algorithm very similar to the Kalman filter [Hal81, Lau81]. Thiele's filter is equivalent to the Kalman filter for the special case of a scalar state, scalar measurement, state transition and measurement matrices both equal to unity, and deterministic initial state. Thiele also proposed a way to estimate the variances of the state and measurement noise, a precursor to adaptive filtering.

It is interesting to note that most of the early contributors to estimation theory were primarily astronomers rather than mathematicians. They used mathematics as a means to an end. Then, as now, the most outstanding and lasting contributions to theory were driven by practical engineering interests. "There is nothing so practical as a good theory" [Lew51, page 169].

Wiener and Kolmogorov's work in the 1940s was similar to the Kalman filter (see Section 3.4). However, their work did not arise within the context of state-space theory. It is more statistical in nature than Kalman filtering, and requires knowledge of covariances such as  $E(x_i x_j^T)$  and  $E(y_i x_j^T)$ . In order to implement a Wiener filter in a closed form, the theory assumes that the state and measurements are stationary random processes. Furthermore, Wiener filtering is a steady-state process; that is, it assumes that the measurements have been generated from the infinite past. The 1950s saw a lot of work on relaxing the assumptions of the Wiener filter [Zad50, Boo52]. NASA spent several years investigating Wiener theory in the 1950s, but could not see any practical way to implement it in space navigation problems [Sch81].

Later in the 1950s, work began on replacing the covariance knowledge required by the Wiener filter with state-space descriptions. The results of this work were algorithms that are very close to the Kalman filter as we know it today. Work in this direction at Johns Hopkins University was motivated by missile tracking and appeared in unpublished work as early as 1956 [Spa88]. Peter Swerling's work at the RAND Corporation in the late 1950s was motivated by satellite orbit estimation [Swe59]. Swerling essentially developed (and published in 1959) the Kalman filter for the case of noise-free system dynamics. Furthermore, he considered nonlinear system dynamics and measurement equations (because of his application). Similar to the dispute between Gauss and Legendre regarding credit for the development of least squares, there has been a smaller dispute regarding credit for the development of the Kalman filter. After the Kalman filter became well known, Peter Swerling wrote a letter to the *AIAA Journal* claiming credit for the algorithm that bears Kalman's name [Swe63]. For the most part, Swerling's claim has been ignored, but his place in the development of the Kalman filter will surely be remembered. He wrote an appendix to [Bro98] comparing his work with Kalman's. Ruslan Stratonovich in the USSR also obtained the Kalman filter equations in 1960. Richard Battin independently developed the Kalman filter equations from a maximum likelihood point of view. He published his results internally at MIT in

1961 (MIT Instrumentation Laboratory Report R-341), and in the open literature in 1962 [Bat62].

Results similar to Kalman filtering were also derived prior to 1960 in fields other than engineering. For example, work as early as 1950 in statistics and economics resulted in a recursive least squares “Kalman filter” for the case of constant parameter estimation with noisy measurements [Pla50, Thi61]. More details about the relationship between this early work and the Kalman filter can be found in [Did85, Wel87].

Rudolph Kalman developed the discrete-time Kalman filter that we presented in this book in 1960 [Kal60]. Kalman and Bucy developed the continuous-time Kalman filter (discussed in Chapter 8) in 1961 [Kal61].

In view of all the earlier work along the same lines, why was the filter named after Kalman? There were probably several factors that contributed to this [Spa88]. First of all, Kalman wrote his papers in a relatively straightforward way and did not focus on any specific applications. Other similar papers were more application oriented, which tended to obscure the fundamental nature of the theory.<sup>3</sup> Second, Kalman discussed the duality between state estimation and optimal control, which allowed him to specify mathematical conditions for filter stability. Third, the increasing popularity of digital computers at the time of Kalman’s papers helped bring his work into the mainstream. Finally, Kalman made his work known to NASA, which needed just such an estimator for the Apollo space program [Sch81, McG85]. The use of the Kalman filter for the Apollo program was facilitated by Stanley Schmidt, who was the branch chief of the NASA Ames Dynamic Analysis Branch in the late 1950s and early 1960s when NASA was beginning feasibility studies of lunar missions. Kalman and Schmidt happened to be close acquaintances during the time that Kalman developed his theory and Schmidt needed a navigation algorithm. During the early 1960s, the Kalman filter was often referred to in papers as the Kalman–Schmidt filter [Bel67]. It is something of an accident of history that the filter was named after Rudolph Kalman, although it is difficult to overstate his contributions to the development of modern control and estimation theory.

Additional interesting notes on the early history of the Kalman filter can be found in [Sor70, Kai74, Lai74, Bat82, Hut84, Gre01].

<sup>3</sup>This can be a lesson for researchers today. As engineers our goal is to gear our work towards practical applications. But (as a rule of thumb) it is the more general, theoretical work that has greater influence on the world and stands the test of time.