5. General Theory
of Optimal Hankel-Norm Approximation

In the previous chapter we have studied in some detail the AAK theorem for the special case where the given Hankel matrix $\Gamma$ has finite rank. An elementary proof of the theorem under this assumption was given, and an application of the result to the problem of system reduction has also been illustrated. In this chapter, we will study the general AAK theory, see AAK [1971], where the Hankel matrix $\Gamma$ is not necessarily of finite rank and may not even be compact. We will supply a proof in detail of the general theorem. In order to give an elementary exposition that requires as little knowledge of linear operator theory as possible, our presentation is rather lengthy. As will be seen in the first section below, the solvability of the problem can be easily proved because it is easy to show that any given bounded Hankel matrix $\Gamma$ has a best approximation from the set of Hankel operators with any specified finite rank. We remark, however, that one of the main contributions of the AAK theory is that an explicit closed-form solution to the problem is formulated. The derivation of this formulation is much more difficult and will contribute to the major portion of this chapter.

5.1 Existence and Preliminary Results

In this section, we include some basic results which are needed later in this chapter. We first show that for a given bounded Hankel matrix $\Gamma$, the general problem of best approximation of $\Gamma$ by a Hankel matrix with specified finite rank is always solvable. We then characterize all bounded operators that commute with the shift operator, and apply this characterization to establish an important theorem due to Beurling concerning the $l^2$-closure of the shifts of a fixed $l^2$ sequence. Based on this result, we formulate the operator norm of any bounded Hankel matrix in terms of the outer factor of an analytic function. Finally, we will study various properties of this formulation.
5.1.1 Solvability of a Best Approximation Problem

In this section, we will show that the following best approximation problem is solvable. Let \( G^{[k]} \) denote the family of bounded Hankel matrices with rank not exceeding \( k \), namely,

\[
G^{[k]} := \{ \Gamma: \ \Gamma \text{ a bounded Hankel matrix, } \text{rank}(\Gamma) \leq k \}. \tag{5.1}
\]

For any given bounded Hankel matrix \( \Gamma \) which may have infinite rank and may not even be compact, we consider the optimization problem

\[
\inf_{A \in G^{[k]}} \| \Gamma - A \|, \tag{5.2}
\]

where \( \| \cdot \| = \| \cdot \|_s \) is the usual operator or spectral norm. Define

\[
d_k(\Gamma) = \inf_{A \in G^{[k]}} \| \Gamma - A \|. \tag{5.3}
\]

We will prove that the optimal error of approximation, \( d_k(\Gamma) \), can always be attained, and so the best approximation problem in (5.2) is solvable. Recall that if the given Hankel matrix \( \Gamma \) is normal and has finite rank, then the solvability of this problem has been verified in Sect. 4.3.2 and a closed-form solution was also given in Sects. 4.3.2 and 4.3.3. For an arbitrary bounded Hankel matrix \( \Gamma \), the following theorem gives an expected answer. However, the derivation of a closed-form solution to this general best approximation problem requires a sequence of preliminary results and will not be completed until the end of this chapter.

**Theorem 5.1.** Let \( \Gamma \) be a bounded Hankel matrix and \( G^{[k]} \) be defined as in (5.1). Then for every non-negative integer \( k \), there exists a matrix \( A_k^* \) in \( G^{[k]} \) such that

\[
d_k(\Gamma) = \| \Gamma - A_k^* \|. \tag{5.4}
\]

We will describe a procedure to define \( A_k^* \). Pick any approximating sequence \( \{ A_k^{(n)} \}_{n=1}^{\infty} \) in \( G^{[k]} \), namely,

\[
\lim_{n \to \infty} \| \Gamma - A_k^{(n)} \| = d_k(\Gamma),
\]

and set

\[
A_k^{(n)} = \begin{bmatrix}
\gamma_1^{(n)} & \gamma_2^{(n)} & \gamma_3^{(n)} & \cdots \\
\gamma_2^{(n)} & \gamma_3^{(n)} & \cdots & \cdots \\
\gamma_3^{(n)} & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots
\end{bmatrix}.
\]

Then, for the unique vector \( e_1 = [1 \ 0 \ 0 \ \cdots]^T \in \ell^2 \), we have