BULLETIN OF THE POLISH
ACADEMY OF SCIENCES
MATHEMATICS
Vol. 62, No. 1, 2014

SYSTEMS THEORY, CONTROL

Remarks on the Stabilization Problem for Linear Finite-Dimensional Systems

by

Takao NAMBU

Presented by Jerzy ZABCZYK

Summary. The celebrated 1967 pole assignment theory of W. M. Wonham for linear finite-dimensional control systems has been applied to various stabilization problems both of finite and infinite dimension. Besides existing approaches developed so far, we propose a new approach to feedback stabilization of linear systems, which leads to a clearer and more explicit construction of a feedback scheme.

1. Introduction. Since the celebrated pole assignment theory [7] for linear control systems of finite dimension appeared, the theory has been applied to various stabilization problems, both of finite and infinite dimension, such as the one with boundary output/boundary input scheme (see, e.g., [5] and the references therein).

The symbol H_n, $n = 1, 2, \dots$, hereafter will denote a finite-dimensional Hilbert space with $\dim H_n = n$, equipped with inner product $\langle \cdot, \cdot \rangle_n$ and norm $\| \cdot \|_n$. The symbol $\| \cdot \|_n$ is also used for the $\mathscr{L}(H_n)$-norm. Let A, B, and C be operators in $\mathscr{L}(H_n)$, $\mathscr{L}(\mathbb{C}^N; H_n)$, and $\mathscr{L}(H_n; \mathbb{C}^N)$, respectively. Given A, C, and any set of n complex numbers, $Z = \{\zeta_i\}_{1 \le i \le n}$, the problem is to seek a suitable B such that $\sigma(A + BC) = Z$. Or, given A and B, its algebraic counterpart is to seek a C such that $\sigma(A + BC) = Z$. Stimulated by the result of [7], various approaches and algorithms for computation of B or C have been proposed (see, e.g., [2–4]). As long as the author knows, however, each approach needs much preparation and background in linear algebra to achieve stabilization and determine the necessary parameters. Explicit realizations of B or C sometimes seem complicated. One reason is

2010 *Mathematics Subject Classification*: Primary 93D15; Secondary 93C35, 93B07.
Key words and phrases: stabilization, pole assignment in linear systems, observability, static feedback scheme.

DOI: 10.4064/ba62-1-9

no doubt the complexity of the process of determining B or C that *exactly* satisfy the relation $\sigma(A+BC)=Z$.

Let us describe our control system: The system, consisting of a state $x(\cdot)\in H_n$, output $y=Cx\in\mathbb{C}^N$, and input $u\in\mathbb{C}^N$, is described by a linear differential equation in H_n,

$$\frac{dx}{dt}=Ax+Bu,\qquad y=Cx,\qquad x(0)=x_0\in H_n. \tag{1.1}$$

Here,

$$Bu=\sum_{k=1}^{N}u_kb_k\quad\text{ for } u=(u_1\ \dots\ u_N)^{\mathrm{T}}\in\mathbb{C}^N,$$
$$Cx=(\langle x,c_1\rangle_n\ \dots\ \langle x,c_N\rangle_n)^{\mathrm{T}}\quad\text{ for } x\in H_n,$$

$(\dots)^{\mathrm{T}}$ being the transpose of vectors or matrices. The vectors $c_k\in H_n$ denote given weights of the observation (output); and $b_k\in H_n$ are actuators to be constructed. By setting $u=y$ in (1.1), the control system yields a feedback system,

$$\frac{dx}{dt}=(A+BC)x,\qquad x(0)=x_0\in H_n. \tag{1.2}$$

According to the choice of a basis for H_n, the operators A, B, and C are identified with matrices of suitable size.

Let us assume that $\sigma(A)\cap\mathbb{C}_+\neq\emptyset$, so that the system (1.1) with $u=0$ is unstable. Given a $\mu>0$, the *stabilization problem* for the finite-dimensional control system (1.2) is to seek a B or a C such that

$$\|e^{t(A+BC)}\|_n\le \text{const}\, e^{-\mu t},\qquad t\ge 0. \tag{1.3}$$

The pole assignment theory [7] plays a fundamental role in the above problem, and has been applied so far to various linear systems. The theory is stated as follows: *Let $Z=\{\zeta_i\}_{1\le i\le n}$ be any set of n complex numbers, where some ζ_i may coincide.* Then there exists an operator B such that $\sigma(A+BC)=Z$ if and only if the pair (C,A) is observable. Thus, if the set Z is chosen such that $\max_{\zeta\in Z}\operatorname{Re}\zeta$, say $-\mu_1$ $(=\operatorname{Re}\zeta_1)$, is negative, and if there is *no* generalized eigenspace of $A+BC$ corresponding to ζ_1, we obtain the decay estimate (1.3).

Now we ask: Do we need *all* information on $\sigma(A+BC)$ for stabilization? In fact, to obtain the decay estimate (1.3), it is not necessary to designate *all* elements of the set Z. What is really necessary is the number $-\mu=\max_{\zeta_i\in Z}\operatorname{Re}\zeta_i$, say $=\operatorname{Re}\zeta_1$, and the spectral property that ζ_1 does not allow any generalized eigenspace; the latter is the requirement that *no* factor of algebraic growth in time is added to the right-hand side of (1.3). In fact, when an algebraic growth is added, the decay property becomes a little worse, and the constant (≥ 1) in (1.3) increases. The above operator $A+BC$ also appears, as a *pseudo*-substructure, in stabilization problems for infinite-dimensional linear systems such as parabolic or retarded systems (see, e.g., [5]): These systems are decomposed into two, and understood as composite systems consisting of two states; one belongs to a finite-dimensional subspace, and the other to an infinite-dimensional one. It is impossible, however, to manage the infinite-dimensional substructures. Thus, no matter how *precisely* the finite-dimensional spectrum $\sigma(A+BC)$ could be assigned, it does not exactly dominate the whole structure of infinite dimension. In other words, the assigned spectrum of finite dimension is not necessarily a subset of the spectrum of the infinite-dimensional feedback control system.

In view of the above observations, our aim is to develop a new approach much simpler than in the existing literature, which allows us to construct a desired operator B or a set of actuators b_k ensuring the decay (1.3) in a simpler and more explicit manner (see (2.7) just below Lemma 2.2). The result is, however, not so sharp as in [7] in the sense that it does not generally provide the precise location of the assigned eigenvalues ([1]). From the above viewpoint of infinite-dimensional control theory, however, the result would be meaningful enough, and satisfactory for stabilization.

Our approach is based on a Sylvester equation of finite dimension. Sylvester equations in infinite-dimensional spaces have also been studied extensively (see, e.g., [1] for equations involving only bounded operators), and even unboundedness of the given operators is allowed [5]. The Sylvester equation in this paper is of finite dimension, so that there arises no difficulty caused by the complexity of infinite dimension. Given a positive integer s and vectors $\xi_k \in H_s$, $1 \leq k \leq N$, let us consider the Sylvester equation in H_n:

$$
\begin{aligned}
&XA - MX = \varXi C, \qquad \varXi \in \mathscr{L}(\mathbb{C}^N; H_s), \qquad \text{where} \\
&\varXi u = \sum_{k=1}^{N} u_k \xi_k \qquad \text{for } u = (u_1 \ldots u_N)^{\mathrm{T}} \in \mathbb{C}^N.
\end{aligned} \tag{1.4}
$$

Here, M denotes a given operator in $\mathscr{L}(H_s)$, and ξ_k vectors to be designed in H_s. A possible solution X would belong to $\mathscr{L}(H_n; H_s)$. The approach via Sylvester equations is found, e.g., in [2–4], where, by setting $n = s$, a condition for the existence of the bounded inverse $X^{-1} \in \mathscr{L}(H_n)$ is sought. Choosing an M such that $\sigma(M) \subset \mathbb{C}_-$, it is then proved that

$$A - (X^{-1}\varXi)C = X^{-1}MX, \qquad \sigma(X^{-1}MX) = \sigma(M) \subset \mathbb{C}_-,$$

([1]) In the case where we can choose $N = 1$, our result exactly coincides with the standard pole assignment theory in [7] (see our Proposition 2.3).

the left-hand side of which means a desired perturbed operator. The procedure of its derivation is, however, rather complicated, and the choice of the ξ_k is unclear. In fact, X^{-1} might not exist for some ξ_k.

Our new approach is rather different. Let us characterize the operator A in (1.4). There is a set of eigenpairs $\{-\lambda_i, \varphi_{ij}\}$ with the following properties:

(i) $\sigma(A) = \{-\lambda_i;\ 1 \le i \le n'\ (\le n)\}$, $\lambda_i \ne \lambda_j$ for $i \ne j$; and
(ii) $A\varphi_{ij} = -\lambda_i \varphi_{ij} + \sum_{k<j} \alpha^i_{jk} \varphi_{ik}$, $1 \le i \le n'$, $1 \le j \le m_i$.

Let $P_{-\lambda_i}$ be the projector in H_n corresponding to the eigenvalue $-\lambda_i$. Then we see that $P_{-\lambda_i} u = \sum_{j=1}^{m_i} u_{ij}\varphi_{ij}$ for $u \in H_n$. The restriction of A onto the invariant subspace $P_{-\lambda_i} H_n$ is, in the basis $\{\varphi_{i1}, \dots, \varphi_{im_i}\}$, represented by the $m_i \times m_i$ upper triangular matrix $-\varLambda_i$, where

$$\varLambda_i|_{(j,k)} = \begin{cases} -\alpha^i_{kj}, & j < k, \\ \lambda_i, & j = k, \\ 0, & j > k. \end{cases}$$

If we set $\varLambda_i = \lambda_i + N_i$, the matrix N_i is nilpotent, that is, $N_i^{m_i} = 0$. The minimum integer n such that $\ker N_i^n = \ker N_i^{n+1}$, denoted as l_i, is called the *ascent* of $-\lambda_i - A$. It is well known that the ascent l_i coincides with the order of the pole $-\lambda_i$ of the resolvent $(\lambda - A)^{-1}$. The Laurent expansion of $(\lambda - A)^{-1}$ in a neighborhood of the pole $-\lambda_i \in \sigma(A)$ is expressed as

$$(1.5)\quad \begin{aligned} &(\lambda - A)^{-1} = \sum_{j=1}^{l_i} \frac{K_{-j}}{(\lambda+\lambda_i)^j} + \sum_{j=0}^{\infty} (\lambda+\lambda_i)^j K_j, \quad \text{where} \\ &l_i \le m_i, \quad K_j = \frac{1}{2\pi \mathrm{i}} \int_{|\zeta+\lambda_i|=\delta} \frac{(\zeta - A)^{-1}}{(\zeta+\lambda_i)^{j+1}}\, d\zeta, \quad j = 0, \pm 1, \pm 2, \dots. \end{aligned}$$

Note that $K_{-1} = P_{-\lambda_i}$. The set $\{\varphi_{ij};\ 1 \le i \le n',\ 1 \le j \le m_i\}$ forms a basis for H_n. Each $x \in H_n$ is uniquely expressed as $x = \sum_{i,j} x_{ij}\varphi_{ij}$. Let T be a bijection, defined as $Tx = (x_{11}\ x_{12}\ \dots\ x_{n'm_{n'}})^{\mathrm{T}}$. Then A is identified with the upper triangular matrix $-\varLambda$;

$$(1.6)\qquad TAT^{-1} = -\varLambda = -\operatorname{diag}(\varLambda_1\ \dots\ \varLambda_{n'}).$$

We turn to the operator M in (1.4). Let η_{ij}, $1 \le i \le n$, $1 \le j \le \ell_i$, be an orthonormal basis for H_s. Then necessarily $s = \sum_{i=1}^n \ell_i \ge n$. Every vector $v \in H_s$ is expressed as $v = \sum_{i=1}^n \sum_{j=1}^{\ell_i} v_{ij}\eta_{ij}$, where $v_{ij} = \langle v, \eta_{ij}\rangle_s$. Let $\{\mu_i\}_{1=1}^n$ be a set of positive numbers such that $0 < \mu_1 < \dots < \mu_n$, and set

$$(1.7)\quad Mv = -\sum_{i=1}^n \sum_{j=1}^{\ell_i} \mu_i v_{ij}\eta_{ij} \quad \text{for} \quad v = \sum_{i=1}^n \sum_{j=1}^{\ell_i} v_{ij}\eta_{ij}, \qquad v_{ij} = \langle v, \eta_{ij}\rangle_s.$$

It is apparent that (i) $\sigma(M) = \{-\mu_i\}_{i=1}^n$; and (ii) $(\mu_i + M)\eta_{ij} = 0$, $1 \le i \le n$, $1 \le j \le \ell_i$. The operator M is self-adjoint, and negative-definite,

$$\langle Mv, v\rangle_s = -\sum_{i=1}^{n}\sum_{j=1}^{n_i} \mu_i |v_{ij}|^2 \le -\mu_1 \|v\|_s^2.$$

Let $Q_{-\mu_i}$ be the projector in H_s corresponding to the eigenvalue $-\mu_i \in \sigma(M)$, say $Q_{-\mu_i} v = \sum_{j=1}^{\ell_i} v_{ij}\eta_{ij}$ for $v = \sum_{i,j} v_{ij}\eta_{ij}$. We put an additional condition on M in (1.7):

$$\sigma(A) \cap \sigma(M) = \emptyset. \tag{1.8}$$

Assuming (1.8), we derive our first result. Since the proof is carried out in exactly the same manner as in [5], it is omitted.

PROPOSITION 1.1. *Suppose that the condition* (1.8) *is satisfied. Then the Sylvester equation* (1.4) *admits a unique operator solution* $X \in \mathscr{L}(H_n; H_s)$. *The solution* X *is expressed as*

$$\begin{aligned} Xu &= \frac{-1}{2\pi \mathrm{i}} \int_{\Gamma} (\lambda - M)^{-1} \Xi C (\lambda - A)^{-1} u \, d\lambda = -\sum_{\lambda \in \sigma(M)} Q_\lambda \Xi C (\lambda - A)^{-1} u \\ &= \sum_{i=1}^{n} Q_{-\mu_i} \Xi C (\mu_i + A)^{-1} u, \end{aligned}$$

where Γ *denotes a Jordan contour encircling* $\sigma(M)$ *in its inside, with* $\sigma(A)$ *outside* Γ. *The above first expression is the so called Rosenblum formula* [1].

Our main results are stated as Theorem 2.1 and Lemma 2.2 in the next section, where a more explicit and concrete expression than ever before of a set of stabilizing actuators b_k in (1.2) is obtained. As we see in the next section, an advantage of considering the operator $X \in \mathscr{L}(H_n; H_s)$ with $s \ge n$ is that the bounded inverse $(X^*X)^{-1}$ is ensured under a reasonable assumption on the operator Ξ. A numerical example is also given. Finally, Proposition 2.3 is stated, where our feedback scheme exactly coincides with the standard pole assignment theory [7] in the case where we can choose $N = 1$.

2. Main results. We assume that $\sigma(A) \cap \mathbb{C}_+ \ne \emptyset$, so that the semigroup e^{tA}, $t \ge 0$, is *unstable*. We construct suitable actuators $b_k \in H_n$ in (1.2) such that $e^{t(A+BC)}$ has a preassigned decay rate, say $-\mu_1$ (see (1.7)). The operator $(C\ CA\ \dots\ CA^{n-1})^{\mathrm{T}}$ belongs to $\mathscr{L}(H_n; \mathbb{C}^{nN})$. Recall that the observability condition on the pair (C, A) is that it is injective, in other words, $\ker(C\ CA\ \dots\ CA^{n-1})^{\mathrm{T}} = \{0\}$. Throughout the section, the condition (1.8) is assumed in the Sylvester equation (1.4). Then we obtain one of the main results:

Theorem 2.1. *Assume that*

$$\begin{aligned} &\ker (C\ CA\ \dots\ CA^{n-1})^{\mathrm{T}} = \{0\}, \\ &\ker Q_{-\mu_i} \Xi = \{0\}, \quad 1 \le i \le n. \end{aligned} \tag{2.1}$$

Then $\ker X = \{0\}$.

Proof. Let $Xu = 0$. In view of Proposition 1.1, we see that

$$Q_{-\mu_i} \Xi C(\mu_i + A)^{-1} u = 0, \quad 1 \le i \le n.$$

Since $\ker Q_{-\mu_i}\Xi = \{0\}$, $1 \le i \le n$, by (2.1), we obtain

$$\begin{aligned} C(\mu_i + A)^{-1} u = 0, \quad 1 \le i \le n, \quad \text{or} \\ \langle (\mu_i + A)^{-1} u, c_k \rangle_n = 0, \quad 1 \le k \le N,\ 1 \le i \le n. \end{aligned} \tag{2.2}$$

Set $f_k(\lambda; u) = \langle (\lambda + A)^{-1} u, c_k \rangle_n$. By recalling that $T(\lambda - A)^{-1} T^{-1} = (\lambda + \Lambda)^{-1}$ (see (1.6)), $f_k(\lambda; u)$ is rewritten as $\langle (\lambda + \Lambda)^{-1} Tu, (T^{-1})^* c_k \rangle_{\mathbb{C}^n}$. Each element of the $n \times n$ matrix $(\lambda + \Lambda)^{-1}$ is a rational function of λ; its denominator is a polynomial of order n, and the numerator at most of order $n-1$. This means that each $f_k(\lambda; u)$ is a rational function of λ, the denominator of which is a polynomial of order n, and the numerator of order $n-1$. Since the numerator of f_k has *at least* n distinct zeros μ_i, $1 \le i \le n$, by (2.2), we conclude that

$$f_k(\lambda; u) = \langle (\lambda + A)^{-1} u, c_k \rangle_n = 0, \quad -\lambda \in \rho(A),\ 1 \le k \le N.$$

Let $c \in \rho(A)$, and set $A_c = c - A$. In view of the identity

$$(\lambda + A)^{-1} = A_c (\lambda + A)^{-1} A_c^{-1} = -A_c^{-1} + (\lambda + c)(\lambda + A)^{-1} A_c^{-1},$$

let us introduce a series of rational functions $f_k^l(\lambda; u)$, $l = 0, 1, \dots$, as

$$f_k^0(\lambda; u) = f_k(\lambda; u), \qquad f_k^{l+1}(\lambda; u) = \frac{f_k^l(\lambda; u)}{\lambda + c}, \quad l = 0, 1, \dots.$$

It is easily seen that

$$f_k^l(\lambda; u) = \langle (\lambda + A)^{-1} A_c^{-l} u, c_k \rangle_n - \sum_{i=1}^{l} \frac{1}{(\lambda + c)^i} \langle A_c^{-(l+1-i)} u, c_k \rangle_n, \tag{2.3}$$

and

$$f_k^l(\lambda; u) = 0, \quad \lambda \in -\rho(A) \setminus \{-c\},\ 1 \le k \le N,\ l \ge 0.$$

In view of the Laurent expansion (1.5) of $(\lambda - A)^{-1}$ in a neighborhood of $-\lambda_i$, we obtain

$$\begin{aligned} 0 &= f_k(\lambda; u) \\ &= -\sum_{j=1}^{l_i} \frac{\langle K_{-j} u, c_k \rangle_n}{(-\lambda + \lambda_i)^j} - \sum_{j=0}^{\infty} (-\lambda + \lambda_i)^j \langle K_j u, c_k \rangle_n, \quad 1 \le k \le N, \end{aligned}$$

in a neighborhood of λ_i. Calculation of the residue of $f_k(\lambda; u)$ at λ_i implies that

$$(2.4)\qquad \begin{aligned} \langle K_{-1}u, c_k\rangle_n = \langle P_{-\lambda_i}u, c_k\rangle_n = 0, \quad 1 \le i \le n',\ 1 \le k \le N, \quad \text{or} \\ CP_{-\lambda_i}u = 0, \quad 1 \le i \le n'. \end{aligned}$$

As for $f_k^l(\lambda; u)$, $l \ge 1$, we have a similar expression in a neighborhood of λ_i,

$$\begin{aligned} f_k^l(\lambda; u) = -\sum_{j=1}^{l_i} \frac{\langle K_{-j}A_c^{-l}u, c_k\rangle_n}{(-\lambda+\lambda_i)^j} - \sum_{j=0}^{\infty}(-\lambda+\lambda_i)^j \langle K_j A_c{}^{-l}u, c_k\rangle_n \\ - \sum_{i=1}^{l} \frac{1}{(\lambda+c)^i} \langle A_c^{-(l+1-i)}u, c_k\rangle_n = 0 \end{aligned}$$

by (2.3). Note that $K_{-1}A_c^{-l}u = P_{-\lambda_i}A_c^{-l}u = A_c^{-l}P_{-\lambda_i}u$. Calculation of the residue of $f_k^l(\lambda; u)$ at λ_i similarly implies that

$$\begin{aligned} \langle K_{-1}A_c^{-l}u, c_k\rangle_n = \langle A_c^{-l}P_{-\lambda_i}u, c_k\rangle_n = 0, \quad 1 \le i \le n',\ 1 \le k \le N, \quad \text{or} \\ CA_c^{-l}P_{-\lambda_i}u = 0, \quad 1 \le i \le n',\ l \ge 1. \end{aligned}$$

Combining these with the above relation (2.4), we see that

$$(2.5)\qquad (C\ CA_c^{-1}\ \dots\ CA_c^{-(n-1)})^{\mathrm{T}} P_{-\lambda_i}u = 0, \quad 1 \le i \le n'.$$

It is clear that $\ker(C\ CA\ \dots\ CA^{n-1})^{\mathrm{T}} = \ker(C\ CA_c\ \dots\ CA_c^{n-1})^{\mathrm{T}}$, where $A_c = c - A$. Thus, by the first condition of (2.1), it is easily seen that

$$\ker(C\ CA_c^{-1}\ \dots\ CA_c^{-(n-1)})^{\mathrm{T}} = \ker(C\ CA\ \dots\ CA^{n-1})^{\mathrm{T}} = \{0\}.$$

Thus, (2.5) immediately implies that $P_{-\lambda_i}u = 0$ for $1 \le i \le n'$, and finally that $u = 0$. ∎

By Theorem 2.1, there is a positive constant such that

$$\|Xu\|_s \ge \text{const}\, \|u\|_n, \qquad \forall u \in H_n.$$

The derivation of the above positive lower bound of $\|Xu\|_s$ is due to a specific nature of finite-dimensional spaces. The operator $X^*X \in \mathscr{L}(H_n)$ is self-adjoint, and positive-definite. In fact, by the relation

$$\text{const}\, \|u\|_n^2 \le \|Xu\|_s^2 = \langle Xu, Xu\rangle_s = \langle X^*Xu, u\rangle_n \le \|X^*Xu\|_n \|u\|_n,$$

we see that $\|X^*Xu\|_n \ge \text{const}\, \|u\|_n$. Thus the bounded inverse $(X^*X)^{-1} \in \mathscr{L}(H_n)$ exists. We go back to the Sylvester equation (1.4). Setting $X^*X = \mathscr{X} \in \mathscr{L}(H_n)$ and $X^*MX = \mathscr{M} \in \mathscr{L}(H_n)$, we obtain the relation

$$\begin{aligned} A - (X^*X)^{-1}X^*MX = (X^*X)^{-1}X^*\varXi C, \quad \text{or} \\ A - \sum_{k=1}^{N} \langle\cdot, c_k\rangle_n \mathscr{X}^{-1}X^*\xi_k = \mathscr{X}^{-1}\mathscr{M}. \end{aligned}$$

Both operators $\mathscr{X}$ and $\mathscr{M}$ are self-adjoint, but $\mathscr{X}^{-1}\mathscr{M}$ is not. The following assertion is the second of our main results, and leads to a stabilization result:

Lemma 2.2. *Assume that* (2.1) *is satisfied. Then* $\sigma(\mathscr{X}^{-1}\mathscr{M})$ *is contained in* $\mathbb{R}^1_-$. *Actually,*

$$-\lambda_* = \max \sigma(\mathscr{X}^{-1}\mathscr{M}) \le -\mu_1. \tag{2.6}$$

In addition, there is no generalized eigenspace for any $\lambda \in \sigma(\mathscr{X}^{-1}\mathscr{M})$.

Remark. By Lemma 2.2, we obtain a decay estimate

$$\begin{aligned} \|\exp t(A - (X^*X)^{-1}X^*\varXi C)\|_n &= \|\exp t(\mathscr{X}^{-1}\mathscr{M})\|_n \\ &\le \mathrm{const}\, e^{-\mu_1 t}, \qquad t \ge 0. \end{aligned} \tag{2.7}$$

In fact, the last assertion of the lemma ensures that no algebraic growth in time arises in the semigroup, regarding the greatest eigenvalue. Thus, the set of actuators $b_k = -(X^*X)^{-1}X^*\xi_k$, $1 \le k \le N$, in other words, $B = -(X^*X)^{-1}X^*\varXi$, explicitly gives the desired set of actuators in (1.2).

Proof of Lemma 2.2. Since $\mathscr{X}$ is positive-definite, we can find a non-unique bijection $\mathscr{U} \in \mathscr{L}(H_n)$ such that

$$\mathscr{X} = X^*X = \mathscr{U}^*\mathscr{U},$$

the so called Cholesky factorization. Define $\mathscr{M}' = (\mathscr{U}^*)^{-1}\mathscr{M}\mathscr{U}^{-1} = (\mathscr{U}^{-1})^*\mathscr{M}\mathscr{U}^{-1}$. Then $\mathscr{M}' \in \mathscr{L}(H_n)$ is a self-adjoint operator, enjoying some properties similar to those of $\mathscr{X}^{-1}\mathscr{M}$. In fact, let $\lambda \in \sigma(\mathscr{X}^{-1}\mathscr{M})$, or $(\lambda\mathscr{X} - \mathscr{M})u = 0$ for some $u \ne 0$. Then, since

$$\begin{aligned} 0 = (\lambda\mathscr{U}^*\mathscr{U} - \mathscr{M})u &= \mathscr{U}^*(\lambda - (\mathscr{U}^*)^{-1}\mathscr{M}\mathscr{U}^{-1})\mathscr{U}u \\ &= \mathscr{U}^*(\lambda - \mathscr{M}')\mathscr{U}u = 0, \end{aligned}$$

we see that λ belongs to $\sigma(\mathscr{M}')$. The converse relation is also correct, which means that

$$\sigma(\mathscr{X}^{-1}\mathscr{M}) = \sigma(\mathscr{M}') \subset \mathbb{R}^1.$$

Inequality (2.6) is achieved by applying the well known min-max principle to $\mathscr{M}'$, or more directly by the following observation: Let $\lambda \in \sigma(\mathscr{X}^{-1}\mathscr{M})$, and $(\lambda\mathscr{X} - \mathscr{M})u = 0$ for some $u \ne 0$. Then

$$\lambda\|Xu\|_s^2 = \lambda\langle\mathscr{X}u, u\rangle_n = \langle\mathscr{M}u, u\rangle_n = \langle MXu, Xu\rangle_s \le -\mu_1\|Xu\|_s^2,$$

from which (2.6) immediately follows, since $Xu \ne 0$.

Next let us show that there is *no* generalized eigenspace for $\lambda \in \sigma(\mathscr{X}^{-1}\mathscr{M})$. Let $(\lambda - \mathscr{X}^{-1}\mathscr{M})^2 u = 0$ for some $u \ne 0$. Setting $v =$

$(\lambda - \mathscr{X}^{-1}\mathscr{M})u$, we calculate

$$\begin{aligned} 0 &= \mathscr{X}(\lambda - \mathscr{X}^{-1}\mathscr{M})^2 u = (\lambda\mathscr{X} - \mathscr{M})v \\ &= (\lambda\mathscr{U}^*\mathscr{U} - \mathscr{M})v = \mathscr{U}^*(\lambda - (\mathscr{U}^*)^{-1}\mathscr{M}\mathscr{U}^{-1})\mathscr{U}v \\ &= \mathscr{U}^*(\lambda - \mathscr{M}')w = 0, \qquad w = \mathscr{U}v, \end{aligned}$$

or $(\lambda - \mathscr{M}')w = 0$. On the other hand, since

$$\begin{aligned} w = \mathscr{U}v &= \mathscr{U}(\lambda - \mathscr{X}^{-1}\mathscr{M})u = \mathscr{U}(\lambda - \mathscr{U}^{-1}(\mathscr{U}^*)^{-1}\mathscr{M})u \\ &= (\lambda - (\mathscr{U}^*)^{-1}\mathscr{M}\mathscr{U}^{-1})\mathscr{U}u = (\lambda - \mathscr{M}')\mathscr{U}u, \end{aligned}$$

we see that

$$0 = (\lambda - \mathscr{M}')w = (\lambda - \mathscr{M}')^2\mathscr{U}u, \qquad \mathscr{U}u \neq 0.$$

But $\mathscr{M}'$ is self-adjoint, so that there is no generalized eigenspace for $\lambda \in \sigma(\mathscr{M}')$. Thus, $\mathscr{U}u$ turns out to be an eigenvector of $\mathscr{M}'$ for λ, and

$$\begin{aligned} 0 &= \mathscr{U}^*(\lambda - \mathscr{M}')\mathscr{U}u = \mathscr{U}^*(\lambda - (\mathscr{U}^*)^{-1}\mathscr{M}\mathscr{U}^{-1})\mathscr{U}u \\ &= (\lambda\mathscr{U}^*\mathscr{U} - \mathscr{M})u = (\lambda\mathscr{X} - \mathscr{M})u. \end{aligned}$$

This means that u is an eigenvector of $\mathscr{X}^{-1}\mathscr{M}$ for λ. ∎

The following example shows that $\lambda_* = -\max\sigma(\mathscr{X}^{-1}\mathscr{M})$ does not generally coincide with the prescribed μ_1.

Example. Let $n = 3$, and set $H_3 = \mathbb{C}^3$, so that A is a 3×3 matrix. Let $A = -\operatorname{diag}(a\ a\ b)$, where $a, b \leq 0$ and $a \neq b$. Since $n = 3$, $n' = 2$, $m_1 = 2$, and $m_2 = 1$, we choose $N = 2$, $s = 6$, $H_6 = \mathbb{C}^6$, and $\ell_1 = \ell_2 = \ell_3 = 2$. As for the operator $C \in \mathscr{L}(\mathbb{C}^3; \mathbb{C}^2)$, let us consider the case, for example, where $c_1 = (1\ 0\ 1)^{\mathrm{T}}$ and $c_2 = (0\ 1\ 0)^{\mathrm{T}}$. The operator C is a 2×3 matrix given by $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. The pair (C, A) is then observable, and the first condition of (2.1) is satisfied.

To consider the Sylvester equation (1.4), let $\{\eta_{ij};\ 1 \leq i \leq 3,\ j = 1, 2\}$ be a standard basis for $\mathbb{C}^6$ such that $\eta_{11} = (1\ 0\ 0\ \dots\ 0)^{\mathrm{T}}$, $\eta_{12} = (0\ 1\ 0\ \dots\ 0)^{\mathrm{T}}$, $\eta_{21} = (0\ 0\ 1\ \dots\ 0)^{\mathrm{T}}$, $\dots$, and $\eta_{32} = (0\ \dots\ 0\ 1)^{\mathrm{T}}$. Set $M = -\operatorname{diag}(\mu_1\ \mu_1\ \mu_2\ \mu_2\ \mu_3\ \mu_3)$ for $0 < \mu_1 < \mu_2 < \mu_3$. In the operator Ξ given by $\Xi u = u_1\xi_1 + u_2\xi_2$ for $(u_1\ u_2)^{\mathrm{T}} \in \mathbb{C}^2$, set $\xi_1 = (1\ 0\ 1\ 0\ 1\ 0)^{\mathrm{T}}$ and $\xi_2 = (0\ 1\ 0\ 1\ 0\ 1)^{\mathrm{T}}$. Then we see that $\ker Q_{-\mu_i}\Xi = \{0\}$, $1 \leq i \leq 3$, and the second condition of (2.1) is satisfied. The unique solution $X \in \mathscr{L}(\mathbb{C}^3; \mathbb{C}^6)$ to the Sylvester equation (1.4) is a 6×3 matrix described as ($u = (u_{11}\ u_{12}\ u_{21})^{\mathrm{T}} \in \mathbb{C}^3$)

$$Xu = \begin{pmatrix} \langle(\mu_1 + A)^{-1}u, c_1\rangle \\ \langle(\mu_1 + A)^{-1}u, c_2\rangle \\ \hline \langle(\mu_2 + A)^{-1}u, c_1\rangle \\ \langle(\mu_2 + A)^{-1}u, c_2\rangle \\ \hline \langle(\mu_3 + A)^{-1}u, c_1\rangle \\ \langle(\mu_3 + A)^{-1}u, c_2\rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{\mu_1 - a} & 0 & \frac{1}{\mu_1 - b} \\ 0 & \frac{1}{\mu_1 - a} & 0 \\ \frac{1}{\mu_2 - a} & 0 & \frac{1}{\mu_2 - b} \\ 0 & \frac{1}{\mu_2 - a} & 0 \\ \frac{1}{\mu_3 - a} & 0 & \frac{1}{\mu_3 - b} \\ 0 & \frac{1}{\mu_3 - a} & 0 \end{pmatrix} \begin{pmatrix} u_{11} \\ u_{12} \\ u_{21} \end{pmatrix},$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{C}^3$. Setting, for computational convenience,

$$\alpha = \left(\frac{1}{\mu_1 - a} \ \frac{1}{\mu_2 - a} \ \frac{1}{\mu_3 - a} \right)^{\mathrm{T}}, \quad \beta = \left(\frac{1}{\mu_1 - b} \ \frac{1}{\mu_2 - b} \ \frac{1}{\mu_3 - b} \right)^{\mathrm{T}},$$
$$1 = (1\ 1\ 1)^{\mathrm{T}},$$

we see that

$$(X^* X)^{-1} = \frac{1}{\gamma} \begin{pmatrix} |\beta|^2 & 0 & -\langle\alpha, \beta\rangle \\ 0 & |\beta|^2 - \langle\alpha, \beta\rangle^2/|\alpha|^2 & 0 \\ -\langle\alpha, \beta\rangle & 0 & |\alpha|^2 \end{pmatrix},$$

where $\gamma = |\alpha|^2|\beta|^2 - \langle\alpha, \beta\rangle^2$. By noting that $X^*\xi_1 = (\langle\alpha, 1\rangle\ 0\ \langle\beta, 1\rangle)^{\mathrm{T}}$ and $X^*\xi_2 = (0\ \langle\alpha, 1\rangle\ 0)^{\mathrm{T}}$, the matrix $A-(X^*X)^{-1}X^*\varXi C$ is concretely described as

$$- \operatorname{diag}(a\ a\ b)$$
$$-\frac{1}{\gamma} \left(\begin{array}{c|c|c} |\beta|^2\langle\alpha, 1\rangle - \langle\alpha, \beta\rangle\langle\beta, 1\rangle & 0 & |\beta|^2\langle\alpha, 1\rangle - \langle\alpha, \beta\rangle\langle\beta, 1\rangle \\ \hline 0 & \langle\alpha, 1\rangle(|\beta|^2 - \langle\alpha, \beta\rangle^2/|\alpha|^2) & 0 \\ \hline |\alpha|^2\langle\beta, 1\rangle - \langle\alpha, \beta\rangle\langle\alpha, 1\rangle & 0 & |\alpha|^2\langle\beta, 1\rangle - \langle\alpha, \beta\rangle\langle\alpha, 1\rangle \end{array} \right).$$

It is apparent that one of the eigenvalues of this matrix is the $(2, 2)$-element:

$$-a - \frac{\langle\alpha, 1\rangle}{\gamma} \left(|\beta|^2 - \frac{\langle\alpha, \beta\rangle^2}{|\alpha|^2} \right) = -a - \frac{\langle\alpha, 1\rangle}{|\alpha|^2},$$

and is certainly smaller than $-\mu_1$. Note that

$$0 < \lambda_* - \mu_1 \leq \frac{1}{|\alpha|^2} \left(\frac{\mu_2 - \mu_1}{(\mu_2 - a)^2} + \frac{\mu_3 - \mu_1}{(\mu_3 - a)^2} \right) \to 0, \qquad \mu_2, \mu_3 \to \infty.$$

The other eigenvalues are those of the matrix

$$(2.8)\qquad -\frac{1}{\gamma}\begin{pmatrix} |\beta|^2\langle\alpha,1\rangle-\langle\alpha,\beta\rangle\langle\beta,1\rangle+\gamma a & |\beta|^2\langle\alpha,1\rangle-\langle\alpha,\beta\rangle\langle\beta,1\rangle \\ |\alpha|^2\langle\beta,1\rangle-\langle\alpha,\beta\rangle\langle\alpha,1\rangle & |\alpha|^2\langle\beta,1\rangle-\langle\alpha,\beta\rangle\langle\alpha,1\rangle+\gamma b \end{pmatrix}.$$

To see that these eigenvalues are generally smaller than $-\mu_1$, let us consider a numerical example: Let $(\mu_1\ \mu_2\ \mu_3) = (2\ 3\ 4)$, $a = 0$, and $b = -1$. Then

$$\alpha = \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \end{pmatrix}^{\mathrm{T}}, \quad \beta = \begin{pmatrix} \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}^{\mathrm{T}}, \quad |\alpha|^2 = \frac{61}{144}, \quad |\beta|^2 = \frac{769}{3600},$$
$$\langle\alpha,\beta\rangle = \frac{3}{10}, \quad \langle\alpha,1\rangle = \frac{13}{12}, \quad \langle\beta,1\rangle = \frac{47}{60},$$
$$\gamma = |\alpha|^2|\beta|^2 - \langle\alpha,\beta\rangle^2 = \frac{253}{518400}.$$

One of the eigenvalues $-a - \langle\alpha,1\rangle/|\alpha|^2$ is $-156/61 < -2\ (= -\mu_1)$. The matrix (2.8) is then

$$\frac{-1}{253}\begin{pmatrix} -1860 & -1860 \\ 3540 & 3287 \end{pmatrix},$$

the eigenvalues of which are denoted as ζ_1 and ζ_2. Then $\zeta_2 < -156/61 < \zeta_1 < -2 = -\mu_1$, and thus $-\lambda_* = \zeta_1 < -\mu_1 = -2$.

We close this paper with the following remark: There is a case where λ_* coincides with μ_1. Following [6], let us consider (1.2) in the space $H_n = \mathbb{C}^n$ (see (1.6)). All operators A, B, and C are then matrices of respective sizes. Let $\sigma(A)$ consist only of simple eigenvalues, so that $m_i = 1$, $1 \le i \le n$, and $n = n'$. Thus we can choose $N = 1$, $\ell_i = 1$, $1 \le i \le n$, and thus $s = n$. The operator in (2.7) is written as $A - (X^*X)^{-1}X^*\varXi C$, where $\varXi u = u\xi$ for $u \in \mathbb{C}^1$, and $C = \langle\cdot, c\rangle_n$, $c = (c_1\ \dots\ c_n)^{\mathrm{T}} \in \mathbb{C}^n$. The observability condition then turns out to be $c_i \ne 0$, $1 \le i \le n$. Let us consider the Sylvester equation (1.4) in $H_s = \mathbb{C}^n$. By setting $\xi = (1\ 1\ \dots\ 1)^{\mathrm{T}} \in \mathbb{C}^n$, the solution X to (1.4) is an $n \times n$ matrix, and has a bounded inverse:

$$X = \varPhi\tilde{C}, \qquad \text{where}$$
$$\varPhi = \left(\frac{1}{\mu_i - \lambda_j};\ \begin{matrix} i \downarrow & 1,\dots,n, \\ j \rightarrow & 1,\dots,n \end{matrix}\right) \quad \text{and} \quad \tilde{C} = \mathrm{diag}\,(c_1\ \dots\ c_n).$$

Thus, $A - (X^*X)^{-1}X^*\varXi C = A - X^{-1}\xi c^{\mathrm{T}}$. We have shown in [6] that, given a set $\{\mu_i\}_{1\le i\le n}$, there is a unique $h \in \mathbb{C}^n$ such that $\sigma(A - hc^{\mathrm{T}}) = \{-\mu_i\}_{1\le i\le n}$, and that h is expressed as

$$h = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \\ \vdots \\ h_n \end{pmatrix} = \frac{-1}{\Delta} \begin{pmatrix} \frac{1}{c_1}\Delta_1 f(\lambda_1) \\ -\frac{1}{c_2}\Delta_2 f(\lambda_2) \\ \frac{1}{c_3}\Delta_3 f(\lambda_3) \\ \vdots \\ (-1)^{n-1}\frac{1}{c_n}\Delta_n f(\lambda_n) \end{pmatrix}, \quad \text{where } f(\lambda) = \prod_{i=1}^{n}(\lambda - \mu_i),$$

$$\Delta = \prod_{1 \le i < j \le n} (\lambda_i - \lambda_j), \qquad \Delta_k = \prod_{\substack{1 \le i < j \le n \\ i,j \ne k}} (\lambda_i - \lambda_j), \quad 1 \le k \le n.$$

Proposition 2.3. *Suppose in Lemma 2.2 that $\sigma(A)$ consists only of simple eigenvalues. Set $\xi = (1\ 1\ \ldots\ 1)^{\mathrm{T}}$ as above. Then $X^{-1}\xi = h$, and thus $\lambda_* = \mu_1$. In fact, $\sigma(A - (X^*X)^{-1}X^*\Xi C) = \{-\mu_i\}_{1 \le i \le n}$.*

Proof. The relation $X^{-1}\xi = h$ is rewritten as

$$-\Delta \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \Phi \hat{C} \begin{pmatrix} \frac{1}{c_1}\Delta_1 f(\lambda_1) \\ -\frac{1}{c_2}\Delta_2 f(\lambda_2) \\ \frac{1}{c_3}\Delta_3 f(\lambda_3) \\ \vdots \\ (-1)^{n-1}\frac{1}{c_n}\Delta_n f(\lambda_n) \end{pmatrix} = \Phi \begin{pmatrix} \Delta_1 f(\lambda_1) \\ -\Delta_2 f(\lambda_2) \\ \Delta_3 f(\lambda_3) \\ \vdots \\ (-1)^{n-1}\Delta_n f(\lambda_n) \end{pmatrix}.$$

In other words, we show that

$$\text{(2.9)} \quad -\sum_{j=1}^{n} \frac{(-1)^{j-1}\Delta_j f(\lambda_j)}{\mu_i - \lambda_j} = \sum_{j=1}^{n} (-1)^{j-1}\Delta_j \overbrace{\prod_{\substack{1 \le \ell \le n \\ \ell \ne i}} (\lambda_j - \mu_\ell)}^{(=\lambda_j^{n-1} + \cdots)} = \Delta, \quad 1 \le i \le n.$$

The left-hand side of (2.9), a polynomial of λ_i, $1 \le i \le n$, is in particular a polynomial of λ_1 of order $n-1$, and the coefficient of λ_1^{n-1} is $\Delta_1 = \prod_{2 \le i < j \le n}(\lambda_i - \lambda_j)$. For $j < k$, let us compare the jth and the kth terms. The following lemma is elementary:

Lemma 2.4 ([6]). *Let $1 \le j < k \le n$. In the product Δ_k, a polynomial of $\{\lambda_i\}_{i \ne k}$, set $\lambda_j = \lambda_k$. Then,*

$$\Delta_k = (-1)^{k-1+j}\Delta_j.$$

In the left-hand side of (2.9), set $\lambda_j = \lambda_k$. Since the terms other than the jth and the kth contain the factor $\lambda_j - \lambda_k$, they become 0. The kth term is then

$$(-1)^{k-1}\Delta_k \prod_{\substack{1\le\ell\le n\\ \ell\ne i}} (\lambda_k-\mu_\ell) = (-1)^{k-1}(-1)^{k-1-j}\Delta_j \prod_{\substack{1\le\ell\le n\\ \ell\ne i}} (\lambda_k-\mu_\ell)$$
$$= -(-1)^{j-1}\Delta_j \prod_{\substack{1\le\ell\le n\\ \ell\ne i}} (\lambda_j-\mu_\ell) = -(\text{the } j\text{th term}).$$

Thus the left-hand side of (2.9) has factors $\lambda_j - \lambda_k$, $j < k$, and is written as $c\Delta$. But $c\Delta$ is a polynomial of λ_1 of order $n-1$, and the coefficient of λ_1^{n-1} is $c\Delta_1$. This means that $c = 1$, and the proof of relation (2.9) is now complete. ■

Acknowledgments. The author thanks the anonymous referee for helpful comments on the original manuscript. The revised version has become more readable.

References

[1] R. Bhatia and P. Rosenthal, *How and why to solve the operator equation $AX-XB=Y$*, Bull. London Math. Soc. 29 (1997), 1–21.

[2] S. P. Bhattacharyya and E. de Souza, *Pole assignment via Sylvester's equation*, Systems Control Lett. 1 (1982), 261–263.

[3] E. K. Chu, *A pole-assignment algorithm for linear state feedback*, Systems Control Lett. 7 (1986), 289–299.

[4] K. Datta, *The matrix equation $XA-BX=R$ and its applications*, Linear Algebra Appl. 109 (1988), 91–105.

[5] T. Nambu, *Alternative algebraic approach to stabilization for linear parabolic boundary control systems*, Math. Control Signals Systems 26 (2014), 119–144.

[6] T. Nambu, *Algebraic multiplicities arising from static feedback control systems of parabolic type*, Numer. Funct. Anal. Optim. (2014) (online).

[7] W. M. Wonham, *On pole assignment in multi-input controllable linear systems*, IEEE Trans. Automat. Control 12 (1967), 660–665.

Takao Nambu
Department of Applied Mathematics
Graduate School of System Informatics
Kobe University
Nada, Kobe 657-8501, Japan
E-mail: nambu@kobe-u.ac.jp

Received November 18, 2013;
received in final form June 30, 2014 (7947)