International Journal of Pure and Applied Mathematics
Volume 38 No. 3 2007, 365-384

THE DUALITIES OF LYAPUNOV AND RICCATI MATRICES IN LINEAR QUADRATIC CONTROL THEORY

Chein-Shan Liu
Department of Mechanical and Mechatronic Engineering
National Taiwan Ocean University
Keelung, 202-24, TAIWAN, R.O.C.
e-mail: csliu@mail.ntou.edu.tw

Abstract: The properties of linearly quadratically controlled systems with complete state are crystallized in the canonical relations of symplecticity. The dualities of the Lyapunov matrices and the Riccati matrices are all derived from the symplectic relations. The complete state controller is shown to be superior, as the optimality indicates, to the conventional state feedback controller. The performance improvement is derived analytically with a quantitative formula.

AMS Subject Classification: 93B52, 93C05
Key Words: active control, linear quadratic optimal control, Hamiltonian matrix, Riccati matrix, Lyapunov matrix, canonical symplectivity

1. Introduction

The linear quadratic (LQ) optimal control methodologies provide a complete multivariable design and synthesis theory [6]. However, the conventional theory gives only an optimal control law for the linear plant without considering the external disturbance, that is,

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \quad \forall t \in [t_0, t_f], \qquad \mathbf{x}(t_0) = \mathbf{x}_0, \tag{1}$$

where $[t_0, t_f]$ is a time interval during which the plant is under a control force $\mathbf{u}(t)$.

Received: April 11, 2007

Upon maximizing the following performance index:

$$J = \frac{1}{2}\int_{t_0}^{t_f} [\mathbf{x}^T(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}^T(t)\mathbf{R}\mathbf{u}(t)]dt, \tag{2}$$

where $\mathbf{Q} \in \mathbb{R}^{n\times n}$ and $\mathbf{R} \in \mathbb{R}^{u\times u}$ are positive semidefinite and positive definite, respectively, and the superscript T stands for the transpose, the optimal control law is found to be

$$\mathbf{u}(t) = -\mathbf{R}^{-1}\mathbf{B}^T\mathbf{R}_x(t)\mathbf{x}(t) \quad \forall t \in [t_0, t_f], \tag{3}$$

where $\mathbf{R}_x$ is a Riccati matrix obtained by solving the following differential Riccati equation:

$$\dot{\mathbf{R}}_x + \mathbf{Q} + \mathbf{R}_x\mathbf{A} + \mathbf{A}^T\mathbf{R}_x - \mathbf{R}_x\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{R}_x = \mathbf{0}, \quad \mathbf{R}_x(t_f) = \mathbf{0}. \tag{4}$$

Equation (3) presents a state feedback (SF) control law. Unfortunately, the control law (3) upon applied to an externally disturbed plant as shown below in (7) is not the optimal one.

In the conventional state feedback control theory, equation (4) is solved numerically backwards in time, and with normal values of weighting matrices and structural properties, the Riccati matrix $\mathbf{R}_x(t)$ remains constant almost over the entire duration $[t_0, t_f]$ except that for very near the terminal time t_f; therefore, we usually set $\mathbf{R}_x$ to be a constant matrix satisfying the algebraic Riccati equation:

$$\mathbf{Q} + \mathbf{R}_x\mathbf{A} + \mathbf{A}^T\mathbf{R}_x - \mathbf{R}_x\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{R}_x = \mathbf{0}. \tag{5}$$

For its wide application in the modern control theory, there are many techniques to solve the above algebraic Riccati equation [2, 3, 4].

In Section 2 we show a linear quadratic control theory under external disturbances on the plant and stress its Hamiltonian form, which characterizes the controlled structures under disturbances. The properties of the Hamiltonian matrices are investigated in Section 3, and with the aid of these properties two kinds of decompositions of the systems are explored in Section 4. The performance improvement of the complete state (CS) controller over the conventional state feedback controller is studied in Section 5. Even, it is trivial that the conventional SF control will produce worse results than the complete state control, there exists no proof of the performance difference of these two controllers in the existing literature. We are going to give an explicit formula to assess the performance difference. We limit our scope to the mathematical properties of the controlled structure and do not consider its practical aspects.

2. LQ Optimal Control Against External Disturbance

In recent two decades a significant amount of effort has been devoted to the active control of engineering structures subject to external disturbances [1, 5, 9, 11, 12, 14, 13]; for a comprehensive review see Meirovitch and Stemple [9], Soong [11], and Yang and Soong [14].

Consider a linear engineering structure with n_s degrees of freedom subject to an external disturbance $\mathbf{w}(t)$ and a control force $\mathbf{u}(t)$. The equation of motion of the structure may be written as

$$\mathbf{M}\ddot{\mathbf{x}}_s(t) + \mathbf{C}\dot{\mathbf{x}}_s(t) + \mathbf{K}\mathbf{x}_s(t) = \mathbf{B}_s\mathbf{u}(t) + \mathbf{E}_s\mathbf{w}(t), \tag{6}$$

with some prescribed initial conditions $\mathbf{x}_s(t_0)$ and $\dot{\mathbf{x}}_s(t_0)$. Here t is a current time and t_0 is an initial time. The column matrix $\mathbf{x}_s$ is the relative displacement. A superposed dot indicates the time differentiation. The symmetric matrices $\mathbf{M}$, $\mathbf{C}$, $\mathbf{K}$ are, respectively, the mass, damping, and stiffness matrices of the structure. To be a structure $\mathbf{M}$ and $\mathbf{K}$ are positive definite and $\mathbf{C}$ is positive semidefinite. This structural dynamical problem of n_s equations of motion together with the initial conditions on n_s displacements and n_s velocities can be transformed to a state space description given as follows [5, 11]:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{E}\mathbf{w}(t), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \tag{7}$$

where

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{x}_s(t) \\ \dot{\mathbf{x}}_s(t) \end{bmatrix}, \quad \mathbf{x}_0 = \begin{bmatrix} \mathbf{x}_s(t_0) \\ \dot{\mathbf{x}}_s(t_0) \end{bmatrix} \tag{8}$$

and

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{M}^{-1}\mathbf{B}_s \end{bmatrix}, \ \mathbf{E} = \begin{bmatrix} \mathbf{0} \\ \mathbf{M}^{-1}\mathbf{E}_s \end{bmatrix}. \tag{9}$$

As usual $\mathbf{0}$ and $\mathbf{I}$ denote the zero and identity matrices, respectively, of appropriate orders as indicated in the context. Let $n = 2n_s$. In (7), the state $\mathbf{x} \in \mathbb{R}^n$, the disturbance $\mathbf{u} \in \mathbb{R}^u$, and the control $\mathbf{w} \in \mathbb{R}^w$ are all functions of time; the initial state $\mathbf{x}(t_0) \in \mathbb{R}^n$ is a constant column matrix; and $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times u}$, $\mathbf{E} \in \mathbb{R}^{n\times w}$ are constant matrices.

In order to minimize the performance index J in (2), which is subjected to the constraint (7), we can construct the Lagrangian

$$L = J + \int_{t_0}^{t_f} \boldsymbol{\lambda}^T(t)[\mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{E}\mathbf{w}(t) - \dot{\mathbf{x}}(t)]dt \tag{10}$$

and minimize L. The column matrix $\boldsymbol{\lambda} \in \mathbb{R}^n$ of n Lagrange multipliers is often referred to as the costate. The necessary conditions for the optimal functions

$\mathbf{x}(t)$, $\mathbf{u}(t)$, $\boldsymbol{\lambda}(t)$ that extremize L are readily found to be

$$\dot{\boldsymbol{\lambda}}(t) = -\mathbf{A}^T\boldsymbol{\lambda}(t) - \mathbf{Q}\mathbf{x}(t), \quad \boldsymbol{\lambda}(t_f) = \mathbf{0}, \tag{11}$$

$$\mathbf{u}(t) = -\mathbf{R}^{-1}\mathbf{B}^T\boldsymbol{\lambda}(t), \tag{12}$$

as well as the state equation (7). The control force would not only depend on the state but also on the internal costate, which is to be made clear later.

Thus, when the controller is governed by the optimal control law (12), the optimal state $\mathbf{x}(t)$ and costate $\boldsymbol{\lambda}(t)$ of the controlled structure under disturbance are governed by

$$\frac{d}{dt}\begin{bmatrix} \mathbf{x}(t) \\ \boldsymbol{\lambda}(t) \end{bmatrix} = \mathbf{H}\begin{bmatrix} \mathbf{x}(t) \\ \boldsymbol{\lambda}(t) \end{bmatrix} + \begin{bmatrix} \mathbf{E}\mathbf{w}(t) \\ \mathbf{0} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{x}(t_0) \\ \boldsymbol{\lambda}(t_f) \end{bmatrix} = \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{0} \end{bmatrix}, \tag{13}$$

where $\mathbf{H}$ is the system matrix of the controlled structure defined as

$$\mathbf{H} := \begin{bmatrix} \mathbf{A} & \mathbf{N} \\ -\mathbf{Q} & -\mathbf{A}^T \end{bmatrix}, \tag{14}$$

and

$$\mathbf{N} := -\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T \tag{15}$$

is negative semidefinite. Hence, $\mathbf{H}$ is a Hamiltonian matrix, which is, by definition, a square matrix satisfying

$$(\mathbf{J}\mathbf{H})^T = \mathbf{J}\mathbf{H}, \tag{16}$$

where

$$\mathbf{J} := \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix}. \tag{17}$$

As such, (13) is a constant coefficient linear Hamiltonian system defined in a symplectic space $\mathcal{V}$ endowed with the canonical metric $\mathbf{J}$. The point of the symplectic space $\mathcal{V}$ is

$$\boldsymbol{\nu} := \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix}, \tag{18}$$

which may be called the complete state [5], consisting of the state $\mathbf{x}$ and the costate $\boldsymbol{\lambda}$. In the complete state space $\mathcal{V}$, the complete state $\boldsymbol{\nu} \in \mathcal{V}$ of the controlled structure is governed by

$$\dot{\boldsymbol{\nu}}(t) = \mathbf{H}\boldsymbol{\nu}(t) + \mathbf{f}(t), \tag{19}$$

where

$$\mathbf{f}(t) := \begin{bmatrix} \mathbf{E}\mathbf{w}(t) \\ \mathbf{0} \end{bmatrix}.$$

In the conventional SF control theory the control force (3) does not meet the optimality condition if the external disturbance does not vanish. Although this methodology prevails in the optimal control literature known as a linear quadratic regulator, it is not the optimal one when the structure is subjected to an external disturbance. Conversely, in the complete state control theory the control force (12) is regulated by the costate which takes the external disturbance into account. In Section 5 we are going to evaluate the difference of that two control algorithms.

3. Symplectic Properties of Controlled Structures

We now investigate the properties of the controlled structure, i.e. the properties of the linear Hamiltonian system in the symplectic space of the complete states. The eigenvalues of a real Hamiltonian matrix are of four types [10, 8]: (1) the quadruples of truly complex eigenvalues $\pm\gamma_1 \pm \delta_1 i, \cdots, \pm\gamma_t \pm \delta_t i$, (2) the pairs of real eigenvalues $\pm\alpha_1, \cdots, \pm\alpha_r$, (3) the pairs of purely imaginary eigenvalues $\pm\beta_1 i, \cdots, \pm\beta_s i$, and (4) the eigenvalue 0.

It is known that for any real Hamiltonian matrix $\mathbf{H} \in \mathbb{R}^{2n\times 2n}$ there exists a symplectic matrix $\mathbf{\Psi} \in \mathbb{C}^{2n\times 2n}$, which is, by definition, a square matrix satisfying

$$\mathbf{\Psi}^T\mathbf{J}\mathbf{\Psi} = \mathbf{J}, \tag{20}$$

or equivalently,

$$\mathbf{\Psi}\mathbf{J}\mathbf{\Psi}^T = \mathbf{J}, \tag{21}$$

such that $\mathbf{H}$ is similar to a Hamiltonian matrix $\mathbf{\Omega} \in \mathbb{C}^{2n\times 2n}$, namely,

$$\mathbf{H}\mathbf{\Psi} = \mathbf{\Psi}\mathbf{\Omega}. \tag{22}$$

Equation (22) is called the relation of symplectic similarity, and (20) and (21) the relations of symplectic orthogonality.

Moreover, if the real Hamiltonian matrix $\mathbf{H}$ does not have the eigenvalue zero,[1] that is, $\mathbf{H}$ is nonsingular, it has been shown [15] that, by slightly modifying the Jordan canonical form, the $\mathbf{\Omega}$ matrix in the relation (22) can be chosen as the simplest Hamiltonian matrix

$$\mathbf{\Omega} = \begin{bmatrix} \mathbf{\Omega}_1 & \mathbf{0} \\ \mathbf{0} & -\mathbf{\Omega}_1^T \end{bmatrix}, \tag{23}$$

where $\mathbf{\Omega}_1 \in \mathbb{C}^{n\times n}$ is a block diagonal matrix, each diagonal block of which is a

[1] Note that the particular arrangement (23) is possible if the eigenvalue type (4) is excluded.

usual Jordan block corresponding to an eigenvalue of $\mathbf{H}$ with positive modulus and argument $\frac{\pi}{2} < \theta \leq \frac{3\pi}{2}$. Correspondingly, a slight modification of the matrix of generalized eigenvectors (namely, adjusting the signs and column positions of the generalized eigenvectors of $\mathbf{H}$) can result in a symplectic matrix $\mathbf{\Psi}$. Accordingly, we make available a (complex) canonical form of the relation of symplectic similarity (22), in which the complex-valued $\mathbf{\Omega} \in \mathbb{C}^{2n\times 2n}$ is such a modified Jordan matrix as to be a Hamiltonian matrix in the form (23), and the complex-valued $\mathbf{\Psi} \in \mathbb{C}^{2n\times 2n}$ is such a modified generalized eigenvector matrix as to be symplectic and hence to satisfy both the symplectic orthogonality relations (20) and (21). These altogether may be called the *complex canonical relations of symplecticity.*

Furthermore, if more restrictively the real Hamiltonian matrix $\mathbf{H}$ is nonsingular,[2] it has been shown [10, 7] that the $\mathbf{\Omega}$ matrix in the relation (22) can be chosen as a *real* Hamiltonian matrix in the form (23) but in which $\mathbf{\Omega}_1$ is taken in *real* Jordan form corresponding to those eigenvalues of $\mathbf{H}$ which have negative real parts. Correspondingly, $\mathbf{\Psi}$ is a *real* symplectic matrix. Accordingly, we make available a (real) canonical form of the relation of symplectic similarity (22), in which the real-valued $\mathbf{\Omega} \in \mathbb{R}^{2n\times 2n}$ is such a modified real Jordan matrix as to be a Hamiltonian matrix in the form (23), and the real-valued $\mathbf{\Psi} \in \mathbb{R}^{2n\times 2n}$ is such a modified generalized eigenvector matrix as to be symplectic and hence to satisfy both the symplectic orthogonality relations (20) and (21). These altogether will be referred to as the *real canonical relations of symplecticity.*

In the last three paragraphs we have step by step considered three levels of $\mathbf{H}$'s. This progressive consideration is for the sake of clarity. However, what we really need in the remainder of this section is the last two — the real Hamiltonian matrix $\mathbf{H}$ is nonsingular. Now, corresponding to the particular arrangement (23), let us partition $\mathbf{\Psi}$ into four $n \times n$ submatrices:

$$\mathbf{\Psi} = \begin{bmatrix} \mathbf{\Psi}_{xx} & \mathbf{\Psi}_{x\lambda} \\ \mathbf{\Psi}_{\lambda x} & \mathbf{\Psi}_{\lambda\lambda} \end{bmatrix}, \tag{24}$$

and define the pairs of the Lyapunov matrices, the Riccati matrices, the similarity matrices and the closed-loop matrices, successively, as follows:

$$\mathbf{L}_x := \mathbf{\Psi}_{xx}\mathbf{\Psi}_{x\lambda}^T, \tag{25}$$

$$\mathbf{L}_\lambda := \mathbf{\Psi}_{\lambda\lambda}\mathbf{\Psi}_{\lambda x}^T, \tag{26}$$

$$\mathbf{R}_x := \mathbf{\Psi}_{\lambda x}\mathbf{\Psi}_{xx}^{-1}, \tag{27}$$

[2]Here, the eigenvalue types (3) and (4) are excluded, so the particular arrangement (23) is valid.

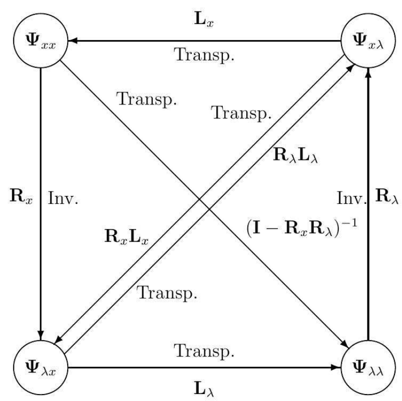

Figure 1: The relation diagram shows the dualities between the two Ricatti matrices and the two Lyapunov matrices.

$$\mathbf{R}_\lambda := \mathbf{\Psi}_{x\lambda}\mathbf{\Psi}_{\lambda\lambda}^{-1}, \tag{28}$$
$$\mathbf{S}_x := \mathbf{\Psi}_{\lambda\lambda}\mathbf{\Psi}_x^T, \tag{29}$$
$$\mathbf{S}_\lambda := \mathbf{\Psi}_x\mathbf{\Psi}_{\lambda\lambda}^T, \tag{30}$$
$$\mathbf{A}_x := \mathbf{A}+\mathbf{N}\mathbf{R}_x, \tag{31}$$
$$\mathbf{A}_\lambda := \mathbf{A}^T+\mathbf{Q}\mathbf{R}_\lambda. \tag{32}$$

In Figure 1 the first six relations are represented by a flow diagram, in which the dualities between $\mathbf{R}_x$ and $\mathbf{R}_\lambda$, and $\mathbf{L}_x$ and $\mathbf{L}_\lambda$ can be seen. With the last four definitions, the symplectic similarity relation (22) between the two Hamiltonian matrices $\mathbf{H}$ in the form (14) and $\mathbf{\Omega}$ in the form (23) is equivalent to the eigenproblems for $\mathbf{A}_x$ and $\mathbf{A}_\lambda$ and the algebraic Riccati equations for $\mathbf{R}_x$ and $\mathbf{R}_\lambda$, respectively, as follows (see Appendix A):

$$\mathbf{A}_x\mathbf{\Psi}_{xx} = \mathbf{\Psi}_{xx}\mathbf{\Omega}_1, \tag{33}$$
$$\mathbf{A}_\lambda\mathbf{\Psi}_{\lambda\lambda} = \mathbf{\Psi}_{\lambda\lambda}\mathbf{\Omega}_1^T, \tag{34}$$
$$\mathbf{Q}+\mathbf{R}_x\mathbf{A}+\mathbf{A}^T\mathbf{R}_x+\mathbf{R}_x\mathbf{N}\mathbf{R}_x = \mathbf{0}, \tag{35}$$
$$\mathbf{N}+\mathbf{A}\mathbf{R}_\lambda+\mathbf{R}_\lambda\mathbf{A}^T+\mathbf{R}_\lambda\mathbf{Q}\mathbf{R}_\lambda = \mathbf{0}. \tag{36}$$

In addition, the symplectic orthogonality relation (20) is equivalent to

$$\mathbf{\Psi}_{xx}^T\mathbf{\Psi}_{\lambda\lambda} = \mathbf{\Psi}_{\lambda x}^T\mathbf{\Psi}_{x\lambda}+\mathbf{I}, \tag{37}$$
$$\mathbf{\Psi}_{xx}^T\mathbf{\Psi}_{\lambda x} = \mathbf{\Psi}_{\lambda x}^T\mathbf{\Psi}_{xx}, \tag{38}$$
$$\mathbf{\Psi}_{\lambda\lambda}^T\mathbf{\Psi}_{x\lambda} = \mathbf{\Psi}_{x\lambda}^T\mathbf{\Psi}_{\lambda\lambda}. \tag{39}$$

The last two equations can be used to prove the symmetries of the Riccati

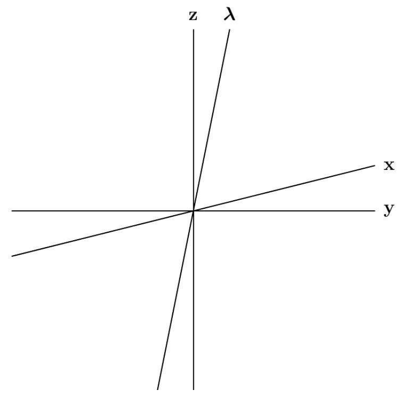

Figure 2: A schematic diagram showing that the changes of bases and coordinates in the symplectic space $\mathcal{V}$ of the complete states $\mathcal{V}$, the coordinates $(\mathbf{x}, \boldsymbol{\lambda})$ being first changed to $(\mathbf{x}, \mathbf{z}_x)$ and then to $(\mathbf{y}_x, \mathbf{z}_x)$. The perpendicularity of the $\mathbf{y}_x$-axis and the $\mathbf{z}_x$-axis hints schematically the decoupling $\mathcal{V} = \mathcal{V}_{y_x} \oplus \mathcal{V}_{z_x}$ of the evolution of the internal state $\mathbf{y}_x \in \mathcal{V}_{y_x}$ and the internal costate $\mathbf{z}_x \in \mathcal{V}_{z_x}$.

matrices $\mathbf{R}_x$ and $\mathbf{R}_\lambda$. The symplectic orthogonality relation (21) is equivalent to

$$\boldsymbol{\Psi}_{\lambda\lambda}\boldsymbol{\Psi}_{xx}^T = \boldsymbol{\Psi}_{\lambda x}\boldsymbol{\Psi}_{x\lambda}^T + \mathbf{I}, \tag{40}$$

$$\boldsymbol{\Psi}_{xx}\boldsymbol{\Psi}_{x\lambda}^T = \boldsymbol{\Psi}_{x\lambda}\boldsymbol{\Psi}_{xx}^T, \tag{41}$$

$$\boldsymbol{\Psi}_{\lambda\lambda}\boldsymbol{\Psi}_{\lambda x}^T = \boldsymbol{\Psi}_{\lambda x}\boldsymbol{\Psi}_{\lambda\lambda}^T. \tag{42}$$

The last two equations can be used to prove the symmetries of the Lyapunov matrices $\mathbf{L}_x$ and $\mathbf{L}_\lambda$.

In the above we have shown that the Riccati matrices $\mathbf{R}_x$ and $\mathbf{R}_\lambda$ defined by (27) and (28) satisfying the algebraic Riccati equations (35) and (36) is merely a direct consequence of the canonical relations of symplecticity. In Appendix A we prove the following pair of Lyapunov equations:

$$\mathbf{N} + \mathbf{A}_x\mathbf{L}_x + \mathbf{L}_x\mathbf{A}_x^T = \mathbf{0}, \tag{43}$$

$$\mathbf{Q} + \mathbf{A}_\lambda\mathbf{L}_\lambda + \mathbf{L}_\lambda\mathbf{A}_\lambda^T = \mathbf{0}, \tag{44}$$

and in the proofs we observe that the Lyapunov matrices $\mathbf{L}_x$ and $\mathbf{L}_\lambda$ defined by (25) and (26) satisfying the Lyapunov equations (43) and (44) is also a consequence of the canonical relations of symplecticity.

From (29), (40), (27) and (25) it follows that

$$\mathbf{S}_x = \mathbf{I} + \mathbf{R}_x \mathbf{L}_x. \tag{45}$$

Similarly, from (30), (40), (28) and (26) it follows that

$$\mathbf{S}_\lambda = \mathbf{I} + \mathbf{R}_\lambda \mathbf{L}_\lambda. \tag{46}$$

In Appendix A we also prove the following similarities between $\mathbf{A}_\lambda$ and $\mathbf{A}_x$:

$$\mathbf{S}_x \mathbf{A}_x^T = \mathbf{A}_\lambda \mathbf{S}_x, \tag{47}$$

$$\mathbf{S}_\lambda \mathbf{A}_\lambda^T = \mathbf{A}_x \mathbf{S}_\lambda. \tag{48}$$

In this section we have seen that the real canonical relations of symplecticity imply the properties of the controlled structures, summarize the *ten identities* (33)-(36) and (43)-(48) and facilitate the calculations of the *eight matrices* (25)-(32).

4. Decompositions of the Symplectic Space

With the properties obtained in the last section, we seek in this section the changes of variables, more precisely, the changes of bases and coordinates in the symplectic space of the complete states. It is preferable not only theoretically but also computationally for a change of variables to preserve the Hamiltonian structure. In Subsections 4.1 and 4.2 the Hamiltonian matrix $\mathbf{H}$ is assumed to be nonsingular.

4.1. Block Triangulation

To have such a structure-preserving transformation (in fact a symplectic basis change), let us define the internal costate [5] by

$$\mathbf{z}_x(t) := \boldsymbol{\lambda}(t) - \mathbf{R}_x \mathbf{x}(t), \tag{49}$$

where the Riccati matrix $\mathbf{R}_x$ can be calculated directly from its definition (27), when $\boldsymbol{\Psi}$ is determined by solving the $\mathbf{H}$ eigenproblem and then putting the generalized eigenvectors in $\boldsymbol{\Psi}$ according to the canonical relations of symplecticity. In other words, we perform a symplectic transformation

$$\begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{R}_x & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{z}_x \end{bmatrix} \tag{50}$$

such that the state-costate equation (13) is transformed, upon utilizing the Riccati equation (35), to the state-internal costate equation:

$$\frac{d}{dt}\begin{bmatrix}\mathbf{x}(t)\\ \mathbf{z}_x(t)\end{bmatrix}=\begin{bmatrix}\mathbf{A}_x & \mathbf{N}\\ \mathbf{0} & -\mathbf{A}_x^T\end{bmatrix}\begin{bmatrix}\mathbf{x}(t)\\ \mathbf{z}_x(t)\end{bmatrix}+\begin{bmatrix}\mathbf{E}\mathbf{w}(t)\\ -\mathbf{R}_x\mathbf{E}\mathbf{w}(t)\end{bmatrix}. \tag{51}$$

Since $\mathbf{N}$ is symmetric, the system matrix of (51), like $\mathbf{H}$, is also a Hamiltonian matrix; hence, it is still a constant coefficient linear Hamiltonian system.

4.2. Block Diagonalization

To eliminate the coupling in (51), let us define the internal state

$$\mathbf{y}_x(t) := \mathbf{x}(t) - \mathbf{L}_x\mathbf{z}_x(t) \tag{52}$$

and perform another symplectic transformation

$$\begin{bmatrix}\mathbf{x}\\ \mathbf{z}_x\end{bmatrix}=\begin{bmatrix}\mathbf{I} & \mathbf{L}_x\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{y}_x\\ \mathbf{z}_x\end{bmatrix} \tag{53}$$

such that, upon utilizing the Lyapunov equation (43), (51) is transformed to

$$\frac{d}{dt}\begin{bmatrix}\mathbf{y}_x(t)\\ \mathbf{z}_x(t)\end{bmatrix}=\begin{bmatrix}\mathbf{A}_x & \mathbf{0}\\ \mathbf{0} & -\mathbf{A}_x^T\end{bmatrix}\begin{bmatrix}\mathbf{y}_x(t)\\ \mathbf{z}_x(t)\end{bmatrix}+\begin{bmatrix}\mathbf{S}_x^T\mathbf{E}\mathbf{w}(t)\\ -\mathbf{R}_x\mathbf{E}\mathbf{w}(t)\end{bmatrix}, \tag{54}$$

where $\mathbf{S}_x^T = \mathbf{I} + \mathbf{L}_x\mathbf{R}_x$ has been used. Again like $\mathbf{H}$, the system matrix of (54) is also a Hamiltonian matrix, as is easily verified; hence, it remains to be a constant coefficient linear Hamiltonian system. Moreover, the equations of $\mathbf{y}_x$ and $\mathbf{z}_x$ are decoupled. See Figure 2 for a schematic explanation of the coordinate changes which result in the decoupled decomposition of the complete state space $\mathcal{V}$ into the internal state space $\mathcal{V}_{y_x}$ and the internal costate space $\mathcal{V}_{z_x}$.

Mathematically speaking, the symplectic space $\mathcal{V}$ of the complete states is the direct sum $\mathcal{V} = \mathcal{V}_{y_x} \oplus \mathcal{V}_{z_x}$ of the invariant subspaces $\mathcal{V}_{y_x}$ and $\mathcal{V}_{z_x}$. We have furnished a rather detailed derivation of the formulae available in the invariant space $(\mathbf{z}_x, \mathbf{y}_x)$. However, there correspondingly exists a dual representation in the invariant space $(\mathbf{z}_\lambda, \mathbf{y}_\lambda)$ (see Appendix B):

$$\frac{d}{dt}\begin{bmatrix}\mathbf{y}_\lambda(t)\\ \mathbf{z}_\lambda(t)\end{bmatrix}=\begin{bmatrix}\mathbf{A}_\lambda^T & \mathbf{0}\\ \mathbf{0} & -\mathbf{A}_\lambda\end{bmatrix}\begin{bmatrix}\mathbf{y}_\lambda(t)\\ \mathbf{z}_\lambda(t)\end{bmatrix}+\begin{bmatrix}\mathbf{E}\mathbf{w}(t)\\ -\mathbf{L}_\lambda\mathbf{E}\mathbf{w}(t)\end{bmatrix}. \tag{55}$$

Again like $\mathbf{H}$, the system matrix of (55) is also a constant coefficient linear Hamiltonian matrix, as is easily verified.

From (50), (53), (B2) and (B5) the following relations can be proved:

$$\mathbf{y}_x = (\mathbf{S}_\lambda^2 - \mathbf{L}_x\mathbf{L}_\lambda)\mathbf{y}_\lambda, \quad \mathbf{y}_\lambda = (\mathbf{I} - \mathbf{R}_\lambda\mathbf{R}_x)\mathbf{y}_x, \tag{56}$$

$$\mathbf{z}_x = (\mathbf{I}-\mathbf{R}_x\mathbf{R}_\lambda)\mathbf{z}_\lambda, \quad \mathbf{z}_\lambda = (\mathbf{S}_x^2-\mathbf{L}_\lambda\mathbf{L}_x)\mathbf{z}_x, \tag{57}$$

where $\mathbf{S}_x$ and $\mathbf{S}_\lambda$ are defined in (29) and (30), respectively. The above two equations characterize the transformations between these two invariant spaces $(\mathbf{z}_x, \mathbf{y}_x)$ and $(\mathbf{z}_\lambda, \mathbf{y}_\lambda)$.

5. Difference of Performance Indices

In this section the performances of the state feedback (SF) and the complete state (CS) controllers are compared analytically. In the conventional SF control theory the internal costate $\mathbf{z}_x(t)$ is neglected; that is, the conventional SF controller is regulated by the response state alone and not by $\mathbf{z}_x(t)$ at all. Consequently, (49) and (12) are approximated by

$$\boldsymbol{\lambda}(t) = \mathbf{R}_x\mathbf{x}(t), \tag{58}$$

$$\mathbf{u}(t) = -\mathbf{R}^{-1}\mathbf{B}^T\mathbf{R}_x\mathbf{x}(t). \tag{59}$$

As mentioned in Section 1, the control law due to (59) does not meet the optimality condition if the external disturbance does not vanish. Although this methodology prevails in the optimal control literature known as a linear quadratic regulator (LQR), it is not the optimal one when the structure is subjected to an external disturbance.

By using

$$\int_{t_0}^{t_f}\frac{d}{d\tau}[\mathbf{x}^T(\tau)\mathbf{R}_x\mathbf{x}(\tau)]d\tau = \mathbf{x}^T(t_f)\mathbf{R}_x\mathbf{x}(t_f)-\mathbf{x}^T(t_0)\mathbf{R}_x\mathbf{x}(t_0), \tag{60}$$

and the first equation in (7) and (4), we may rewrite the performance index (2) as

$$J = \frac{1}{2}\mathbf{x}^T(t_0)\mathbf{R}_x\mathbf{x}(t_0) + \int_{t_0}^{t_f}\mathbf{w}^T(\tau)\mathbf{E}^T\mathbf{R}_x\mathbf{x}(\tau)d\tau + \frac{1}{2}\int_{t_0}^{t_f}[\mathbf{u}(\tau)+\mathbf{R}^{-1}\mathbf{B}^T\mathbf{R}_x\mathbf{x}(\tau)]^T\mathbf{R}[\mathbf{u}(\tau)+\mathbf{R}^{-1}\mathbf{B}^T\mathbf{R}_x\mathbf{x}(\tau)]d\tau, \tag{61}$$

where the term $-\mathbf{x}^T(t_f)\mathbf{R}_x\mathbf{x}(t_f)/2$ disappears due to $\mathbf{R}_x(t_f)=\mathbf{0}$.

The state equations for CS and SF controls are, respectively,

$$\dot{\mathbf{x}}_{CS}(t) = \mathbf{A}\mathbf{x}_{CS}(t)+\mathbf{B}\mathbf{u}_{CS}(t)+\mathbf{E}\mathbf{w}(t), \tag{62}$$

$$\dot{\mathbf{x}}_{SF}(t) = \mathbf{A}\mathbf{x}_{SF}(t)+\mathbf{B}\mathbf{u}_{SF}(t)+\mathbf{E}\mathbf{w}(t), \tag{63}$$

and the corresponding control forces are, respectively,

$$\mathbf{u}_{CS}(t) = -\mathbf{R}^{-1}\mathbf{B}^T\mathbf{R}_x\mathbf{x}_{CS}(t)-\mathbf{R}^{-1}\mathbf{B}^T\mathbf{z}_x(t), \tag{64}$$

$$\mathbf{u}_{SF}(t) = -\mathbf{R}^{-1}\mathbf{B}^T\mathbf{R}_x\mathbf{x}_{SF}(t). \tag{65}$$

Therefore, the performance index for the CS controller is

$$J_{CS} = \frac{1}{2}\mathbf{x}_{CS}^T(t_0)\mathbf{R}_x\mathbf{x}_{CS}(t_0) - \frac{1}{2}\int_{t_0}^{t_f} \mathbf{z}_x^T(\tau)\mathbf{N}\mathbf{z}_x(\tau)d\tau + \int_{t_0}^{t_f} \mathbf{w}^T(\tau)\mathbf{E}^T\mathbf{R}_x\mathbf{x}_{CS}(\tau)d\tau, \tag{66}$$

and that for the SF controller is

$$J_{SF} = \frac{1}{2}\mathbf{x}_{SF}^T(t_0)\mathbf{R}_x\mathbf{x}_{SF}(t_0) + \int_{t_0}^{t_f} \mathbf{w}^T(\tau)\mathbf{E}^T\mathbf{R}_x\mathbf{x}_{SF}(\tau)d\tau. \tag{67}$$

Define

$$\Delta J := J_{SF} - J_{CS} \tag{68}$$

as the difference of the performance indices, and

$$\Delta\mathbf{x}(t) := \mathbf{x}_{SF}(t) - \mathbf{x}_{CS}(t), \qquad \Delta\mathbf{x}(t_0) = \mathbf{0} \tag{69}$$

as the difference of the states for the SF and CS controls. With the aid of (62), (63), (64) and (65) the differential equation of the state difference is found to be

$$\frac{d}{dt}\Delta\mathbf{x}(t) = \mathbf{A}_x\Delta\mathbf{x}(t) - \mathbf{N}\mathbf{z}_x(t), \tag{70}$$

where $\mathbf{z}_x$ is governed by

$$\dot{\mathbf{z}}_x(t) = -\mathbf{A}_x^T\mathbf{z}_x(t) - \mathbf{R}_x\mathbf{E}\mathbf{w}(t). \tag{71}$$

The solution of (70) can be deduced from (71) and the second equation in (69) as follows:

$$\Delta\mathbf{x}(t) = \int_{t_0}^{t}\int_{t_f}^{\tau} e^{\mathbf{A}_x(t-\tau)}\mathbf{N}e^{-\mathbf{A}_x^T(\tau-\eta)}\mathbf{R}_x\mathbf{E}\mathbf{w}(\eta)d\eta d\tau. \tag{72}$$

In view of (66)-(68) and the second equation in (69), the difference of the performance indices can be written as

$$\Delta J = \frac{1}{2}\int_{t_0}^{t_f} \mathbf{z}_x^T(\xi)\mathbf{N}\mathbf{z}_x(\xi)d\xi + \int_{t_0}^{t_f} \mathbf{w}^T(\eta)\mathbf{E}^T\mathbf{R}_x\Delta\mathbf{x}(\eta)d\eta. \tag{73}$$

Inserting the solution of (71) with $\mathbf{z}_x(t_f) = \mathbf{0}$ and (72) into the above equation leads to

$$\Delta J = \frac{1}{2}\int_{t_0}^{t_f}\int_{t_f}^{\xi}\int_{t_f}^{\xi} \mathbf{w}^T(\tau)\mathbf{E}^T\mathbf{R}_x e^{-\mathbf{A}_x(\xi-\tau)}\mathbf{N}e^{-\mathbf{A}_x^T(\xi-\eta)}\mathbf{R}_x\mathbf{E}\mathbf{w}(\eta)d\tau d\eta d\xi$$

$$+\int_{t_0}^{t_f}\int_{t_0}^{\eta}\int_{t_f}^{\xi}\mathbf{w}^T(\tau)\mathbf{E}^T\mathbf{R}_x e^{-\mathbf{A}_x(\xi-\tau)}\mathbf{N}e^{-\mathbf{A}_x^T(\xi-\eta)}\mathbf{R}_x\mathbf{E}\mathbf{w}(\eta)d\tau d\xi d\eta. \quad (74)$$

Let us define an energy-like measure (L_2-norm) of the internal costate:

$$J_{\mathbf{z}} := -\frac{1}{2}\int_{t_0}^{t_f}\mathbf{z}_x^T(\tau)\mathbf{N}\mathbf{z}_x(\tau)d\tau. \quad (75)$$

Recalling that $\mathbf{N}$ is negative semidefinite; hence, $J_{\mathbf{z}}$ is a nonnegative functional of $\mathbf{z}_x$. Substituting the solution of (71) with $\mathbf{z}_x(t_f) = \mathbf{0}$ into the above equation, we have

$$J_{\mathbf{z}} = \frac{1}{2}\int_{t_0}^{t_f}\int_{t_f}^{\xi}\int_{t_f}^{\xi} f(\xi,\tau,\eta)d\tau d\eta d\xi, \quad (76)$$

where

$$f(\xi,\tau,\eta) := -\mathbf{w}^T(\tau)\mathbf{E}^T\mathbf{R}_x e^{-\mathbf{A}_x(\xi-\tau)}\mathbf{N}e^{-\mathbf{A}_x^T(\xi-\eta)}\mathbf{R}_x\mathbf{E}\mathbf{w}(\eta). \quad (77)$$

By inspection, $f(\xi,\tau,\eta)$ is symmetric with respect to the last two variables, i.e., $f(\xi,\tau,\eta) = f(\xi,\eta,\tau)$, and $J_{\mathbf{z}} > 0$, $\forall\ \mathbf{z}_x \neq \mathbf{0}$. Accordingly, (74) can be written as

$$\Delta J = \frac{1}{2}\int_{t_f}^{t_0}\int_{t_f}^{\xi}\int_{t_f}^{\xi} f(\xi,\tau,\eta)d\tau d\eta d\xi + \int_{t_f}^{t_0}\int_{t_0}^{\eta}\int_{t_f}^{\xi} f(\xi,\tau,\eta)d\tau d\xi d\eta. \quad (78)$$

By using the triple integral identity (C1) derived in Appendix C, (78) reduces to

$$\Delta J = \int_{t_f}^{t_0}\int_{\eta}^{t_0}\int_{\tau}^{t_0} f(\xi,\tau,\eta)d\xi d\tau d\eta + \int_{t_f}^{t_0}\int_{t_0}^{\eta}\int_{t_f}^{\xi} f(\xi,\tau,\eta)d\tau d\xi d\eta, \quad (79)$$

and (75) reduces to

$$J_{\mathbf{z}} = -\int_{t_f}^{t_0}\int_{\eta}^{t_0}\int_{\tau}^{t_0} f(\xi,\tau,\eta)d\xi d\tau d\eta. \quad (80)$$

By splitting the integral $\int_{t_0}^{\eta}$ of the second term on the right-hand side of (79) into $\int_{t_0}^{t_f} + \int_{t_f}^{\eta}$, we have

$$\Delta J = \int_{t_f}^{t_0}\int_{\eta}^{t_0}\int_{\tau}^{t_0} f(\xi,\tau,\eta)d\xi d\tau d\eta$$
$$+ \int_{t_f}^{t_0}\int_{t_0}^{t_f}\int_{t_f}^{\xi} f(\xi,\tau,\eta)d\tau d\xi d\eta + \int_{t_f}^{t_0}\int_{t_f}^{\eta}\int_{t_f}^{\xi} f(\xi,\tau,\eta)d\tau d\xi d\eta. \quad (81)$$

Interchanging the integration order of the above two inner integrals in the second term and reversing the integration direction of the outer integral, we obtain

$$\Delta J = \int_{t_f}^{t_0}\int_{\eta}^{t_0}\int_{\tau}^{t_0} f(\xi,\tau,\eta)d\xi d\tau d\eta$$
$$- \int_{t_f}^{t_0}\int_{t_f}^{t_0}\int_{\tau}^{t_0} f(\xi,\tau,\eta)d\xi d\tau d\eta + \int_{t_f}^{t_0}\int_{t_f}^{\eta}\int_{t_f}^{\xi} f(\xi,\tau,\eta)d\tau d\xi d\eta. \quad (82)$$

Furthermore, by splitting the second integral $\int_{t_f}^{t_0}$ in the second term of the right-hand side into $\int_{t_f}^{\eta} + \int_{\eta}^{t_0}$, and canceling the common terms, we have

$$\Delta J = -\int_{t_f}^{t_0}\int_{t_f}^{\eta}\int_{\tau}^{t_0} f(\xi,\tau,\eta)d\xi d\tau d\eta + \int_{t_f}^{t_0}\int_{t_f}^{\eta}\int_{t_f}^{\xi} f(\xi,\tau,\eta)d\tau d\xi d\eta. \quad (83)$$

Now, interchanging the integration order of the above two inner integrals of the second term on the right-hand side, we obtain

$$\begin{aligned}\Delta J &= \int_{t_f}^{t_0}\int_{t_f}^{\eta}\int_{t_0}^{\tau} f(\xi,\tau,\eta)d\xi d\tau d\eta + \int_{t_f}^{t_0}\int_{t_f}^{\eta}\int_{\tau}^{\eta} f(\xi,\tau,\eta)d\xi d\tau d\eta \\ &= \int_{t_f}^{t_0}\int_{t_f}^{\eta}\int_{t_0}^{\eta} f(\xi,\tau,\eta)d\xi d\tau d\eta. \end{aligned} \quad (84)$$

By interchanging the integration order of the two outer integrals, we thus have

$$\Delta J = \int_{t_f}^{t_0}\int_{\tau}^{t_0}\int_{t_0}^{\eta} f(\xi,\tau,\eta)d\xi d\eta d\tau. \quad (85)$$

By interchanging the dummy indices τ and η, and utilizing the symmetric property of $f(\xi,\tau,\eta)$, we obtain

$$\Delta J = -\int_{t_f}^{t_0}\int_{\eta}^{t_0}\int_{\tau}^{t_0} f(\xi,\tau,\eta)d\xi d\tau d\eta. \quad (86)$$

Comparing the above result with (80), we eventually obtain

$$\Delta J = J_{\mathbf{z}} > 0, \quad \forall\ \mathbf{z}_x \neq \mathbf{0}. \quad (87)$$

One direct consequence is

$$J_{SF} - J_{CS} = J_{\mathbf{z}} > 0, \quad \forall\ \mathbf{z}_x \neq \mathbf{0}. \quad (88)$$

This result says that J_{CS} is always smaller than J_{SF}, that is,

$$J_{CS} < J_{SF}, \quad \forall\ \mathbf{z}_x \neq \mathbf{0}. \quad (89)$$

When the external disturbance is zero, i.e. $\mathbf{w} = \mathbf{0}$, one has $\mathbf{z}_x(t) = \mathbf{0}$, $t \in [t_0, t_f]$ by (71) and $\mathbf{z}_x(t_f) = \mathbf{0}$, and thus $J_{\mathbf{z}} = 0$ by (75), implying $J_{SF} = J_{CS}$. For this case the state feedback control law (59) is the best one; however, it cannot meet the optimality condition when the external disturbance appears.

The above proof does not recourse to any special assumption about the external disturbance $\mathbf{w}(t)$; therefore, the optimality of CS control and its superiority over SF control are confirmed for any external disturbance, no matter deterministic or random, Gaussian or non-Gaussian.

6. Concluding Remarks

It has been shown that the properties of the controlled structures are crystallized in the canonical relations of symplecticity, which, among others, summarize ten identities, including the Riccati and Lyapunov equations, and facilitate the calculations of six matrices, including the Riccati and Lyapunov matrices. The governing equation of the controlled structure is a linear Hamiltonian system, which can be block triangualized or block diagonalized. Although by itself sufficing to describe the linear structure, the state $\mathbf{x}$ must be supplemented with the internal costate $\mathbf{z}_x$ (or the costate $\boldsymbol{\lambda}$) for a complete state description of the LQ system, which contains the linear structure and the quadratic-form optimal controller. The complete state controller takes full advantage of the information of the actual disturbance.

The complete state controller rigorously satisfies the optimality condition. The performance improvement has been proved to be proportional to the L_2-norm of the internal costate. This genuinely optimal law improved the performance without exception when the linear structure is controlled quadratically to against the external disturbances.

References

[1] M. Abdel-Rohman, V.H. Quintana, H.H. Leipholz, Optimal control of civil engineering structures, *J. Eng. Mech. ASCE*, **106** (1980), 57-73.

[2] P. Benner, H. Fa β bender, An implicitly restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem, *Linear Algebra Appl.*, **263** (1997), 75-111.

[3] R. Byers, Solving the algebraic Riccati equation with the matrix sign function, *Linear Algebra Appl.*, **85** (1987), 267-279.

[4] E.K.-W. Chu, H.-Y. Fan, W.-W. Lin, A structure-preserving doubling algorithm for continuous-time algebraic Riccati equations, *Linear Algebra Appl.*, **396** (2005), 55-80.

[5] H.-K. Hong, C.-S. Liu, D.-Y. Liou, Complete state LQ optimal control of earthquake-excited structures, *Proc. Nat. Sci. Counc. ROC, A*, **18** (1994), 386-399.

[6] M.A. Johnson, M.J. Grimble, Recent trends in linear optimal quadratic multivariable control system design, *IEE Proceedings D, Control Theory and Application*, **134** (1987), 53-73.

[7] A.J. Laub, K. Meyer, Canonical forms for symplectic and Hamiltonian matrices, *Celestial Mech.*, **9** (1974), 213-238.

[8] C.-S. Liu, Nonlinear Lorentzian and Hamiltonian formulations and their relations, *Int. J. Appl. Math.*, **10** (2002), 59-97.

[9] L. Meirovitch, T.J. Stemple, Nonlinear control of structures in earthquakes, *J. Eng. Mech. ASCE*, **123** (1997), 1090-1095.

[10] K.R. Meyer, G.R. Hall, *Introduction to Hamiltonian Dynamical Systems and the N-body Problem*, Springer-Verlag, New York (1992).

[11] T.T. Soong, *Active Structural Control: Theory, Practise*, Longman Scientific and Technical, Essex, England (1990).

[12] J.N. Yang, A. Akbarpour, P. Ghaemmaghami, New optimal control algorithms for structural control, *J. Eng. Mech. ASCE*, **113** (1987), 1369-1387.

[13] J.T.P. Yao, S. Sae-Ung, Active control of building structures, *J. Eng. Mech. ASCE*, **104** (1978), 335-350.

[14] J.N. Yang, T.T. Soong, Recent advances in active control of civil engineering structures, *Prob. Eng. Mech.*, **3** (1988), 179-188.

[15] W. Zhong, On the reciprocal theorem and adjoint symplectic orthogonal relation, *Acta Mech. Sinica*, **24** (1992), 432-437, In Chinese.

Appendix A

Proof of (33)-(36). Substituting (14) for $\mathbf{H}$, (24) for $\mathbf{\Psi}$ and (23) for $\mathbf{\Omega}$ into (22), we obtain the following four equations:

$$\mathbf{A}\mathbf{\Psi}_{xx} + \mathbf{N}\mathbf{\Psi}_{\lambda x} = \mathbf{\Psi}_{xx}\mathbf{\Omega}_1, \tag{A1}$$

$$\mathbf{Q}\mathbf{\Psi}_{x\lambda} + \mathbf{A}^T\mathbf{\Psi}_{\lambda\lambda} = \mathbf{\Psi}_{xx}\mathbf{\Omega}_1^T, \tag{A2}$$

$$\mathbf{Q}\boldsymbol{\Psi}_{xx} + \mathbf{A}^T\boldsymbol{\Psi}_{\lambda x} = -\boldsymbol{\Psi}_{\lambda x}\boldsymbol{\Omega}_1, \tag{A3}$$

$$\mathbf{A}\boldsymbol{\Psi}_{x\lambda} + \mathbf{N}\boldsymbol{\Psi}_{\lambda\lambda} = -\boldsymbol{\Psi}_{x\lambda}\boldsymbol{\Omega}_1^T. \tag{A4}$$

Inserting $\boldsymbol{\Psi}_{\lambda x} = \mathbf{R}_x\boldsymbol{\Psi}_{xx}$ obtained from (27) into (A1) and recalling the definition (31) we get (33). Similarly, inserting $\boldsymbol{\Psi}_{x\lambda} = \mathbf{R}_\lambda\boldsymbol{\Psi}_{\lambda\lambda}$ obtained from (28) into (A2) and recalling the definition (32) we get (34). Now, inserting $\boldsymbol{\Psi}_{\lambda x} = \mathbf{R}_x\boldsymbol{\Psi}_{xx}$ into (A3), replacing $\boldsymbol{\Psi}_{xx}\boldsymbol{\Omega}_1$ by $\mathbf{A}_x\boldsymbol{\Psi}_{xx}$ due to (33), then using (31) for $\mathbf{A}_x$ and deleting $\boldsymbol{\Psi}_{xx}$ on both sides we get (35). Finally, inserting $\boldsymbol{\Psi}_{x\lambda} = \mathbf{R}_\lambda\boldsymbol{\Psi}_{\lambda\lambda}$ into (A4), replacing $\boldsymbol{\Psi}_{\lambda\lambda}\boldsymbol{\Omega}_1^T$ by $\mathbf{A}_\lambda\boldsymbol{\Psi}_{\lambda\lambda}$ due to (34), then using (32) for $\mathbf{A}_\lambda$ and deleting $\boldsymbol{\Psi}_{\lambda\lambda}$ on both sides we get (36). □

Proof of (43). Multiplying (37) by $\boldsymbol{\Psi}_{xx}^{-T}$ from the left and by $\boldsymbol{\Psi}_{\lambda\lambda}^{-1}$ from the right and considering the definition of $\mathbf{R}_\lambda$ and the definition and symmetry of $\mathbf{R}_x$, we have

$$\mathbf{I} = \mathbf{R}_x\mathbf{R}_\lambda + \boldsymbol{\Psi}_{xx}^{-T}\boldsymbol{\Psi}_{\lambda\lambda}^{-1}, \tag{A5}$$

which, upon considering the definition of $\mathbf{R}_\lambda$ and the definition and symmetry of $\mathbf{L}_x$, changes to

$$\mathbf{I} = \mathbf{R}_x\mathbf{R}_\lambda + \mathbf{L}_x^{-1}\mathbf{R}_\lambda \quad \text{or} \quad \mathbf{R}_\lambda^{-1} = \mathbf{R}_x + \mathbf{L}_x^{-1}. \tag{A6}$$

Multiplying (36) from the left and the right by $\mathbf{R}_\lambda^{-1}$ and noting the above equation, give

$$(\mathbf{R}_x + \mathbf{L}_x^{-1})\mathbf{N}(\mathbf{R}_x + \mathbf{L}_x^{-1}) + (\mathbf{R}_x + \mathbf{L}_x^{-1})\mathbf{A} + \mathbf{A}^T(\mathbf{R}_x + \mathbf{L}_x^{-1}) + \mathbf{Q} = \mathbf{0}.$$

Using (35) and multiplying from both the left and the right by $\mathbf{L}_x$, we obtain

$$\mathbf{N} + (\mathbf{A} + \mathbf{N}\mathbf{R}_x)\mathbf{L}_x + \mathbf{L}_x(\mathbf{A}^T + \mathbf{R}_x\mathbf{N}) = \mathbf{0}.$$

In view of (31) and the symmetries of $\mathbf{N}$ and $\mathbf{R}_x$, the above equation is just (43), thus completing the proof. □

Proof of (44). Upon considering the definition of $\mathbf{L}_\lambda$ and the definition and symmetry of $\mathbf{R}_x$, (A5) changes to

$$\mathbf{I} = \mathbf{R}_x\mathbf{R}_\lambda + \mathbf{R}_x\mathbf{L}_\lambda^{-1} \quad \text{or} \quad \mathbf{R}_x^{-1} = \mathbf{R}_\lambda + \mathbf{L}_\lambda^{-1}. \tag{A7}$$

Multiplying (35) from both the left and the right by $\mathbf{R}_x^{-1}$ and noting the above equation, give

$$(\mathbf{R}_\lambda + \mathbf{L}_\lambda^{-1})\mathbf{Q}(\mathbf{R}_\lambda + \mathbf{L}_\lambda^{-1}) + \mathbf{A}(\mathbf{R}_\lambda + \mathbf{L}_\lambda^{-1}) + (\mathbf{R}_\lambda + \mathbf{L}_\lambda^{-1})\mathbf{A}^T + \mathbf{N} = \mathbf{0}.$$

Using (36) and multiplying from both the left and the right by $\mathbf{L}_\lambda$, we have

$$\mathbf{Q} + (\mathbf{A}^T + \mathbf{Q}\mathbf{R}_\lambda)\mathbf{L}_\lambda + \mathbf{L}_\lambda(\mathbf{A} + \mathbf{R}_\lambda\mathbf{Q}) = \mathbf{0}.$$

In view of (32) and the symmetries of $\mathbf{Q}$ and $\mathbf{R}_\lambda$, the above equation is nothing but (44), thus completing the proof. □

Proof of (47). From (32) and (36) it follows that

$$\mathbf{A}_\lambda = -\mathbf{R}_\lambda^{-1}(\mathbf{N} + \mathbf{A}\mathbf{R}_\lambda). \tag{A8}$$

Using

$$\mathbf{R}_\lambda = \mathbf{L}_x(\mathbf{I} + \mathbf{R}_x\mathbf{L}_x)^{-1},$$

which follows from (A6), (A8) changes to

$$\mathbf{A}_\lambda = -(\mathbf{I} + \mathbf{R}_x\mathbf{L}_x)\mathbf{L}_x^{-1}[\mathbf{N} + \mathbf{A}\mathbf{L}_x(\mathbf{I} + \mathbf{R}_x\mathbf{L}_x)^{-1}], \tag{A9}$$

which can be rearranged to

$$\mathbf{A}_\lambda = -(\mathbf{I} + \mathbf{R}_x\mathbf{L}_x)\mathbf{L}_x^{-1}[\mathbf{N} + (\mathbf{A} + \mathbf{N}\mathbf{R}_x)\mathbf{L}_x](\mathbf{I} + \mathbf{R}_x\mathbf{L}_x)^{-1}. \tag{A10}$$

In view of (31) and (43) we obtain

$$\mathbf{A}_\lambda = (\mathbf{I} + \mathbf{R}_x\mathbf{L}_x)\mathbf{L}_x^{-1}\mathbf{L}_x\mathbf{A}_x^T(\mathbf{I} + \mathbf{R}_x\mathbf{L}_x)^{-1}.$$

This is just (47) by the identity $\mathbf{L}_x^{-1}\mathbf{L}_x = \mathbf{I}$. So the proof is completed. □

Proof of (48). Comparing the first equations in (A6) and (A7) we obtain

$$\mathbf{R}_\lambda\mathbf{L}_\lambda = \mathbf{L}_x\mathbf{R}_x. \tag{A11}$$

Taking the transpose of (47) and using the above equation we can prove (48). □

Appendix B

To solve the optimal control problem, we seek in this appendix another type of change of variables, more precisely, a change of bases and coordinates in the symplectic space of the complete states dual to $(\mathbf{y}_x, \mathbf{z}_x)$, that is $(\mathbf{y}_\lambda, \mathbf{z}_\lambda)$. For this purpose let us define the internal state

$$\mathbf{y}_\lambda(t) := \mathbf{x}(t) - \mathbf{R}_\lambda\boldsymbol{\lambda}(t), \tag{B1}$$

where the Riccati matrix $\mathbf{R}_\lambda$ can be calculated directly from its definition (28). In other words, we perform a symplectic transformation

$$\begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{R}_\lambda \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{y}_\lambda \\ \boldsymbol{\lambda} \end{bmatrix}, \tag{B2}$$

such that the state-costate equation (13) is transformed, upon utilizing the Riccati equation (36), to the internal costate-costate equation:

$$\frac{d}{dt}\begin{bmatrix} \mathbf{y}_\lambda \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_\lambda^T & \mathbf{0} \\ -\mathbf{Q} & -\mathbf{A}_\lambda \end{bmatrix} \begin{bmatrix} \mathbf{y}_\lambda \\ \boldsymbol{\lambda} \end{bmatrix} + \begin{bmatrix} \mathbf{E}\mathbf{w}(t) \\ \mathbf{0} \end{bmatrix}. \tag{B3}$$

Since $\mathbf{Q}$ is symmetric, the system matrix of (B3), like $\mathbf{H}$, is also a Hamiltonian matrix, as is easily verified; hence, (B3) remains to be a constant coefficient linear Hamiltonian system.

Let us define the internal costate by

$$\mathbf{z}_\lambda(t) := \boldsymbol{\lambda}(t) - \mathbf{L}_\lambda \mathbf{y}_\lambda(t), \tag{B4}$$

and perform another symplectic transformation

$$\begin{bmatrix} \mathbf{y}_\lambda \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{L}_\lambda & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{y}_\lambda \\ \mathbf{z}_\lambda \end{bmatrix}, \tag{B5}$$

such that, upon utilizing the Lyapunov equation (44), (B3) is further transformed to (55).

Appendix C

For an integrable function $F(t, \xi, \zeta)$ with the symmetric property $F(t, \xi, \zeta) = F(t, \zeta, \xi)$, the following triple-integral identity holds:

$$\int_\tau^t \int_\tau^\eta \int_\tau^\eta F(\eta, \xi, \zeta) d\xi d\zeta d\eta = 2 \int_\tau^t \int_\zeta^t \int_\xi^t F(\eta, \xi, \zeta) d\eta d\xi d\zeta. \tag{C1}$$

Proof. First, we have

$$\int_\tau^t \int_\zeta^t F(t, \xi, \zeta) d\xi d\zeta = \int_\tau^t \int_\tau^\xi F(t, \xi, \zeta) d\zeta d\xi. \tag{C2}$$

By using $F(t, \xi, \zeta) = F(t, \zeta, \xi)$ one has

$$\begin{aligned} \int_\tau^t \int_\tau^t F(t, \xi, \zeta) d\xi d\zeta &= 2 \int_\tau^t \int_\zeta^t F(t, \xi, \zeta) d\xi d\zeta \\ &= 2 \int_\tau^t \int_\tau^\xi F(t, \xi, \zeta) d\zeta d\xi. \end{aligned} \tag{C3}$$

The second equality follows from (C2), and we merely require to prove the first equality, from which we have

$$\int_\tau^t \int_\tau^\zeta F(t, \xi, \zeta) d\xi d\zeta = \int_\tau^t \int_\zeta^t F(t, \xi, \zeta) d\xi d\zeta.$$

Applying (C2) on the right-hand side we obtain

$$\int_\tau^t \int_\tau^\zeta F(t, \xi, \zeta) d\xi d\zeta = \int_\tau^t \int_\tau^\xi F(t, \xi, \zeta) d\zeta d\xi.$$

Interchanging the indices ξ and ζ in the left-hand side we get

$$\int_\tau^t \int_\tau^\xi F(t,\zeta,\xi) d\zeta d\xi = \int_\tau^t \int_\tau^\xi F(t,\xi,\zeta) d\zeta d\xi.$$

Because of $F(t,\zeta,\xi) = F(t,\xi,\zeta)$ the above equation is proved; hence, we have proved the first equality in (C3).

Now, we use the Leibnitz's rule in the following two equations:

$$\frac{d}{dt}\int_\tau^t \int_\tau^\eta \int_\tau^\eta F(\eta,\xi,\zeta) d\xi d\zeta d\eta = \int_\tau^t \int_\tau^t F(t,\xi,\zeta) d\xi d\zeta, \tag{C4}$$

$$\frac{d}{dt}\int_\tau^t \int_\zeta^t \int_\xi^t F(\eta,\xi,\zeta) d\eta d\xi d\zeta = \int_\tau^t \int_\zeta^t F(t,\xi,\zeta) d\xi d\zeta. \tag{C5}$$

By (C3) we have

$$\frac{d}{dt}\int_\tau^t \int_\tau^\eta \int_\tau^\eta F(\eta,\xi,\zeta) d\xi d\zeta d\eta = 2\frac{d}{dt}\int_\tau^t \int_\zeta^t \int_\xi^t F(\eta,\xi,\zeta) d\eta d\xi d\zeta, \tag{C6}$$

which is then integrated to obtain (C1). □