RM-76-59

# INVARIANT THEORY, THE RICCATI GROUP, AND LINEAR CONTROL PROBLEMS

J. Casti

July 1976

## Preface

One of the central themes of current research in mathematical system theory is the use of invariant theoretic-techniques to shed light on certain structural issues. The theory of invariants has been employed to characterize minimal-parameter canonical forms for constant, finite-dimensional linear systems. The objective of the present paper is to extend these basic results to the case in which a quadratic cost functional is adjoined. A complete, independent set of invariants is found for the so-called "linear-quadratic-gaussian" (LQG) problem.

# Invariant Theory, the Riccati Group, and Linear Control Problems


## Abstract

The classical algebraic theory of invariants is applied to the linear-quadratic-gaussian control problem to derive a canonical form under a certain matrix transformation group. The particular group of transformations, termed here the "Riccati group," is induced from the matrix Riccati equation characterizing the LQG problem solution.

Examples of the invariant-theoretic approach are given along with a discussion of topics meriting further study, including geometric interpretation of the group orbits, extension of the Riccati group, and connections with the generalized X-Y functions.


## 1. Introduction

One of the principal themes of nineteenth-century mathematics was the theory of invariants of transformation groups. This theory, developed to a high degree of sophistication by Hilbert, Cayley, Sylvester, Gordan, Dickson, and other, has been more or less relegated to the backwaters of mathematical activity until rather recently. A fascinating account of the rise and fall of the theory from the sociological point of view is given in [1,2].

In connection with renewed interest in invariant theory on the part of modern algebraists, there has been a corresponding interest in the use of invariant theoretic-techniques to shed light on certain algebraic issues arising in system theory. Since 1970, a number of papers have appeared [cf. 3-7] in which the general techniques developed by the nineteenth-century algebraists have been employed to characterize certain minimal-parameter canonical forms for constant, finite-dimensional linear systems. This work has been an outgrowth of one of the central themes of current research in mathematical system theory, namely, the use of modern algebra and differential geometry as tools to explore system-theoretic issues. This point of view, pioneered by Kalman, Hermann, Arbib, Brockett, and others, has developed from an offshoot of mathematical control and filtering theory to a legitimate branch of applied mathematics during the past decade. A good account of the current state of affairs may be found in [8].

Broadly speaking, past use of invariant theory in linear system problems has been confined to the following general set-up: a system $\Sigma$ is given by the dynamics

$$\dot{x} = Fx + Gu \quad ,$$

and the output

$$y(t) = Hx \quad ,$$

where x, u, and y are n, m, and p-dimensional vectors, respecti\ with F, G, H being constant matrices of appropriate sizes. Thus $\Sigma$ is characterized by the matrix triple (F,G,H). In addition, a certain group $\mathcal{G}$ of transformations is given which acts on the triple (F,G,H) to produce a new system $(\bar{F},\bar{G},\bar{H})$, i.e.,

$$\Sigma = (F,G,H) \overset{\mathcal{G}}{\rightarrow} (\bar{F},\bar{G},\bar{H}) = \bar{\Sigma} \quad .$$

In essence (details later), the theory of invariants is used to find a set of polynomials in the components of F, G, H which remain "invariant" under application of transformations from $\mathcal{G}$ to $\Sigma$. A principle objective is to use the transformations from $\mathcal{G}$ in such a way that $\bar{\Sigma}$ assumes a "simpler" form than $\Sigma$. Usually, "simpler" has been interpreted to mean that $\bar{\Sigma}$ contains a minimal number of parameters, consistent with the group $\mathcal{G}$.

In the classical terminology, a collection of polynomials $\{p_\alpha\}$, $\alpha \in A$ (A an index set), forming a set of invariants under $\mathcal{G}$ is called:

i) <u>independent</u> if no algebraic relations exist among the polynomials;

ii) <u>complete</u> if $p_\alpha(q) = p_\alpha(q')$ for all $\alpha \in A$ implies $q = gq'$ for some $g \in \mathcal{G}$.

The basic work cited above has been devoted to the determination of a complete, independent set of invariants and the associated canonical forms for the linear system $\Sigma$ using various choices of the group $\mathcal{G}$.

Our objective in this report is to extend the basic results obtained for the linear system $\Sigma$ to the case in which a quadratic cost functional

$$J = \int_t^T [(x,Qx) + 2(x,Su) + (u,Ru)]ds + (x(T),P_0x(T)) \quad ,$$

with $Q = Q'$, $R > 0$, $P_0 = P_0'$, is adjoined. Thus, we deal with the question of determining a complete, independent set of invariants for the so-called "linear-quadratic-gaussian" (LQG) problem. It will be shown that choosing $\mathcal{G}$ to be what we term the "Riccati group", such a set of invariants may be obtained for the LQG problem.

The general plan of the paper is to present, first of all, a brief review of the basic ideas of invariant theory with special emphasis on their use in linear system problems. We then give a formulation of the LQG problem which is particularly suitable for our purposes and define the transformations comprising the "Riccati" group. The next section contains the main results of the paper, namely a complete, independent set of invariants and the associated canonical form for the LQG problem. The paper concludes with a discussion of unresolved issues, as well as the connections of the current results with the generalized X-Y functions introduced in [9-11].

## 2. Invariant Theory and Linear Systems*

In general terms, invariant theory addresses the following situation: a fixed group $\mathcal{G}$ acts on certain mathematical "quantities" q. An <u>(absolute)</u> <u>algebraic</u> <u>invariant</u> of q with respect

---

* A more complete treatment of these matters is given in [22].

to $\mathcal{G}$ is a polynomial $p \in k[z_1,...,z_N]$ in the N "variables" constituting q such that $p(q) = p(gq)$ for all $g \in \mathcal{G}$ (here k represents an arbitrary, but fixed coefficient field). If p is an integer-valued function of q having the same property, then we call p an arithmetic invariant. If $p(gq) = (\det g)^{\omega}p(q)$, p is called a projective invariant of weight $\omega$, in the case where $\mathcal{G}$ is a matrix group (or a matrix representation of an abstract group).

A simple example of the foregoing set-up is the classical Grassmann variety of algebraic geometry which, following [12], can be expressed in the following manner. Let q correspond to rectangular matrices of size $n \times m$ over a field $k (m \geq n)$. Let $\mathcal{G}$ be the general linear group $GL(k^n)$, which acts on q by left multiplication, i.e., $q \to Tq$, $T \in \mathcal{G}$.

Let $\{q_i\}$ denote the columns of q and let $\det(q_{i_1},...,q_{i_n})$ be the determinant formed from exactly n (not necessarily distinct) columns of q. Consider the case $\det(q_1,...,q_n) \neq 0$. By left multiplication by the inverse of the matrix $(q_1,...,q_n)$, q becomes a matrix in which the block of first n columns is the identity matrix. Moreover, the remaining elements of q are then given (through Cramer's rule) by determinants of the type discussed above divided by $\det(q_1,...,q_n)$. These ratios are easily seen to be absolute invariants of $GL(k^n)$. Each such transformed matrix is in a canonical form and corresponds to a single orbit with respect to the Zariski neighborhood $\det(q_1,...,q_n) \neq 0$. Over other neighborhoods, say, $\det(q_2,...,q_{n+1}) \neq 0$, a similar construction of a canonical form can be carried out. Thus, the global structure is obtained by "piecing together" local canonical forms determined by different Zariski neighborhoods.

By considering the Plücker map $q \to (\det(q_1,...,q_n),..., \det(q_{m-n+1},...,q_m))$, one can look at the orbits as a point set in a projective space of high dimension. It can be shown that this point set forms a projective variety, the so-called Grassmann variety (see [13,14] for details). Thus, since there is a 1-1 correspondence between the points of this variety and the n-dimensional subspaces of m-dimensional space, we see that the ratios of determinants described above uniquely characterize the orbits generated by the group $GL(k^n)$.

The simplest system-theoretic example is that associated with a coordinate transformation in the state space. Let $\Sigma = (F,G,H)$ denote, as usual, the internal description of a constant, linear dynamical system. Take q as the triple of matrices $\Sigma = (F,G,H)$ and let $\mathcal{G} = GL(k^n)$ act on $\Sigma$ by a change of basis in the state space $X = k^n$, i.e.,

$$F \to TFT^{-1}$$

$$G \to TG \tag{1}$$

$$H \to HT^{-1}$$

for $T \in GL(k^n)$. The external properties of $\Sigma$ are preserved under this action since the infinite sequence of matrices

$$A_j = HF^{j-1}G \quad , \qquad j = 1,2,\ldots \quad ,$$

is fixed under $\mathcal{G}$. Consequently, every element of every matrix $A_j$ is an algebraic invariant of $(F,G,H)$ with respect to $GL(k^n)$.

It can be proved that the infinite sequence $\mathcal{S} = \{A_1,A_2,\ldots\}$ may be generated in a finitary way whenever $\mathcal{S}$ has a finite-dimensional realization; for instance, the matrices $A_1,\ldots,A_n$ and the numbers $\alpha_1,\ldots,\alpha_n$ (the coefficients of the characteristic polynomical of F in any minimal realization of $\mathcal{S}$) will generate all the elements of $\mathcal{S}$[cf.15]. Furthermore, it is clear that the $\alpha_1,\ldots,\alpha_n$ can be expressed as rational functions of $A_1,\ldots,A_{2n}$, and so it follows that all elements of $\mathcal{S}$ can be generated by algebraic operations from a finite number among them.

Since the (absolute) invariants form a ring, the above considerations lead to the result that

<u>The ring of invariants of a dynamical system $\Sigma$ with respect to the group $GL(k^n)$ can be viewed as being equivalent to the external description of $\Sigma$.</u>

The following foundational result, due to Hilbert, is now suggested.

<u>Hilbert Basis Theorem</u>. <u>The ring of invariants under the action of $GL(k^n)$ has a finite basis</u>.

(In fact, it is believed that this result is true for almost all groups, but the precise conditions are not yet known - this is Hilbert's $14^{th}$ problem.) In view of the above discussion, it follows that it should be possible to analyze almost all finite-dimensional systems by finitary methods, since the external description of such systems can be presumably identified with a particular ring of invariants.

In recent years, major steps forward in pursuing the above program have been made in [3,4,6] by enlarging the group of transformations $\mathcal{G}$ acting on $\Sigma$. In addition to the transformation (1), it is interesting to consider the additional transformations

$$G \to GV^{-1} \quad , \qquad (V \in GL(k^m)) \quad , \tag{2}$$

$$F \to F - GL \quad , \qquad (L = \text{arbitary } m\times n \text{ matrix over } k) \quad . \tag{3}$$

The three transformations (1)-(3) together generate a group acting on $\Sigma$ which is usually called the <u>feedback</u> group. Since this group is bigger than $GL(k^n)$, $\Sigma$ has fewer invariants. For example, the coefficients of the characteristic polynomial of F can be <u>arbitrarily</u> altered by the transformation (3). It is known that if (F,G) is controllable, then under the feedback group the pair (F,G) has only <u>arithmetic</u> invariants. These quantities have already been computed in 1890 by Kronecker during his investigations of the much more general problem of singular pencils of matrices [16]. In the mid-1960's, invariants under various subgroups of the feedback group were computed by Popov and are summarized in [4]. The invariant-theoretic point of view greatly simplifies the precise presentation of these results and was first given in [3]. Other notable work in this direction, particularly as it relates to the so-called "decoupling" problem of linear systems was given in the early 1970's by Wonham, Morse, and their colleagues [6].

A major problem remaining in this area is the determination of a complete set of invariants of (F,G,H) under the feedback group. It is easy to see that the full feedback group is not needed to obtain the Kronecker canonical form for (F,G); in fact, since (F,G) has $n^2 + mn$ parameters and the feedback group has $n^2 + nm + m^2$ parameters, exactly $m^2$ parameters are "left over" by means of which the feedback group can act on H. This action is rather complicated, however, and it is not at all clear how to obtain invariants and canonical forms in this case. Some fragmentary results are reported, though, in [17].

## 3. The Linear-Quadratic Gaussian Problem and the Riccati Group

As noted in the introductory section, we shall be concerned with the problem of minimizing (over u) the quadratic criterion

$$J = \int_t^T [(x,Qx) + 2(x,Su) + (u,Ru)]ds + (x(T),P_0x(T)) \quad , \qquad (4)$$

where $Q = Q'$, $P_0 = P_0'$, $R > 0$, and x and u are n, m-dimensional vectors connected by the linear equations

$$\dot{x} = Fx + Gu \quad , \qquad x(t) = c \quad . \qquad (5)$$

All matrices are assumed to be real, constant matrices of appropriate sizes.

Well known [15] results show that the minimizing control $u^*(t)$ is given in feedback form as

$$u^*(t) = -K(t)x(t) \quad ,$$

where $K(t) = R^{-1}(G'P(t)+S')$ with $P(t)$ being the solution of the matrix Riccati equation

$$-\dot{P} = Q + PF + F'P - (PG+S)R^{-1}(PG+S)' \quad , \qquad t < T \quad , \qquad (6)$$

$$P(T) = P_0 \quad .$$

In addition, the minimal value of the quadratic criterion is expressed through P as

$$J_{min}(t) = (c,P(t)c) \quad .$$

Thus, from an algebraic point of view, we may regard the LQG problem as being specified by the set of matrices $\mathcal{Q} = (F,G,Q,R,S,P_0)$. Our goal will be to let a certain naturally induced group act on $\mathcal{Q}$ and to calculate a complete set of independent, algebraic invariants of $\mathcal{Q}$ under this group. Simultaneously, we shall also obtain a canonical form for the LQG problem.

Before passing to the definition of the group , we make a short digression into the underlying theory of the LQG problem as the results will be essential for our subsequent development.

If one assumes at the outset that the optimal control law $u^*(t)$ has the feedback form $u^* = -K(t)x$, with $K(t)$ an unknown matrix function to be determined, then substitution of $u^*$ into the criterion, together with addition to the criterion function of the identity

$$(\dot{x}-Fx+GKx, P(t)(\dot{x}-Fx+GKx)) = 0 \quad ,$$

valid for all $P(t) \geq 0$, transforms the original problem to simpler form. The modified problem, following an integration by parts, is to minimize

$$\mathrm{tr}\left\{[I \ \ -K]\begin{bmatrix} F'P + PF + Q + \dot{P} & PG + S \\ G'P + S' & R \end{bmatrix}\begin{bmatrix} I \\ -K \end{bmatrix} W\right\} \quad ,$$

where $W \geq 0$ is arbitrary. The minimization is over all matrix functions $K(t)$, $P(t)$. Since $W \geq 0$ is arbitrary, it is not difficult to see that the minimizing P and K satisfy the equations given above. This approach is detailed in [28,29].

What is of interest in this approach as far as the aims of this paper are concerned, is the matrix

$$\mathscr{M}(P) = \begin{bmatrix} Q + PF + F'P & PG + S \\ G'P + S' & R \end{bmatrix} ,$$

which will play an essential role in the development of our canonical form for the LQG problem.

Returning to the algebraic problem, it is clear that several transformations can act on $\mathscr{Q}$ in such a way that the effect of a transformation on the <u>solution</u> of the problem can be directly computed. Such transformations can be viewed as "simplifying" the probelm of solving (6).

We shall work with the following set of transformations of $\mathscr{Q}$, conviently termed the <u>Riccati group</u> $\mathscr{R}$:

(I) change of basis T in the state space,

(II) change of basis V in the control space,

(III) application of an arbitrary feedback law L,

(IV) addition of a constant matrix M to P.

Thus, the Riccati group $\mathscr{R}$ equals the feedback group plus Type IV transformations.

The action of $\mathscr{R}$ on the problem data is easily calculated from the Riccati equation (6) (hence, the nomenclature for $\mathscr{R}$). Explicitly, we have

$$(F,G,Q,R,S,P_0) \xrightarrow{(I)} (TFT^{-1},TG,T'^{-1}QT^{-1},R,T'^{-1}S,T'^{-1}P_0T^{-1}) ,$$

$$\det T \neq 0 ,$$

$$(F,G,Q,R,S,P_0) \xrightarrow{(II)} (F,GV^{-1},Q,V'^{-1}RV^{-1},SV^{-1},P_0) ,$$

$$\det V \neq 0 ,$$

$$(F,G,Q,R,S,P_0) \xrightarrow{(III)} (F-GL,G,Q+L'RL-L'S'-SL,R,S-L'R,P_0) \quad ,$$

L arbitrary ,

$$(F,G,Q,R,S,P_0) \xrightarrow{(IV)} (F,G,Q-F'M-MF,R,S-MG,P_0+M) \quad ,$$

$M = M'$ .

There are a few scattered papers in the literature concerning the choice of the simplest canonical form for $\mathcal{Q}$ under a group generated by one or more of the transformations (I)-(IV) just described. For example, in [18] it is shown that in the single-input case (G = single column matrix) exactly n parameters are needed to describe the problem and these parameters may be chosen to be the diagonal entries of the matrix Q. In other words, all single-input quadratic optimization problems can be reduced to the solution of an n-parameter family of Riccati equations plus the transformations (I)-(IV). More extensive results have been obtained recently in [19].

## 4. Invariants and LQG Canonical Forms

The triangular factorization of the matrix $\mathcal{M}(P)$ as

$$\mathcal{M}(P) = \begin{bmatrix} \phi & \Gamma \\ 0 & \psi \end{bmatrix} \begin{bmatrix} \phi' & 0 \\ \Gamma' & \psi' \end{bmatrix} \quad ,$$

immediately leads to the conclusion that the optimal feedback gain function K(t) is related to the triangular factors as

$$K(t) = \psi^{-1}(t)\Gamma'(t) \quad .$$

In addition, it is easily verified that $-\dot{P} = \phi\phi'$. Hence, solution of the LQG problem would be greatly simplified by using the Riccati group to transform $\mathcal{M}(P)$ to a form which facilitates the triangular decomposition of $\mathcal{M}(P)$. Since there are also deep connections between the triangular factors and the "generalized X-Y" functions [20], such an approach will also shed new light on these important quantities. Thus, our plan is clear: we shall utilize the Riccati group to reduce $\mathscr{M}(P)$ to a simpler form, at the same time obtaining an LQG-canonical form for the matrices comprising $\mathscr{Q}$.

We begin with the problem

$$\mathscr{Q} = (F, G, Q, R, S, P_0) \quad ,$$

and assume that $P_0$ is not a solution of the algebraic Riccati equation (this assumption is only for convenience and its removal will be discussed later). In view of our desire to triangularly factor $\mathscr{M}(P)$, we shall use the transformation group $\mathscr{R}$ to bring $\mathscr{M}$ to a form as near diagonal as possible.

Application of Type (I) and (II) transformations gives

$$\mathscr{M}_{I-II}(P_0) = \left[\begin{array}{c|c} T'^{-1}(Q+P_0F+F'P_0)T^{-1} & T'^{-1}(P_0G+S)V^{-1} \\ \hline V'^{-1}(G'P_0+S')T^{-1} & V'^{-1}RV^{-1} \end{array}\right] \quad . \tag{7}$$

Further, use of Type (III) and (IV) transformations leads to

$$\mathscr{M}_{I-IV}(P_0) = \mathscr{M}_{I-II}(P_0) + \left[\begin{array}{c|c} \begin{array}{c} L'R_{I-II}L - L'(P_0G+S)'_{I-II} - (P_0G+S)_{I-II}L \\ - MF_{I-III} - F'_{I-III}M \end{array} & -L'R_{I-II} - MG_{I-III} \\ \hline -R_{I-II}L - G'_{I-III}M & 0 \end{array}\right] , \tag{8}$$

where the Roman numeral subscripts denote the form obtained after application of the indicated sequence of transformations.

In light of (7)-(8), we choose the Type (I) transformation T such that

$$T'^{-1}QT^{-1} = \text{diag}_p(\pm 1) \quad ,$$

where the subscript p indicates that the first p positions equal

±1, the remainder being zero, p = rank Q. Such a choice of T is always possible since Q is symmetric [21].

Next, we select the Type (II) transformation V so that

$$V'^{-1}RV^{-1} = I \quad .$$

As with T, such a choice is possible since, by the original assumption, R > 0. As a result of these choices we have

$$\mathcal{M}_{I-II}(P_0) = \left[\begin{array}{c|c} \text{diag}_p(\pm 1) + T'^{-1}(P_0F+F'P_0)T^{-1} & T'^{-1}(P_0G+S)V^{-1} \\ \hline V'^{-1}(G'P_0+S')T^{-1} & I_m \end{array}\right] ,$$

where $I_m$ is the m×m identity matrix.

The feedback transformation law L is now selected in order to "zero out" the off-diagonal blocks. This involves selecting L such that

$$R_{I-II}L = V'^{-1}(G'P_0+S')_{I-II}T^{-1} \quad ,$$

or, since $R_{I-II} = I_m$,

$$L = V'^{-1}(G'P_0+S')_{I-II}T^{-1} \quad .$$

Finally, we come to the choice of the Type (IV) transformation M. Recalling (8) and the choices of T, V, and L , for arbitrary M = M' we find that

$$\mathcal{M}_{I-IV}(P_0) = \left[\begin{array}{c|c} \begin{array}{c}\text{diag}_p(\pm 1) + T'^{-1}(P_0F+F'P_0-L'L)T^{-1} \\ -MF_{I-III} - F'_{I-III}M\end{array} & -MG_{I-III} \\ \hline -G'_{I-III}M & I_m \end{array}\right] .$$

Thus, we choose M such that

(i) $MG_{I-III} = 0$ (this implies rank $M \leq n - m$),

(ii) $-MF_{I-III} - F'_{I-III}M + T'^{-1}(P_0F+F'P_0-L'L)T^{-1} =$

$$-\left[\begin{array}{c|c} \mathrm{diag}_{n-m}(\pm 1) & X_1 \\ \hline X_1' & X_2 \end{array}\right],$$

where the entries in $X_1$ and $X_2$ are, in general, non-zero elements determined by M. Such a choice of M is clearly possible since only the last m columns say, and m rows are fixed by condition (i) and $P_0$ is not a solution of the algebraic Riccati equation.

Thus, the final form for $\mathscr{M}$ is

$$\mathscr{M}_{I-IV}(P_0) = \left[\begin{array}{c|c} \mathrm{diag}_p(\pm 1) - \left[\begin{array}{c|c} \mathrm{diag}_{n-m}(\pm 1) & X_1 \\ \hline X_1' & X_2 \end{array}\right] & 0 \\ \hline 0 & I_m \end{array}\right],$$

where the first diagonal matrix has p non-zero entries, the second has n-m non-zero entries.

The invariants of $\mathscr{Q}$ under the action of the Riccati group $\mathscr{R}$ can now be seen to be

a) the entries of the matrices $X_1$ and $X_2$,

and

b) the entries of the matrix $(P_0)_{I-IV}$, i.e., the numbers $T'^{-1}P_0T^{-1} + M$.

This is in complete agreement with a count of the number of "degrees of freedom" in $\mathscr{Q}$ versus that in $\mathscr{R}$, i.e., the entries in $\mathscr{Q}$ represent $N_1 = n^2 + nm + n(n+1)/2 + m(m+1)/2 + mn + n(n+1)/2$ independent quantities while in $\mathscr{R}$ there are $N_2 = n^2 + m^2 + mn + n(n+1)/2$ numbers to be chosen in the transformations (I)-(IV). Thus, $N_1 - N_2 = m(m+1)/2 + mn - m^2 + n(n+1)/2$, precisely the number of entries in the matrices $X_1$, $X_2$, and $(P_0)_{I-IV}$. Thus, the foregoing choices of T, V, L, and M have removed as many degrees of freedom as possible from the problem $\mathscr{Q}$.

Since the transformations (I)-(IV) above are determined only by the original problem data $\mathscr{Q}$, it is manifestly clear that the elements $X_1$, $X_2$, and $(P_0)_{I-IV}$ form a complete set of invariants for $\mathscr{Q}$ under $\mathscr{R}$, i.e., any two problems $\mathscr{Q}_1$ and $\mathscr{Q}_2$ leading to the same set of invariants lie on the same $\mathscr{R}$ orbit. It only remains to verify that this set of invariants is independent, i.e., that some $\mathscr{Q}$ can be reconstructed given the elements of $X_1$, $X_2$, $(P_0)_{I-IV}$ and a set of transformations from $\mathscr{R}$.

Determination of $\mathscr{Q}$ is carried out in the following steps:

1. Choose arbitrary, but fixed elements T, V, L such that $\det T \neq 0$, $\det V \neq 0$. Also, select $M = M'$ such that rank $M \leq n - m$.

2. Determine $P_0$ from the invariants $(P_0)_{I-IV}$ and the transformations M and T; explicitly,

$$P_0 = T'[(P_0)_{I-IV} - M]T \quad .$$

3. Determine Q from $\text{diag}_p(\pm 1)$ and T, i.e.,

$$Q = T'\text{diag}_p(\pm 1)T \quad .$$

4. Using the invariants $X_1$ and $X_2$, determine F from M, T, L, $P_0$, and $\text{diag}_p(\pm 1)$ by solving the equation

$$T'^{-1}(P_0F+F'P_0-L'L)T^{-1} - MTFT^{-1} - T'^{-1}F'T'M =$$

$$-\left[\begin{array}{cc|c} \text{diag}_{n-m}(\pm 1) & X_1 & \\ X_1' & X_2 & 0 \\ \hline 0 & & I_m \end{array}\right] \quad .$$

Here we have made use of the requirement that $MG_{I-III} = 0$ in order to remove G from the coefficient of M. The above equation is uniquely solved as a linear system for the components of F.

5. Determine the precise entries in $G_{I-III}$ by solving the system

$$MG_{I-III} = 0 \quad ,$$

$$G_{I-III}L = 0 \quad .$$

(Recall that step 4 was carried out knowing only that the columns of $G_{I-III}$ are contained in the kernel of M. The second equation above uniquely determines these columns.) Knowing $G_{I-III}$, we calculate G as

$$G = T^{-1}G_{I-III}V \quad .$$

6. Determine R as

$$R = V'V \quad .$$

7. Determine S from the relation

$$R_{I-II}L = V'^{-1}(G'P_0+S')_{I-II}T^{-1} \quad ,$$

which, after a minor calculation, yields

$$S' = V'^2LT^2 - G'P_0 \quad .$$

The final point to consider in connection with the invariants given above is the case in which $P_0$ <u>is</u> a solution of the algebraic Riccati equation

$$Q + P_0F + F'P_0 - (P_0G+S)R^{-1}(P_0G+S)' = 0 \quad . \tag{9}$$

Upon examination of the steps leading from the original problem $\mathcal{Q}$ to the canonical form of $\mathcal{M}$, it is easily verified that the choice of L indicated above leads to the result

$$\mathrm{diag}_p(\pm 1) + T'^{-1}(P_0F+F'P_0-L'L)T^{-1} = 0 \quad .$$

Thus, if $P_0$ solves the algebraic Riccati equation (9),

$$\mathcal{M}_{I-IV} = \left[\begin{array}{c|c} -MF_{I-III} - F'_{I-III}M & -MG_{I-III} \\ \hline -G'_{I-III}M & I_m \end{array}\right] \quad .$$

If the pair (F,G) is controllable, then we choose $M \equiv 0$ to zero out the upper left block (since controllability plus the particular choice of L implies $F_{I-III}$ is a stability matrix, i.e., the equation $MF_{I-III} + F'_{I-III}M = 0$ has no nontrivial solution M). Finally, we see that this leads to

$$\mathcal{M}_{I-IV} = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & I_m \end{array}\right] \quad ,$$

i.e., $P_0$ being a solution of the algebraic Riccati equation implies that the invariants $X_1$ and $X_2$ vanish identically. The converse, namely that $X_1 = 0$, $X_2 = 0$, $M = 0$ implies $P_0$ is a solution of the algebraic Riccati equation, is also easily established.

In summary, we have established the following basic result.

Algebraic Invariant Theorem for the LQG Problem. Given the LQG problem data $\mathcal{D} = (F,G,Q,R,S,P_0)$, with $P_0$ not a solution of the algebraic Riccati equation (9), a complete, independent set of algebraic invariants of $\mathcal{D}$ under the Riccati group $\mathcal{R}$ is given by the $m(m+1)/2 + mn - m^2 + n(n+1)/2$ elements of the three matrices $X_1$, $X_2$, and $(P_0)_{I-IV}$, where $X_1$ and $X_2$ appear in the canonical matrix

$$\mathcal{M} = \left[\begin{array}{c|c} \mathrm{diag}_p(\pm 1) - \left[\begin{array}{c|c} \mathrm{diag}_{n-m}(\pm 1) & X_1 \\ \hline X_1' & X_2 \end{array}\right] & 0 \\ \hline 0 & I_m \end{array}\right] ,$$

<u>and $(P_0)_{I-IV} = T'^{-1} P_0 T^{-1} + M$.</u>

<u>If $P_0$ is a solution of the algebraic Riccati equation (9), then the invariants $X_1$ and $X_2$ vanish identically and the only non-fixed invariants are the elements of $(P_0)_{I-IV}$.</u>

## 5. An Example: A Lur'e-Lefschetz-Letov System

In his book on the Lur'e problem [23], Letov made extensive use of the properties of the following completely controllable system

$$F = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \quad , \qquad G = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad ,$$

where the $\lambda_i$ are all non-zero and distinct. We shall compute the system invariants with a quadratic cost functional having

$$Q = I \quad , \qquad R = I \quad , \qquad S = 0 \quad , \qquad P_0 = 0 \quad ,$$

the system dynamics being in the Lur'e-Lefschetz-Letov form given above.

Following the steps given in Section 4, we first note that the special forms of Q and R imply that the type (I) and (II) transformations T and V may be chosen as arbitrary orthogonal matrices. In our case, we choose them so as to simplify the calculations. Thus, for the sake of definiteness, we take

$$T = I \quad , \qquad V = I \quad .$$

In view of the fact that there is no terminal cost term ($P_0 = 0$) and no cross-coupling in the problem ($S=0$), the type III transformation L is

$$L = (P_0G)_{I-II} = 0 \quad .$$

Following selection of the type (IV) transformation M, we will have

$$\mathcal{M}_{I-IV} = \left[\begin{array}{c|c} I_n - MF_{I-III} - F'_{I-III}M & -MG_{I-III} \\ \hline -G'_{I-III}M & I_m \end{array}\right] ,$$

where, since $L = 0$, $T = I$, $V = I$,

$$F_{I-III} = F = \text{diag}(\lambda_1, \ldots, \lambda_n) \quad ,$$

$$G_{I-III} = G = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} .$$

We now show that it is possible to select M such that

$$MG_{I-III} = 0 \quad ,$$

$$I_n - MF_{I-III} - F'_{I-III}M = \left[\begin{array}{c|c} 0 & -X_1 \\ \hline -X'_1 & 1 - X_2 \end{array}\right] ,$$

where the vector $X_1$ and the scalar $X_2$ will be the non-zero invariants of the problem.

Upon partitioning M as

$$M = \left[\begin{array}{c|c} M_{11} & M_{12} \\ \hline M'_{12} & M_{22} \end{array}\right] ,$$

where $M_{11}$ is the size $(n-1)\times(n-1)$, while $M_{12}$ is $(n-1)\times 1$, and $M_{22}$ = scalar, we have

$$MG = \begin{bmatrix} M_{11}G_1 + M_{12}G_2 \\ M'_{12}G_1 + M_{22}G_2 \end{bmatrix} = 0 , \qquad (10)$$

where the size of $G_1$ is $(n-1)\times 1$, and $G_2$ is the scalar 1. Thus, the vector $M_{12}$ and the scalar $M_{22}$ will be fixed by the orthogonality condition (10), leaving $M_{11}$ free. Letting

$$\Lambda_{n-1} = \text{diag}(\lambda_1,\ldots,\lambda_{n-1}) , \qquad \Lambda_n = \lambda_n ,$$

we desire to choose $M_{11}$ such that

$$M_{11}\Lambda_{n-1} + \Lambda_{n-1}M_{11} = I_{n-1}$$

It is easily checked that

$$M_{11} = \text{diag}\left(\frac{1}{2\lambda_1},\frac{1}{2\lambda_2},\ldots,\frac{1}{2\lambda_{n-1}}\right) .$$

The orthogonality condition (10) then yields immediately

$$M_{12} = -\frac{1}{2}\begin{pmatrix} \frac{1}{\lambda_1} \\ \frac{1}{\lambda_2} \\ \vdots \\ \frac{1}{\lambda_{n-1}} \end{pmatrix} , \qquad M_{22} = \frac{1}{2}\sum_{i=1}^{n-1}\frac{1}{\lambda_i} .$$

Finally, from the representation

$$I_n - MF_{I-III} - F'_{I-III}M = \left[\begin{array}{c|c} 0 & -X_1 \\ \hline -X_1' & 1 - X_2 \end{array}\right] ,$$

we obtain the invariants as

$$X_1 = M_{12}\Lambda_n + \Lambda_{n-1}M_{12} ,$$

$$-1 + X_2 = M_{22}\Lambda_n + \Lambda_n M_{22} .$$

Hence,

$$X_1 = -\frac{1}{2}\begin{bmatrix} \frac{\lambda_n}{\lambda_1} + 1 \\ \frac{\lambda_n}{\lambda_2} + 1 \\ \vdots \\ \frac{\lambda_n}{\lambda_{n-1}} + 1 \end{bmatrix} ,$$

$$X_2 = \lambda_n \sum_{i=1}^{n-1} \frac{1}{\lambda_i} + 1 .$$

As a result, we see that the algebraic invariants of the problem are the entries of $X_1$ and $X_2$, together with the entries of

$$(P_0)_{I-IV} = T'^{-1}P_0T^{-1} + M$$

$$= M .$$

The final canonical form for $\mathscr{M}$ is

$$\mathcal{M}_{I-IV} = \left[\begin{array}{c|c} \left[\begin{array}{c|c} 0 & -X_1 \\ \hline -X_1' & 1-X_2 \end{array}\right] & 0 \\ \hline 0 & 1 \end{array}\right] .$$

Knowing T, V, L, and M, it is also now a simple matter to work out the induced canonical forms for F, G, Q, R, S, and $P_0$.

6. Discussion

Several points surrounding the foregoing results suggest themselves for further investigations which, regrettably, require a development beyond the confines of this paper. We briefly touch on a selection of topics which seem particularly pertinent.

1) Connections with Generalized X-Y Functions

The matrix $\mathcal{M}(P)$ has been studied elsewhere [20] in connection with its role in the development of "low-dimensional" algorithms for computing the optimal feedback gain function for the LQG problem. In particular, when the system data $\mathcal{D}$ consists of constant matrices, it has been shown [20] that by factoring $\mathcal{M}$ into its symmetric triangular factors, i.e.,

$$\mathcal{M}(P) = \begin{bmatrix} \phi & \Gamma \\ 0 & \psi \end{bmatrix} \begin{bmatrix} \phi' & 0 \\ \Gamma' & \psi' \end{bmatrix} ,$$

it is possible to identify (modulo an orthogonal transformation) the entries $\phi$ and $\Gamma$ with the generalized X and Y functions discussed in [9-11]. In fact, it was this identification which, by a rather tortuous route, led to the study of $\mathcal{M}$ in connection with the problem of invariants.

It would be of interest to pursue this matter further, at least to the extent of explicitly calculating the associated "canonical" X-Y system with an eye toward further simplification of the LQG problem.

2) <u>Geometric Structure of $\mathscr{Q}$</u>

As pointed out in [22], it is somewhat naive to confine invariant-theoretic studies of system problems solely to obtaining the invariants and the associated canonical form. Such an approach details only the algebraic side of the picture; however, there is also a geometric side which must be pinned down. A complete understanding of the geometric structure of the orbits of $\mathscr{Q}$ under $\mathscr{R}$ will be essential to any hope of obtaining a <u>global</u> theory of the LQG problem.

The above point is already made clear by analogy with the far simpler problem of classifying the orbits of the set of $n\times nm$ matrices under the group consisting of left multiplication by a nonsingular matrix. We have seen that, in this case, the orbits have the geometric structure of a quasiprojective variety, the so-called Grassmann variety. A similar type of geometric structure must be established for the orbits of $\mathscr{Q}$ under $\mathscr{R}$.

3) <u>Enlargement of $\mathscr{R}$</u>

The canonical form presented above for the LQG problem was derived by using up all allowable actions of the Riccati group $\mathscr{R}$. The only way to obtain a "simpler" form (i.e., one with fewer parameters) is to enlarge the transformation group to include more operations. Unfortunately, it is not at all clear how to meaningfully carry out this prescription. Presumably, whatever additional transformations are employed will have physical significance in the context of the LQG set-up and, consequently, must leave the Riccati equation invariant. This is clearly a problem meriting additional study.

4) <u>Applications to Other Areas</u>

The theory presented here is, in actuality, not so much an algebraic theory of control processes but rather it is an algebraic theory of the matrix Riccati equation. Clearly, given any matrix Riccati equation

$$\dot{R} = A + BR + RC + RDR \quad , \qquad R(a) = F \quad ,$$

with A = A', B = C', D = D', F = F', we can form a corresponding LQG problem.

However, there are many areas of mathematical physics and engineering in which a matrix Riccati equation without the above symmetry properties plays the central role. We note the theories of neutron transport and atmospheric radiative transfer as particular examples in this regard [23-27]. In addition, numerous mathematical questions regarding two-point boundary value problems and Fredholm integral equations involve a similar Riccati equation as a vital part of the problem formulation.

In summary, the algebraic results presented here give the possibility for classifying a wide variety of problems in mathematics and physics and for studying their inherent algebraic and geometric structure.

Acknowledgement

The author wishes to express his gratitude to Professor R.E. Kalman for many hours of fruitful discussion regarding the basic issues presented here. In particular, access to the unpublished manuscript [22] has been an essential ingredient in our presentation, especially the material covered in Section 2.

## References


1. Fisher, C.S. "The Death of a Mathematical Theory," Arch. Hist. of Exact Sci. 3(1966), 137-159.

2. Fisher, C.S. "The Last Invariant Theorists," Arch. of European Sociology 8(1967), 216-244.

3. Kalman, R. "Kronecker Invariants and Feedback," in Proc. Conf. on Ord. Diff. Eqs., L. Weiss, ed., Academic Press, New York, 1972.

4. Popov, V. "Invariant Description of Linear, Time-Invariant Controllable Systems," SIAM J. Control 10(1972), 252-264.

5. Weinert, H. "Complete Sets of Invariants for Multivariable Linear Systems," 7th Princeton Conf. on Info. Th. & Systems, Princeton, 1973.

6. Wonham, W., and A. Morse, "Feedback Invariants of Linear Multivariable Systems," Proc. IFAC Symp. on Multivariable Systems, Düsseldorf, October 1971.

7. Rosenbrock, H. State-Space and Multivariable Theory, Wiley-Interscience, New York, 1970.

8. Proc. NATO Institute on Algebraic and Geometric Methods for Nonlinear Systems, R. Brockett and D. Mayne, eds., London, 1973 (published by D. Reidel, Co., 1975).

9. Casti, J. "Matrix Riccati Equations, Dimensionality Reduction, and Generalized X-Y Functions," Utilitas Math. 6(1974), 95-110.

10. Casti, J. "Some Recent Developments in the Theory and Computation of Linear Control Problems," Proc. MRC Conf. on Calculus of Variations and Control Theory, Madison, Wisconsin, 1975 (to be published by Academic Press, New York, 1976).

11. Kailath, T. "Some Chandrasekhar-Type Algorithms for Quadratic Regulators," Proc. IEEE Dec. & Control Conf., New Orleans, 1972.

12. Kalman, R. "Global Structure of the Class of Linear Dynamical Systems of Fixed Dimension," Proc. 8th Princeton Conf. on Info. Th. and Systems, Princeton, 1974.

13. Hodge, W., and D. Pedoe, Methods of Algebraic Geometry, vol. 1, Cambridge University Press, Cambridge, 1947.

14. Dieudonne, J., and J. Carrell, Invariant Theory: Old and New, Academic Press, New York, 1971.

15. Kalman, R., P. Falb, and M. Arbib, Topics in Mathematical Systems Theory, McGraw-Hill, New York, 1969.

16. Gantmacher, F. Theory of Matrices, Chelsea Publishing Co., New York, 1959.

17. Wang, S., and E. Davison, "Canonical Forms of Linear Multivariable Systems," SIAM J. Control & Opt. 14(1976), 236-250.

18. Kreindler, E., and J. Hedrick, "On the Equivalence of Quadratic Loss Functions," Int'l J. Control. 11(1970), 213-222.

19. Bernhard, P., and G. Cohen, "Etude d'une fonction frequéntielle intervenant dans un problème de commande optimale avec une application à la rèduction de la taille de ce problème," Revue RAIRO (1973), 63-85.

20. Casti, J. "Generalized X-Y Functions, the Linear Matrix Inequality, and Triangular Factorization for Linear Control Problems," IIASA RM-76-10, IIASA, February 1976 (submitted to SIAM J. Control).

21. Marcus, M, and H. Minc, A Survey of Matrix Theory and Matrix Inequalities, Prindle, Weber, and Schmidt Publishing Co., Boston, 1964.

22. Kalman, R. "System-Theoretic Aspects of Invariant Theory," November 1974 (unpublished manuscript).

23. Letov, A. Stability in Nonlinear Control Systems, Princeton University Press, Princeton, 1961.

24. Casti, J. "Optimal Filtering and Neutron Transport: Isomorphic Theories?" IIASA RM-75-25, IIASA, June 1975 (submitted to SIAM Review).

25. Chandrasekhar, S. Radiative Transfer, Dover Publishing Co., New York, 1960.

26. Bellman, R., R. Kalaba, and M. Prestrud, Invariant Imbedding and Radiative Transfer in Slabs of Finite Thickness, American Elsevier Co., New York, 1963.

27. Casti, J., and R. Kalaba, Imbedding Methods in Applied Mathematics, Addison-Wesley Publishing Co., Reading, Mass., 1973.

28. Brockett, R. Finite-Dimensional Linear Systems, John Wiley and Sons, New York, 1970.

29. Willems, J. "Least-Squares Stationary Optimal Control and the Algebraic Riccati Equation," IEEE Tran. Auto. Cont. AC-16 (1971), 621-634.