

REMARKS ON THE POLE-SHIFTING PROBLEM OVER RINGS

R. BUMBY, E.D. SONTAG*, H.J. SUSSMANN**
and W. VASCONCELOS**

Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, USA

Communicated by H. Bass
Revised January 1980

Given a square n -matrix F and an n -row matrix G , pole-shifting problems consist in obtaining more or less arbitrary characteristic polynomials for $F + GK$, for suitable ("feedback") matrices K . A review of known facts is given, various partial results are proved, and the case $n=2$ is studied in some detail.

Introduction

Problems that appear in trying to extend linear control results to systems over rings R have attracted considerable attention lately. This interest has been due mainly to applications-oriented motivations (in particular, dealing with delay-differential equations), and partly to a purely algebraic interest. We shall not touch here on the (nonalgebraic) motivations—many can be found in the various references given—save to note that interest in applications lies not with arbitrary rings R but with certain broad classes, such as polynomial rings over \mathbf{R} or \mathbf{C} (delay systems), integers and finite rings (digital systems, coding), rings of suitably smooth real or complex functions (parametrized families of systems), and group algebras with real or complex coefficients (discretized p.d.e.'s on certain manifolds).

In this note, we shall restrict our attention to the problem(s) of modifying the characteristic polynomial of a given system through the use of feedback. A *system* (with m inputs, of dimension n , over the commutative ring R) is just a pair of matrices (F, G) over R , where F is n by n , and G is n by m . A *feedback (matrix)* for this system is any m by n R -matrix K . The *closed-loop* system obtained applying feedback K to the system (F, G) is by definition the new system $(F + GK, G)$. We shall be interested in the *characteristic polynomial* of the system (F, G) , meaning just the characteristic polynomial of F , $\text{ch.p.}(F) = \det(zI - F)$.

* Author supported in part by U.S. Air Force Grant AFOSR F-49620-79-C-0117.

** Authors supported in part by N.S.F. Grants.

The above terminology originates in the study of vector equations

$$(\alpha x)(t) = Fx(t) + Gu(t), \quad (1.1)$$

(where $x(\cdot)$ is the “state variable” and $u(\cdot)$ is the “input” or “control” function, and where α is either a difference or a differentiation operator), under a state-feedback control law

$$u(t) = Kx(t) + v(t), \quad (1.2)$$

where v is a new input, so that the composite (“closed-loop”) system becomes

$$(\alpha x)(t) = (F + GK)x(t) + Gv(t). \quad (1.3)$$

An excellent discussion of modern feedback control topics is given in [26], when $R =$ real or complex numbers. (Extensions to rings appear for instance when F, G are matrices of operators, or when states and inputs are restricted in various ways.) The interest in ch.p. $(F + GK)$ is due to the fact that stability and other dynamic properties of (1.1) depend on ch.p. (F) , and the feedback K is used to change these properties for various control purposes. No explicit use will be made here of the interpretation (1.1)–(1.3). A system will be for us just the above-defined algebraic object.

For systems over a field R , the main “pole-shifting” result is:

Theorem ($R =$ field). *Assume that (F, G) is reachable, i.e. that the columns of G, FG, F^2G, \dots generate R^n . (By Cayley–Hamilton, $(G, FG, \dots, F^{n-1}G)$ is enough). Then,*

(1.4) *For each monic polynomial $p(z)$ in $R[z]$ of degree n , there is some feedback K such that ch.p. $(F + GK) = p(z)$.*

Conversely, if (1.4) is true for a given (F, G) , then this system is reachable.

The proof of the above theorem evolved over many years, (including many generalizations dealing with the possible set of invariant factors of $F + GK$). A short and elegant proof was given by Heymann [8] and we review it below.

Consider the property given by (1.4). We shall call this the *coefficient-assignment* property. It is easy to prove (see below) that reachability is still necessary for (1.4) to hold, over any commutative ring R . In fact, when the number of columns of G (“inputs”) is 1, or (obviously) if the dimension n is 1, the reachability condition is also sufficient. A CA_n -ring will be one for which, (as over fields), reachability is equivalent to (1.4) for every system of dimension $\leq n$; a CA -ring is a CA_n ring for all $n \geq 1$.

It is not hard to prove that semilocal rings are CA -rings, but the problem of deciding if there are any (“nice”) non- CA -rings was open. We shall show below that the rational integers and the polynomial rings over \mathbf{R} are not CA -rings, but that polynomials in one variable over \mathbf{C} are (at least) CA_2 . We leave open the question of

existence of an “ n -stable range”, i.e. whether there is any s with $CA_s = CA$.

For many applications, the following weaker property, which we shall call *pole-assignment*, is enough:

(1.5) For each $\lambda_1, \dots, \lambda_n$ in R , there is some feedback K with $\text{ch.p.}(F + GK) = (z - \lambda_1) \cdots (z - \lambda_n)$.

(The standard terminology “poles” is motivated by the fact that the eigenvalues of F in (1.1) give rise to the poles of the “transfer function” of the system.) It is again true that (1.5) implies reachability of (F, G) . A PA_n -ring will be one for which reachability implies (1.5), for each system of dimension at most n ; a PA -ring if for all n . It is known that principal-ideal domains are PA rings, and an extension of the argument for PID’s will show that elementary-divisor-rings (and some others) are also PA -rings. This will apply in particular to rings of real-analytic functions and others of applied interest. We shall see, however, that polynomial rings over \mathbf{R} in more than one variable (and other non-Bézout rings) fail to give PA -rings.

A final class of rings appears in studying the related single-input control or *feedback-cyclization* property:

(1.6) *There exist a u in R^m and a feedback K such that $(F + GK, Gu)$ is reachable.*

In the present context this property is of interest because it allows reducing a coefficient or pole-assignment problem to the (easy) case $m = 1$. Indeed, if (F, G) is reachable and satisfies (1.6), and if $p(z)$ is given, one may first find a (K_1, u_1) with $(F + GK_1, Gu_1)$ reachable and then, using (1.4) for the new system (now $m = 1$), find a K_2 with $p(z) = \text{ch.p.}(F + GK_1 + Gu_1K_2) = \text{ch.p.}(F + G[K_1 + u_1K_2])$. An FC_n -ring will be one for which (1.6) is true for any reachable system of dimension at most n , an FC -ring if for all n . Note that then $FC_n \subseteq CA_n \subseteq PA_n$. We shall give a characterization of FC_2 -rings among (almost all) principal-ideal domains.

2. The semilocal case

We review the basic results over fields, and a few results over rings given in [23]. Unless otherwise noted, R is an arbitrary commutative ring. We shall not distinguish between linear maps $R^s \rightarrow R^t$ and matrices in the canonical bases of R^s, R^t .

2.1. Lemma. *The pole-assignment property (1.5) implies reachability.*

Proof. When R is a field, one simply notes that $\text{ch.p.}(F + GK)$ is always divisible by $\text{ch.p.}(F_1)$, where F_1 is the map induced by F on $R^n/\text{Reach}(F, G)$; here $\text{Reach}(F, G)$ denotes the span of the columns of $(G, FG, \dots, F^{n-1}G)$. The case of arbitrary R follows from this. It must be proved that the map:

$$A := [G, FG, \dots, F^{n-1}G] : R^{mn} \rightarrow R^n \quad (2.2)$$

is onto. But this holds iff $A(M) := A \otimes (R/M)$ is onto for each maximal ideal M . Since, for each M , $(F(M), G(M))$ satisfies (1.5) over R/M (just lift the λ_i to R), it follows that $A(M)$ is indeed onto for each M . \square

The single-input case ($m = 1$) can be attacked in several ways. An interesting homological approach is given in [27]. An alternative approach is that used classically for $R = \text{field}$, which uses a concept which we shall need later, the *feedback group* $\mathbf{F}_{n,m}$: this is the group, acting on m -input, n -dimensional systems (F, G) , generated by the following three types of transformations:

$$F \rightarrow T^{-1}FT, \quad G \rightarrow T^{-1}G, \quad T \text{ in } \text{GL}(n, R), \quad (2.3a)$$

$$F \rightarrow F + GK, \quad G \rightarrow G, \quad K \text{ in } R^{m \times n}, \quad (2.3b)$$

$$F \rightarrow F, \quad G \rightarrow GB, \quad B \text{ in } \text{GL}(m, R). \quad (2.3c)$$

A considerable body of system-theoretic literature exists regarding problems related to $\mathbf{F}_{n,m}$ when R is a field. For R a complex polynomial ring, an algebraic-geometric study was initiated by Byrnes [3]. Two fundamental facts make this group relevant here. First, the set of reachable (F, G) is invariant under $\mathbf{F}_{n,m}$ (easy) and second, (1.4)–(1.5) are also invariant under the group (clear). The following is standard for R a field, with exactly the same proof (see for instance [14, Chapter 2]):

2.4. Lemma (Control canonical form). *Any single-input ($m = 1$) reachable (F, g) is $\mathbf{F}_{n,m}$ -equivalent to a system of the form*

$$F = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & & \vdots \\ \cdot & & & \ddots & & \cdot \\ \cdot & & & & \ddots & \cdot \\ \cdot & & & & & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 \end{pmatrix}, \quad G = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{pmatrix}.$$

In particular, $\mathbf{F}_{n,m}$ is transitive on single-input systems, and (1.4), (1.5) are true for such systems.

Proof. The reachability condition means that $g, Fg, \dots, F^{n-1}g$ give a basis for R^n . Let

$$\text{ch.p.}(F) = z^n + a_{n-1}z^{n-1} + \cdots + a_0,$$

and define a new basis

$$v_i := F^{n-i}g + a_{n-1}F^{n-i-1}g + \cdots + a_{n-i}g$$

(transformation of type (2.3a) and, in the new basis, apply (transformation of type (2.3b)) $K := (a_0, \dots, a_{n-1})$. \square

Fields are FC (and hence CA, PA) rings. This is proved as follows. If (F, G) is reachable, then, after if necessary reordering the nonzero columns g_1, \dots, g_r of G , $r \leq m$, there is a basis

$$\{g_1, Fg_1, \dots, F^{n_1-1}g_1, g_2, \dots, F^{n_2-1}g_2, \dots, F^{n_r-1}g_r\}$$

with the property that each $F^{n_i}g_i$ is dependent on the vectors to its left. Define $K: R^n \rightarrow R^m$ by:

$$K(F^i g_j) = 0 \quad \text{if } i < n_j - 1,$$

$$K(F^{n_i-1} g_i) = e_{i+1} \quad \text{if } i < r,$$

(where e_i is the i th canonical basis vector), and

$$K(F^{n_r-1} g_r) = \text{arbitrary.}$$

A simple calculation shows that $(F + GK, g_1)$ is reachable, as wanted.

A product R of FC-rings R_i is again an FC-ring, since finding K, u is equivalent to finding corresponding K_i, u_i over the R_i . It follows that a *semi-local ring* R is also an FC-ring. Indeed, for any given system (F, G) one may consider the system (\bar{F}, \bar{G}) obtained reducing modulo the radical of R . Since $R/\text{Rad}(R)$ is a product of fields, it is an FC-ring. So any (\bar{u}, \bar{K}) for (\bar{F}, \bar{G}) lift to a pair (u, K) satisfying (1.6) if (\bar{u}, \bar{K}) does. This shows that certain rings of interest (e.g., finite R), are FC-rings, and also points out the topological aspects of the obstructions to being an FC-ring. A more arithmetic aspect will be clear in Section 4.

3. Pole assignment

Let $F: R^n \rightarrow R^n$ and let S be a submodule of R^n . We shall say that λ is an *eigenvalue of F modulo S* iff there is a unimodular v in R^m with

$$Fv - \lambda v \in S. \tag{3.1}$$

By *unimodular* we mean that $v = (v_1, \dots, v_n)'$ generates a direct summand of R^n isomorphic to R ; denoting by $c(v)$ (the *content* of v) the ideal generated by the entries of the vector (or more generally, a matrix) v , unimodular means that $c(v) = R$.

We shall use the above for systems (F, G) , where $S := G$ is the image of G , and will just say eigenvalue “mod G ”. Its utility lies in the fact that λ is an eigenvalue of F mod G iff λ is an eigenvalue (with unimodular eigenvector) of some F_1 with (F_1, G_1) being $F_{n,m}$ -equivalent to (F, G) . Indeed, if (3.1) holds with $S = G$,

$$Fv - \lambda v = Gu,$$

a projection $R^n \rightarrow R$ on the span of v can be composed with the map $1 \rightarrow u$ to give a

$$K: R^n \rightarrow R^m, \quad Kv = u,$$

so that $(F + GK)v = \lambda v$, as wanted.

3.2. Remark. The following properties are equivalent for a given R :

- (a) For each reachable (F, G) , every λ in R is an eigenvalue of $F \bmod G$;
- (b) For each reachable (F, G) , $F^{-1}G$ has some unimodular element.

Indeed, assume that (b) is true, (F, G) is reachable, and λ is given. Consider $A := F - \lambda I$. Then (A, G) is again reachable. But v is in $A^{-1}G$ iff $Fv = \lambda v \bmod G$. Conversely, (b) is the particular case $\lambda = 0$ of (a). \square

3.3. Proposition. *Assume that R is a PA_k ring and that rank-one projective R -modules are free. Then, for each (F, G) reachable of dimension $\leq k$, there is some unimodular element in $F^{-1}G$.*

Proof. Let (F, G) be reachable of dimension $n \leq k$. Assume rank 1 projectives are free. Since R is a PA_k -ring, and $(F + I, G)$ is also reachable, there is a K such that $A := F + I + GK$ satisfies $\text{ch.p.}(A) = z^{n-1}(z - 1)$. By Cayley–Hamilton, $A^n = A^{n-1}$. Thus A^{n-1} is idempotent and $L := \text{Im}(A^{n-1})$ is a summand of R^n . Since L is also the kernel of $A - I$, which has (locally) a simple eigenvalue at zero, it follows that L has rank 1. Thus L is free, and a generator v of L is in $F^{-1}G$. \square

Recall that R is a *Hermite ring* iff stably free modules P (i.e. $P \oplus R^r$ free for some r) are necessarily free; R is *projective-free* iff finitely generated projectives are free.

3.4. Proposition. *Assume that R is a Hermite ring and that $F^{-1}G$ contains a unimodular element whenever (F, G) is reachable and of dimension $\leq k$. Then, for each reachable (F, G) of dimension $n \leq k$ and each $\lambda_1, \dots, \lambda_n$ in R , there is some (\hat{F}, \hat{G}) which is $F_{n,m}$ -equivalent to (F, G) and there are unimodular v_1, \dots, v_n with*

$$Fv_i = \lambda_i v_i \bmod \langle v_1, \dots, v_{i-1} \rangle, \quad i = 1, \dots, n. \quad (3.5)$$

In particular, R is a PA_k ring.

Proof. Let (F, G) be reachable, $\lambda_1, \dots, \lambda_n$ in R . By Remark 3.2, we may assume $(\bmod F_{n,m})$ that $Fv_1 = \lambda v_1$ for some unimodular v_1 . Since R is Hermite, the quotient $R^n / \langle v_1 \rangle \simeq R^{n-1}$. Further, F induces a map F_1 on R^{n-1} , and together with $G_1 (= G$ followed by the projection $R^n \rightarrow R^{n-1})$ constitutes again a reachable system. By induction on n , there are v_2, \dots, v_n giving (3.5) for F_1 , and this lifts to F in R^n .

When (3.5) holds, $\text{ch.p.}(\hat{F}) = (z - \lambda_1) \cdots (z - \lambda_n)$. Thus R is a PA_k ring. \square

The above diagonal reduction is in fact one of the direct ways known for establishing the pole-shifting result over a field.

3.6. Corollary. *Let R be a projective-free ring. Then R is a PA ring iff $F^{-1}G$ has a unimodular element whenever (F, G) is reachable.*

The following necessary condition is suggested by Morse's proof [19] that $\mathbf{R}[x]$ is a PA -ring:

3.7. Proposition. *Let R be a Bezout ring such that whenever a matrix A has content R , there is a vector v with Av unimodular. Then R is a PA-ring.*

Proof. By reachability of (F, G) , $c(G) = R$. Let g in G be unimodular, $L = R^n / \langle g \rangle \cong R^{n-1}$. Let \hat{F} be the composition of F with the projection onto L . Since R is Bezout, the image of \hat{F} is free and so $\ker \hat{F}$ is a nontrivial summand of R^n . Let v be unimodular in $\ker \hat{F}$. Then v is in $F^{-1}G$, and by Corollary 3.6, R is a PA-ring. \square

The above condition on matrices of content R holds in particular when every matrix A is known to be equivalent to a diagonal matrix, i.e. when for each A ,

$$A = PBQ,$$

with

$$B = \left(\begin{array}{ccc|ccc} d_1 & & & 0 & & 0 \\ & d_2 & & & 0 & 0 \\ & & \ddots & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & d_r & 0 \\ \hline & & & & & 0 \\ 0 & & & & 0 & 0 \end{array} \right)$$

Here $c(A) = c(B) = \langle d_1, \dots, d_r \rangle = R$ implies that $A(Q^{-1}w)$ is unimodular, where w is the column vector $(1, 1, \dots, 1)$. This property implies that R is Bezout, so that R is then a PA-ring. A particular case is that of *elementary divisor rings*, those for which a diagonalization as above always exists with $d_i | d_{i+1}$, $i = 1, \dots, r-1$. These rings were studied by Yohe [28], Leavitt and Whaples [17], Kaplansky [15], and others.

Remark. The existence of an element v as in Proposition 3.7 such that Av is unimodular, for other rings, may also occur if the content of the k th exterior power $A^k A$ of A is also R , $k > \text{Krull dimension of } R$ and projective R -modules are free; see [5]. Probably the required property in Proposition 3.7 is satisfied for all Bezout rings of dimension one. On the other hand, if R is an affine domain over \mathbb{C} of dimension one, it can be proved that it is only satisfied if R is a P.I.D.

3.8. Example. The ring R of real-analytic functions on an open interval I (finite or infinite) is an elementary divisor ring, hence a PA-ring. (These rings appear in studying single-parameter smooth families of systems, or in the algebraic theory of time-varying systems [13]). Using the criterion given by Helmer [7] (who proved that the ring of real *entire* functions is an EDR), we need to show that R is Bezout and that, for each f, g in R , there exists a *relatively prime* part $a = \text{RP}(f, g)$ of f with respect to g , where a divides f and is coprime with g , and such that any nonunit b dividing f/a has no common zero with g . But given f, g , let (C_i, n_i) be the set of

common zeroes with their multiplicities. By the Mittag-Leffler theorem there is an entire function h having precisely these zeroes. Then $a := f/h$ is $\text{RP}(f, g)$. Also, $k := (f/h)^2 + (g/h)^2$ is always nonzero, so a unit in R ; thus

$$h = \left(\frac{f}{hk}\right)f + \left(\frac{g}{hk}\right)g,$$

and $(f, g) = (h)$, proving that R is also Bezout. \square

We now give counterexamples showing that $\mathbf{R}[x, y]$ and $\mathbf{Z}[x]$ are not PA-rings. Both examples will have F invertible, so that Corollary 3.6 concludes that both $F^{-1}G$ and G must have a unimodular column combination.

3.9. Example. Let $R = \mathbf{R}[x, y]$. Consider

$$F := \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad G := \begin{pmatrix} y+x & y+x-1+x^2+y^2 \\ y-x & y-x+1-x^2-y^2 \end{pmatrix}$$

Thus $G = F\hat{G}$, where

$$\hat{G} = \begin{pmatrix} y & y \\ -x & -x+1-x^2-y^2 \end{pmatrix}.$$

We prove that (F, G) is reachable. Consider the matrix (G, FG) and let Δ_{ij} denote the minor corresponding to columns (i, j) . We shall see that $\Delta_{13} = \Delta_{24} = \Delta_{12} = 0$ has no complex solutions. The equation $\Delta_{13} = 0$ gives $x^2 + y^2 = 0$. Calculating mod $x^2 + y^2 = 0$, $\Delta_{24} = -2(2x - 1)$, so $x = \frac{1}{2}$. But $\Delta_{12} = 2y \text{ mod}(x^2 + y^2)$, so also $y = 0$, and this contradicts $x^2 + y^2 = 0$, $x = \frac{1}{2}$. So (F, G) is reachable. If R were a PA-ring, F being invertible, there would exist polynomials P, Q in R with

$$f(x, y) := P \begin{pmatrix} y \\ -x \end{pmatrix} + Q \begin{pmatrix} y \\ -x+1-x^2-y^2 \end{pmatrix}$$

unimodular and hence nonzero when (x, y) is in \mathbf{R}^2 . Restricting $f: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ to the unit circle S^1 ,

$$f(x, y) = a(x, y) \begin{pmatrix} y \\ -x \end{pmatrix},$$

a nonzero tangent vector. Thus the topological degree of

$$f' := f/|f|: S^1 \rightarrow S^1$$

is ± 1 . So f has a zero in the interior of S^1 , contradicting unimodularity of f . \square

Remark. The above example shows also that the ring of continuous functions on \mathbf{R}^2 is not a PA-ring, since only continuity of P, Q , was used above. Moreover, since $F + \lambda I$ has nonzero determinant whenever $\lambda \neq 1 \pm \sqrt{-1}$, the same proof will give that no eigenvalues different from $1 \pm \sqrt{-1}$ can be assigned for this (F, G) under

feedback. It is worthwhile to notice that the same example *does not* provide a negative result when the complex ring $\mathbf{C}[x, y]$ is considered instead: here

$$(x+1-iy)\begin{pmatrix} y \\ -x \end{pmatrix} - \begin{pmatrix} y \\ -x+1-x^2-y^2 \end{pmatrix}$$

is unimodular.

3.10. Example. Another counterexample is provided by $R := \mathbf{Z}[x]$. This is not a PA-ring. Indeed, let

$$F := \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \quad G := \begin{pmatrix} x-2 & 3 \\ -3 & x+2 \end{pmatrix}.$$

We claim that (F, G) is reachable. Write

$$A := \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} (F, G) = \begin{pmatrix} x+1 & 1-x & -3 & x+2 \\ -3 & x+2 & x-5 & x+5 \end{pmatrix}.$$

We prove that no maximal ideal M can contain the minors Δ_{12} , Δ_{23} , and Δ_{24} , of A . Otherwise $\Delta_{12} = x^2 + 5$, $\Delta_{12} + \Delta_{23} = 3(x+2)$, and $\Delta_{23} - \Delta_{24} = x(x+17)$ are again in M . In particular, either 3 or $x+2$ are in M . If 3 is in M and $x+2$ is in M , also $5 = \Delta_{12} - x^2$ belongs to M and $M = R$; but if x is not in M then $x+17$ is in M and so also $x-1$ is, implying that $(x^2+5) - x(x-1) = x+6$ would be in M and therefore x is in M . Assuming that 3 is not in M and $x+2$ is in M leads similarly to a contradiction. Thus (F, G) is reachable.

We claim that there is no unimodular combination of columns of $F^{-1}G$, or equivalently, of G . Consider

$$R^2 \xrightarrow{G} R^2 \rightarrow \text{coker}(G) \rightarrow 0. \tag{3.11}$$

If the image of G contains a unimodular element then $\text{coker}(G)$ is cyclic. We then show that $\text{coker}(G)$ cannot be cyclic. Tensoring by

$$S := R/\det G = \mathbf{Z}[\sqrt{-5}]$$

results in

$$\text{coker}(G \otimes S) = \text{coker}(G) \otimes S = I \text{ cyclic}. \tag{3.12}$$

But $\text{coker}(G \otimes S)$ is $(\mathbf{Z}[\sqrt{-5}])^2$ modulo the relations $(\sqrt{-5}-2)x_1 - 3x_2 = 0$, $3x_1 + (\sqrt{-5}+2)x_2 = 0$, i.e. the ideal $I = (3, \sqrt{-5}-2)$, which is not principal, contradicting (3.12). \square

Remark. An ideal-theoretic obstruction featured in examples such as 3.9 and 3.10 is the following. Let (F, G) be a reachable system of dimension two with entries in a Noetherian ring R , with F invertible. Assume that the ideal J of R generated by a column of G has grade two (see [16]). (If R is a unique factorization domain this means that J is not contained in a principal ideal). Let I be the image of J in the ring $S = R/(\det G)$. Then:

3.13. Lemma. *I² is a principal ideal.*

Proof. The “grade” condition implies that *I* is an invertible ideal of *S* and easily allows its identification with coker *G*, the *R*-module (in fact, *S*-module) obtained as the cokernel of $R^2 \xrightarrow{G} R^2$. Since *F* is invertible, we also have $I \cong \text{coker } FG$. To show *I*² principal, it suffices to prove that $I \oplus I$ can be generated by two elements [16]. For that consider the commutative diagram with exact rows:

$$\begin{array}{ccccccc}
 0 & \rightarrow & L & \rightarrow & R^2 \oplus R^2 & \xrightarrow{G+FG} & R^2 \rightarrow 0 \\
 & & \downarrow & & \downarrow & & \parallel \\
 & & & & G \cong FG & & \\
 & & \downarrow & & \downarrow & & \\
 0 & \rightarrow & R^2 & \rightarrow & R^2 \oplus R^2 & \longrightarrow & R^2 \rightarrow 0
 \end{array}$$

where the top map, $(G + FG)(u \oplus v) = G(u) + FG(v)$, is surjective because the system is reachable, while the bottom surjection is just ordinary “addition”. By the “snake lemma” (see [1, Chapter 1]), it follows that $I \oplus I \cong R^2 / (\text{image } L)$, as desired. \square

The remarks in this section seem to indicate that nonlocal rings of dimension greater than one will in general not be PA-rings. In particular, we leave as an

Open problem. Is $\mathbb{C}[x, y]$ a PA-ring?

4. PA₂- and CA₂-rings

We restrict our attention now to principal-ideal domains. These are always PA-rings but will turn out to be in general non CA- or FC-rings. *For the rest of this paper, R is a PID.* We first note:

4.1. Lemma. *Let (F, G) be reachable, of dimension 2. Then (F, G) is equivalent mod $F_{2,m}$ to some (nonunique) (F, G) of the form*

$$F = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 & : & 0 \\ b & c & : & 0 \end{pmatrix} \tag{4.2}$$

with *b, c, coprime.*

Proof. *R* being a PA-ring, one may assume that *F* has eigenvalues 0, 1. Reasoning as in Proposition 3.3, we may assume that *F* is as above. Reachability implies that each row of *G* is now unimodular. And a further transformation of type (2.3c) gives the first row of *G* as displayed. \square

We then have:

4.3. Proposition. *Assume that $R/(p)$ has characteristic different from 2 for all irreducibles p . Then R is an FC_2 -ring iff the following property holds:*

(4.4) *For each nonzero nonunit c in R and each irreducible p not dividing c , there is a unit ε such that εp is a square mod c .*

Proof. To prove sufficiency, note first that if (4.4) holds as stated then it clearly holds also for any $p = b$ not necessarily irreducible, as long as $(b, c) = 1$. By Lemma 4.1, we restrict our attention to those systems of the form (4.2) with $(b, c) = 1$. If c is a unit or $b = 0$, then G is invertible, and if c is zero, then F is cyclic with respect to the first column of G , so in any of these two cases the result is immediate. We assume then that c is nonzero and not a unit, $b \neq 0$, and look for (K, u) such that $F + GK$ is cyclic with generator Gu . In fact, we shall find (K, u) of the following special form:

$$K = \begin{pmatrix} -by & y \\ 0 & 0 \end{pmatrix}, \quad u = \begin{pmatrix} \alpha \\ 1 \end{pmatrix}. \tag{4.5}$$

In order that $\det(Gu, (F + GK)Gu) = \varepsilon = \text{unit}$, or $b\alpha^2 + c\alpha - \varepsilon = c^2y$, we need that there exist α, ε in R , with ε a unit, such that

$$b\alpha^2 + c\alpha - \varepsilon \equiv 0 \pmod{c^2}. \tag{4.6}$$

Since b is a unit mod c^2 , and since $\text{char}(R/(c^2)) \neq 2$, one can solve (4.6) for α using the quadratic formula if $c^2 + 4b\varepsilon$ is a square mod c^2 . This is now equivalent to $b\varepsilon$ being a square mod c^2 , which we have shown to be equivalent to (4.4), since $(b, c^2) = 1$.

To prove that (4.4) is necessary, let p, c be as in (4.4) and take a in R such that $a = p^{-1} \pmod{c}$. Consider the system

$$F := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad G := \begin{pmatrix} 1 & 0 \\ a & c \end{pmatrix},$$

and assume that there are

$$K = \begin{pmatrix} x & y \\ v & w \end{pmatrix}, \quad u = \begin{pmatrix} \alpha \\ \gamma \end{pmatrix}$$

with

$$\det(Gu, (F + GK)Gu) = \varepsilon = \text{unit}. \tag{4.7}$$

Calculating mod c , (4.7) becomes

$$a\alpha^2 \equiv \varepsilon \pmod{c},$$

so also

$$\alpha^2 \equiv \varepsilon p \pmod{c},$$

as wanted. \square

4.8. Corollary. *If $R/(q)$ has for all q irreducible, characteristic $\neq 2$ and no quadratic extensions, R is an FC_2 -ring.*

Proof. If every element is a square modulo q , it is easy to verify by induction in r that every unit (mod q) is a square modulo q^r for all r . By the Chinese Remainder Theorem, then, (4.4) holds for any c . \square

So $C[x]$ is an FC_2 -ring. For $R[x]$ note that (4.4) implies that $p(x)$ must have constant sign at all roots of c ; thus $p := x - 1$, $c := x(x - 2)$ show that $R[x]$ is not an FC_2 -ring. It is clear from the proof of Proposition 4.3 that an FC_2 -ring satisfies (4.4) whenever $R/(p)$ has non-two characteristic, even if $R/(q)$ has characteristic 2 for other q ; thus Z is not an FC -ring either (e.g.: $c = 15$, $b = 2, 7$, or 13).

4.9. Conjecture. $C[x]$ is an FC -ring.

Turning to (more general) CA_2 -rings, the problem now is to find, for a given (F, G) , and given α, β , a K with $\text{ch.p.}(F + GK) = z^2 - \alpha z + \beta$, i.e. $\text{trace}(F + GK) = \alpha$ and $\det(F + GK) = \beta$. Applying

$$T = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix},$$

we shall use the general form

$$F = \begin{pmatrix} 0 & 0 \\ b & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix} \quad (4.10)$$

rather than the above one. If K is as in (4.7), one needs then

$$x + cw + 1 = \alpha \quad (4.11a)$$

$$x(1 + cw) - y(b + cv) = \beta. \quad (4.11b)$$

Solving for x in (4.11a) and substituting in (4.11b) results in

$$c^2w^2 + c(2 - \alpha)w + (1 + \beta - \alpha) = (-y)(b + cv). \quad (4.12)$$

So there will exist a K as desired iff the left side of (4.12) has a solution in w , modulo $(b + cv)$. This requires that there be a solution of

$$c^2w^2 + c(2 - \alpha)w + (1 + \beta - \alpha) \equiv 0 \pmod{p} \quad (4.13)$$

for each irreducible factor p of $(b + cv)$. Standard techniques from elementary number theory allow to also recover (4.12) from (4.13) together with additional conditions at the primes p with $R/(p)$ of characteristic 2. Since $(b, c) = 1$, we have that $(p, c) = 1$ for the p of (4.13). If also $\text{char } R/(p) \neq 2$, (4.13) is equivalent to the requirement that the discriminant

$$c^2(2 - \alpha)^2 - 4c^2(1 + \beta - \alpha) = c^2(\alpha^2 - 4\beta)$$

be a square modulo p . We may then construct a counterexample to CA_2 by showing that each element of the form $b + cv$ has an irreducible factor p for which $(\alpha^2 - 4\beta)$ is not a square modulo p .

Now consider $R = \mathbf{Z}$. The congruence (4.13) is tested by the Legendre symbol

$$\left(\frac{\alpha^2 - 4\beta}{p}\right).$$

We can guarantee that (4.13) fails for some p dividing $b + cv$ if we force $b + cv$ to be odd and the Jacobi symbol

$$\left(\frac{\alpha^2 - 4\beta}{b + cv}\right) = -1.$$

Quadratic reciprocity allows us to express this last condition in terms of the values taken on by $b + cv$ modulo $\alpha^2 - 4\beta$. (One must also note that the usual quadratic reciprocity formulation deals only with positive integers, but that the law may be still applied formally if either of these two quantities is positive).

Thus if we take $\alpha^2 - 4\beta$ odd and positive, and choose c divisible by $\alpha^2 - 4\beta$ and b such that

$$\left(\frac{b}{\alpha^2 - 4\beta}\right) = -1,$$

then we will have shown that \mathbf{Z} is not a CA -ring. In particular, we may take $\alpha := 1$, $\beta := -1$, $c := 10$, $b := 3$.

The above reasoning can be extended to $\mathbf{R}[x]$. For this we note that there is a (simple) theory of quadratic reciprocity on $\mathbf{R}[x]$, which we now explain. Given polynomials f, p , with p irreducible, we have the *Legendre symbols*

$$\left(\frac{f}{p}\right) := \begin{cases} 0 & \text{if } p/f, \\ -1 & \text{if } p \text{ is linear and } f < 0 \pmod{p}, \\ +1 & \text{otherwise.} \end{cases} \quad (4.14)$$

Note that these satisfy

$$\left(\frac{ca_1a_2}{p}\right) = (-1)^{(\text{sgn } c)(\text{deg } p)} \left(\frac{a_1}{p}\right) \left(\frac{a_2}{p}\right), \quad (4.15)$$

for polynomials a_i and constant c . For arbitrary f, g , $g = c \prod q_i^{r_i}$ with the q_i irreducible, the *Jacobi symbols* are

$$\left(\frac{f}{g}\right) := \prod \left(\frac{f}{q_i}\right)^{r_i} \quad (4.16)$$

4.17. Lemma. Let $f = \varepsilon \prod p_i^{r_i}$, $g = \delta \prod q_j^{s_j}$, then

$$\left(\frac{f}{g}\right) \left(\frac{g}{f}\right) = (-1)^{(\deg f)(\deg g) + (\operatorname{sgn} \varepsilon)(\deg g) + (\operatorname{sgn} \delta)(\deg f)} \quad (4.18)$$

Proof. By (4.15) and the definition (4.16),

$$\left(\frac{f}{g}\right) \left(\frac{g}{f}\right) = (-1)^{(\operatorname{sgn} \varepsilon)(\deg g) + (\operatorname{sgn} \delta)(\deg f)} \prod \left(\frac{P_i}{Q_j}\right)^{r_i s_j},$$

so it is enough to prove the lemma for f, g both monic irreducible. If both f, g have degree 2, there's nothing to prove. If $f(x) = x - a$ and $\deg g = 2$, then $\left(\frac{f}{g}\right) = +1$. But g has no real roots and, (being monic), is positive for large x . Thus $g(a) > 0$ and also $\left(\frac{g}{f}\right) = +1$. If both f, g have degree one, $g = x - b$, then

$$\left(\frac{g}{f}\right) = \operatorname{sgn}(a - b) = -\operatorname{sgn}(b - a) = -\left(\frac{f}{g}\right),$$

as wanted. \square

We apply this to show that $\mathbf{R}[x]$ is *not* a CA-ring. Consider (4.10) with $b := x$, $c := x^2 - 1$, and let $\alpha := 2x$, $\beta := 1$. We need to prove that

$$\alpha^2 - 4\beta = 4(x^2 - 1), \quad \text{or just } x^2 - 1,$$

is not a square mod some prime p dividing $x + (x^2 - 1)v$. In terms of Legendre symbols

$$\left(\frac{x - (x^2 - 1)v}{x^2 - 1}\right) = \left(\frac{x - (x^2 - 1)v}{x - 1}\right) \left(\frac{x - (x^2 - 1)v}{x + 1}\right) = -1.$$

By Lemma 4.17,

$$\left(\frac{x^2 - 1}{x - (x^2 - 1)v}\right) = (-1)^{2d + 0b + 2(\operatorname{sgn} \delta)} (-1) = -1,$$

where d is the degree of $x - (x^2 - 1)v$. Thus

$$\left(\frac{x^2 - 1}{p}\right) = -1$$

for some p dividing $x - (x^2 - 1)v$, as wanted. \square

We close this with the statement, for R a PID, of an

Open problem. Is $\text{FC} = \text{CA}$?

References

- [1] N. Bourbaki, *Commutative Algebra* (Addison-Wesley, Reading, MA, 1972).
- [2] R. Brockett and J.L. Willems, Discretized partial differential equations: examples of control systems defined on modules, *Automatica* 10 (1974) 507–515.
- [3] C. Byrnes, On the control of certain infinite dimensional systems by algebro-geometric techniques, *Amer. J. Math.* (1978).
- [4] S. Eilenberg, *Automata, Languages, and Machines*, Vol. A. (Academic Press, New York, 1974).
- [5] D. Eisenbud and E.G. Evans, Generating modules efficiently: theorems from algebraic K -theory, *J. Algebra*, 27 (1973).
- [6] M. Fliess, *Matrices de Hankel*, *J. Math. Pures Appl.* 53 (1974) 197–224.
- [7] O. Helmer, Divisibility properties of integral functions, *Duke Math. J.* 6 (1940) 345–356.
- [8] M. Heymann, Comments on “Pole assignment in multi-input controllable systems”, *IEEE Trans. Autom. Contr.*, AC-13 (1969) 748–749.
- [9] R. Johnson, *Linear systems over various rings*, Ph.D. dissertation, M.I.T., Cambridge, MA (1973).
- [10] E.W. Kamen, On an algebraic theory of systems defined by convolution operators, *Math. System Theory* 9 (1975) 57–74.
- [11] E.W. Kamen, *Lectures on algebraic systems theory: linear systems over rings*, N.A.S.A. Contractor Report 3016 (1978).
- [12] E.W. Kamen, An operator theory of linear functional differential equations, *J. Diff. Equations* 27 (1978).
- [13] E.W. Kamen, New results in realization theory for linear time-varying analytic systems, *IEEE Trans. Autom. Contr.* (1980) to appear.
- [14] R.E. Kalman, P.L. Falb and M.A. Arbib, *Topics in Mathematical System Theory* (McGraw-Hill, New York, 1969).
- [15] I. Kaplansky, Elementary divisors and modules, *Trans. Amer. Math. Soc.* 66 (1949) 464–491.
- [16] I. Kaplansky, *Commutative Rings* (Univ. Chicago Press, Chicago, IL, 1974).
- [17] W. Leavitt and G. Whaples, On matrices with elements in a principal-ideal ring, *Bull. Amer. Math. Soc.* 55 (1949) 117–118.
- [18] M.M. Matluk and A. Gill, Linear sequential circuits over rings, *Proc. Int. IEEE Conf. on Systems Networks and Computers*, Vol. I. (1971).
- [19] A.S. Morse, Ring models for delay-differential systems, *Automatica*, 12 (1976) 529–531.
- [20] Y. Rouchaleau, *Linear, discrete time, finite-dimensional systems over some classes of commutative rings*, Ph.D. dissertation, Stanford, CA (1972).
- [21] Y. Rouchaleau and E.D. Sontag, On the existence of minimal realizations of linear dynamical systems over Noetherian integral domains, *J. Comput. System Sci.* 18 (1979) 65–75.
- [22] Y. Rouchaleau, B. Wyman and R.E. Kalman, Algebraic structure of linear dynamical systems, III, Realization theory over a commutative ring, *Proc. Nat. Acad. Sci., U.S.A.* 69 (1978) 3404–3406.
- [23] E.D. Sontag, Linear systems over commutative rings: A survey, *Ric. di Automatica* 7 (1976) 1–34.
- [24] E.D. Sontag, On split realization of response maps over rings, *Inform. and Control* 37 (1978) 23–33.
- [25] J.L. Willems, Optimal control of a uniform string of moving vehicles, *Ric. di Automatica* 2 (1971) 184–192.
- [26] W.M. Wonham, *Linear Multivariable Control* (Springer, New York, 1974).
- [27] B. Wyman, Pole placement over integral domains, *Comm. Algebra* (1978).
- [28] C.R. Yohe, Triangular and diagonal forms for matrices over commutative Noetherian rings, *J. Algebra* 6 (1967) 335–368.
- [29] D.C. Youla, The synthesis of networks containing lumped and distributed parameters, *Network and Switching Theory* 11 (1968) 73–133.