Minimal and canonical rational generator matrices for convolutional codes

Forney, Jr., G David; Johannesson, Rolf; Wan, Zhe-Xian

Published in:
IEEE Transactions on Information Theory

DOI:
10.1109/18.556681

1996

Minimal and Canonical Rational Generator Matrices for Convolutional Codes

G. David Forney, Jr., *Fellow, IEEE*, Rolf Johannesson, *Senior Member, IEEE*, and Zhe-xian Wan

***Abstract*— A full-rank $k \times n$ matrix $G(D)$ over the rational functions $F(D)$ generates a rate $R = k/n$ convolutional code $\mathcal{C}$. $G(D)$ is minimal if it can be realized with as few memory elements as any encoder for $\mathcal{C}$, and $G(D)$ is canonical if it has a minimal realization in controller canonical form. We show that $G(D)$ is minimal if and only if for all rational input sequences $\boldsymbol{u}(D)$, the span of $\boldsymbol{u}(D)G(D)$ covers the span of $\boldsymbol{u}(D)$. Alternatively, $G(D)$ is minimal if and only if $G(D)$ is globally zero-free, or globally invertible. We show that $G(D)$ is canonical if and only if $G(D)$ is minimal and also globally orthogonal, in the valuation-theoretic sense of Monna.**

***Index Terms*—Algebraic structure of convolutional codes, minimal generator matrices, canonical generator matrices, behavioral system theory, valuation theory.**

I. Introduction

THE structural properties of convolutional codes and their generator matrices have been investigated in a series of papers by Forney [1]–[3] and Johannesson and Wan [4], [5].

In [1], it was shown that every convolutional code $\mathcal{C}$ can be generated by a polynomial generator matrix $G(D)$ that has a polynomial inverse $G^{-1}(D)$ (is "basic") and that can be realized in controller canonical form with as few memory elements as any encoder for $\mathcal{C}$. The latter property requires that $G(D)$ not only be "basic" but also have a property called the "predictable degree property." More elementary proofs of these results were given in [2] and in [4].

In [4] and [5], the authors also pointed out that the terminology of [1] could be confusing, and improved it by defining a *minimal* generator matrix $G(D)$ as one that can be realized with as few memory elements as any encoder for $\mathcal{C}$, and a *canonical* generator matrix as one that has a minimal realization in controller canonical form. We continue to use this terminology here. In these terms, it is shown in [1] that every convolutional code $\mathcal{C}$ has a canonical polynomial generator matrix $G(D)$, called "minimal" in [1] and "minimal-basic" in [4].

Generalizations of these definitions and results to rational generator matrices, using ideas from valuation theory, are sketched in [2, Appendix] and [3, sec. 3.5]. However, no proofs are given, and some of the claimed extensions are not entirely correct.

Manuscript received December 10, 1995; revised July 1, 1996. The material in this paper was presented in part at the Meeting on Information Theory: Algebraic, Combinatorial and Probabilistic Codes and Coding Techniques, Oberwolfach, Germany, February 18–24, 1996.

G. D. Forney, Jr. is with Motorola, Inc., Mansfield, MA 02048 USA.

R. Johannesson and Z.-x. Wan are with the Department of Information Theory, Lund University, S-221 00 Lund, Sweden.

Publisher Item Identifier S 0018-9448(96)07888-1.

In [5] (see also [6]), some steps were taken toward a more rigorous development of results for rational generator matrices along the lines suggested in [2] and [3]. The purpose of this paper is to complete this development. The rationale for this is twofold. First, rational generator matrices are sometimes preferred as convolutional encoders, e.g., in most trellis coding applications (e.g., V.32 and V.34 modems), and also in some newly developed iterative decoding schemes [7]. Second, it is of mathematical interest to generalize known results for polynomial generator matrices and to understand them in a more general setting. Moreover, such a generalization may suggest even further extensions.

The main results of this paper are a) a generalization of the theory of [1], etc., to rational generator matrices using concepts from valuation theory; b) the new, general minimality test of Theorems 8 and 9, which applies to rational generator matrices; and c) a demonstration that canonicality is the intersection of two independent properties: minimality and the global predictable valuation property (global orthogonality).

In Section II, we review ideas from valuation theory which are useful in the study of rational generator matrices, as originally suggested in [2]. The valuation-theoretic notion of orthogonality due to Monna [8] suggests a natural generalization of the predictable degree property called the "predictable valuation property," which will be discussed in Section III. In Section IV, we establish simple necessary and sufficient conditions for a polynomial generator matrix to be canonical. These results use only elementary linear algebra and an elementary behavioral-theoretic construction of Forney and Trott [9]. In Section V, we give necessary and sufficient conditions for a rational generator matrix to be minimal. Canonicality is revisited in Sections VI and VII, in which we extend our results on canonicality to rational generator matrices of rates $R = 1/n$ and $R = k/n$, respectively. In Section VIII, we use the extended invariant factor theorem to show the equivalence of minimality to global invertibility. A brief discussion in Section IX concludes the paper.

Remark: The appendix of [2] stimulated a flurry of papers in system theory by a group of authors including Kailath [10, ch. 6], Kung [11], Lévy [12], Verghese [13], and Wyman [14]. These papers use the language of valuation theory to provide a unified treatment of the entire pole/zero structure of rational matrices over the complex field, including those at zero and infinity. Most of the results of this paper having to do with globally orthogonal matrices (which [13] calls "globally (column-) reduced") first appeared in [2] or in these papers. In particular, the authors of [13] developed most of our Theorem

5, as well as the greedy construction algorithm for choosing a set of rational generators with minimum defect.

Moreover, Kailath [10, ch. 6], [13] traces the roots of these ideas back to Wedderburn [15], Vekua [16] (see also the references therein), and Gantmacher [17]. Wyman has taken to calling our invariant structure indices "Forney-Wedderburn indices."

This work seems to have been regarded by system theorists as mostly a matter of new language for known concepts, and eventually it died away. However, none of this work appears to have addressed realizations, or in particular minimality. It thus missed what we would regard as the most interesting system-theoretic results of this paper, such as the identification of minimality with global zero-freedom or with global invertibility.

II. Polynomials, Rational Functions, and Valuation Theory

It was suggested in [2] that valuation theory might provide the natural language for generalization of the results of [1] and [2]. In this section we present a brief, self-contained exposition of the few elementary concepts of valuations of the field of rational functions that will be needed in this paper. For a general introduction to valuation theory, see Jacobson [18] or Monna [8].

The set $F[D]$ of polynomials in the indeterminate delay operator D over a field F is a ring which, like the ring of integers, is a principal ideal domain. This implies unique factorization: every nonzero polynomial can be uniquely factored into a product of primes, up to units. The units in $F[D]$ are the set $\mathcal{U}$ of polynomials of degree zero—i.e., the nonzero elements of F. The primes in $F[D]$ are the set $\mathcal{P}$ of monic irreducible polynomials. (A polynomial $p(D) = p_0 + p_1 D + \cdots + p_l D^l$ is said to be monic if $p_l = 1$.) For simplicity we write p for the monic irreducible polynomial $p(D)$ in $\mathcal{P}$. A nonzero polynomial $f(D) \in F[D]$ can be uniquely written as

$$f(D) = u \prod_{p \in \mathcal{P}} p^{e_p(f(D))}$$

for some unit $u \in \mathcal{U}$. The exponents $\{e_p(f(D)), p \in \mathcal{P}\}$ that occur in this unique factorization are called the *valuations* of the polynomial $f(D)$ at the primes p, or the *p-valuations* of $f(D)$. If $f(D) = 0$, then by convention we define $e_p(0) = \infty$ for all p in $\mathcal{P}$.

The set $F(D)$ of rational functions in the indeterminate D over the field F is the field of quotients of $F[D]$, namely, the set of all $r(D) = f(D)/g(D)$ with $f(D),\ g(D) \in F[D]$ and $g(D) \neq 0$. The quotient is made unique by requiring that $g(D)$ be monic and that $f(D)/g(D)$ be reduced to lowest terms by cancellation of common factors. Again, we have unique factorization of rational functions up to units

$$r(D) = f(D)/g(D) = u \prod_{p \in \mathcal{P}} p^{e_p(f(D)) - e_p(g(D))} = u \prod_{p \in \mathcal{P}} p^{e_p(r(D))}$$

for some $u \in \mathcal{U}$, where the p-valuation $e_p(r(D))$ is defined as the difference

$$e_p(r(D)) = e_p(f(D)) - e_p(g(D)).$$

Thus the p-valuation $e_p(r(D))$ of a rational function may be negative for $p \in \mathcal{P}$.

The p-valuation $e_p(r(D))$ for each $p \in \mathcal{P}$ has the defining properties of an exponential nonarchimedean valuation [18]:

i) (uniqueness of 0):

$$e_p(r(D)) = \infty, \text{ if and only if } r(D) = 0;$$

ii) (additivity):

$$e_p(r(D)s(D)) = e_p(r(D)) + e_p(s(D));$$

iii) (strong triangle inequality):

$$e_p(r(D) + s(D)) \geq \min\{e_p(r(D)), e_p(s(D))\}.$$

There is one more nontrivial valuation on $F(D)$, viz., the negative degree function

$$e_{D^{-1}}(r(D)) \stackrel{\text{def}}{=} \begin{cases} \deg g(D) - \deg f(D), & \text{if } r(D) = f(D)/g(D) \neq 0, \\ & \text{where } f(D), g(D) \in F[D]; \\ \infty, & \text{if } r(D) = 0. \end{cases}$$

The reason for calling this function a valuation at $p = D^{-1}$ will be explained below. It is easily verified that this D^{-1}-valuation satisfies properties i)–iii).

We denote by $\mathcal{P}^*$ the set consisting of the elements of $\mathcal{P}$ plus D^{-1}

$$\mathcal{P}^* \stackrel{\text{def}}{=} \mathcal{P} \cup \{D^{-1}\}.$$

It is easy to see that the p-valuations $e_p(r(D))$ are equal to zero for all p in $\mathcal{P}^*$ if and only if $r(D)$ is a unit (a nonzero element of F). Thus all of the p-valuations for p in $\mathcal{P}^*$ are trivial on F. It is shown in valuation theory [18] that there are no other valuations of the rational functions $F(D)$ that are trivial on F.

For $r(D) \neq 0$, from the unique factorization of $r(D)$ we have immediately the important *product formula*, as it is called in valuation theory [18], written here in additive form since we are using exponential valuations

$$\sum_{p \in \mathcal{P}^*} e_p(r(D)) \deg p = 0$$

where the degree of D^{-1} is defined as 1.

Example 1: Let F be the binary field $\mathbb{F}_2$, and let $f(D) = D + D^2 + D^3$. Then

$$e_D(f(D)) = 1, \quad e_{1+D+D^2}(f(D)) = 1, \quad e_{D^{-1}}(f(D)) = -3$$

and all other p-valuations are equal to zero. The reader may verify easily that the product formula holds.

The *delay* of a rational function $r(D)$ is defined as

$$\operatorname{del} r(D) = e_D(r(D)).$$

Similarly, the *degree* of a rational function $r(D)$ may be defined as

$$\deg r(D) = -e_{D^{-1}}(r(D))$$

which for polynomials coincides with the standard definition.

A rational function $r(D)$ is

a) *causal*

$$\text{if } \operatorname{del} r(D) \geq 0, \text{ i.e., if } e_D(r(D)) \geq 0;$$

b) *polynomial*

$$\text{if } e_p(r(D)) \geq 0 \text{ for all } p \in \mathcal{P};$$

c) *finite*

$$\text{if } e_p(r(D)) \geq 0 \text{ for all } p \in \mathcal{P} \text{ except possibly } D.$$

Causal rational functions are also sometimes called proper (particularly when z-transforms are used rather than D-transforms), and finite rational functions are also called *Laurent polynomials*.

A rational function $r(D)$ may be expanded by long division into a formal Laurent series in powers of D and thus identified with a semi-infinite sequence over F that begins with all zeroes; for example

$$1/(1-D) = 1 + D + D^2 + \cdots.$$

In this way, the set of rational functions $F(D)$ may be identified with a subset of the set $F((D))$ of formal Laurent series in D over F, which we shall call the rational formal Laurent series. These are precisely the formal Laurent series that eventually become periodic. The first nonzero term of a formal Laurent series expansion of $r(D)$ in powers of D is the term involving $D^{e_D(r(D))} = D^{\operatorname{del} r(D)}$; i.e., the formal Laurent series in D "starts" at a "time index" equal to the delay $e_D(r(D))$ of $r(D)$.

Alternatively, a rational function may be expanded similarly into a formal Laurent series in D^{-1}; for example

$$1/(1-D) = -D^{-1} - D^{-2} - \cdots.$$

In this way, $F(D)$ may alternatively be identified with a subset of $F((D^{-1}))$. If elements of $F((D^{-1}))$ are identified with semi-infinite sequences over F that finish with all zeroes, then $r(D)$ "ends" at a time equal to the degree $-e_{D^{-1}}(r(D))$ of $r(D)$. This hints at why we use the notation $p = D^{-1}$ for this valuation.

We should emphasize that this second, alternative expansion is a purely mathematical construct, and that when we wish to identify a rational function in $r(D)$ with a physical sequence of elements of F, we shall always use the first formal Laurent series expansion in powers of D.

A finite rational function (Laurent polynomial) may be written as $r(D) = f(D)/D^n$ for $f(D) \in F[D]$ and $n \in Z$. The two expansions of a finite rational function $r(D)$ as formal Laurent series in D and in D^{-1} coincide, and $r(D)$ "starts" at time $\operatorname{del} r(D)$ and "ends" at time $\deg r(D)$.

More generally, a rational function $r(D)$ may be expanded as a formal Laurent series in powers of p for any p in $\mathcal{P}^*$, as follows. Let $r(D) = f(D)/g(D)$, where $f(D)$ and $g(D) \neq 0$ are polynomial. If $f(D) = 0$, then the formal power series in p is simply $f(D) = 0$. If $f(D) \neq 0$, then we may write

$$f(D) = [f(D)]_p p^{e_p(f(D))} + f^{(1)}(D)$$

where $e_p(f(D))$ is the p-valuation of $f(D)$, $[f(D)]_p$ is the residue of $f(D)p^{-e_p(f(D))}$ modulo p, and $f^{(1)}(D)$ is a polynomial (possibly 0) whose p-valuation is greater than $e_p(f(D))$. Iterating this process, possibly indefinitely, we obtain a formal Laurent series in p whose first nonzero term is $[f(D)]_p\, p^{e_p(f(D))}$. Similarly, we may expand the denominator $g(D)$ into a formal Laurent series in p whose first nonzero term is $[g(D)]_p\, p^{e_p(g(D))}$. Then by long division we obtain a formal Laurent series expansion of $r(D)$ in powers of p whose first term is $[r(D)]_p\, p^{e_p(r(D))}$, where

$$e_p(r(D)) = e_p(f(D)) - e_p(g(D))$$
$$[r(D)]_p = [f(D)]_p/[g(D)]_p.$$

This division is well-defined because $[f(D)]_p$ and $[g(D)]_p$ are nonzero residues of polynomials in $F[D]$ modulo p, which may be regarded as elements of the quotient ring $F[D]_p = F[D]/pF[D]$, which is actually a field since p is an irreducible (prime) polynomial.

If $r(D) = 0$, then in addition to $e_p(0) = \infty$, we define $[0]_p = 0$ for all p in $\mathcal{P}^*$.

Note that this general expansion method works perfectly well for $p = D^{-1}$, if we take $[f(D)]_{D^{-1}}$ and $[g(D)]_{D^{-1}}$ to be the coefficients of the highest order terms of $f(D)$ and $g(D)$, respectively, i.e., the coefficients of $D^{\deg f(D)}$ and $D^{\deg g(D)}$, respectively. This explains our use of the notation $p = D^{-1}$ for this valuation. In our previous papers [1], [2], [4], [5], we denoted the highest order coefficient $[f(D)]_{D^{-1}}$ by $[f(D)]_h$.

Example 1 (cont.): For $f(D) = D + D^2 + D^3$ (or indeed for any nonzero polynomial in D), the formal Laurent series in the polynomial D is simply $f(D)$. We have $e_D(f(D)) = 1$, $[f(D)]_D = 1$, and the first nonzero term of the series is $[f(D)]_D D^{e_D(f(D))} = D$. Similarly, for $p = D^{-1}$, we have $e_{D^{-1}}(f(D)) = -3$, $[f(D)]_{D^{-1}} = 1$, and the formal Laurent series in D^{-1} is

$$f(D) = (D^{-1})^{-3} + (D^{-1})^{-2} + (D^{-1})^{-1}$$

whose first nonzero term is

$$[f(D)]_{D^{-1}}(D^{-1})^{e_{D^{-1}}(f(D))} = (D^{-1})^{-3}.$$

For $p = 1 + D$, we have

$$e_{1+D}(f(D)) = 0,\ [f(D)]_{1+D} = 1$$

and the formal Laurent series in $1 + D$ is

$$f(D) = (1+D)^0 + (1+D)^3$$

whose first nonzero term is

$$[f(D)]_{1+D}(1+D)^{e_{1+D}(f(D))} = (1+D)^0.$$

For $p = 1 + D + D^2$, we have

$$e_{1+D+D^2}(f(D)) = 1,\ [f(D)]_{1+D+D^2} = D$$

and the formal Laurent series in $1 + D + D^2$ is simply

$$f(D) = D(1 + D + D^2)^1$$

whose first and only nonzero term is

$$[f(D)]_{1+D+D^2}(1 + D + D^2)^{e_{1+D+D^2}(f(D))} = D(1 + D + D^2)^1.$$

Example 2: Let $F = \mathbb{F}_2$ and

$$r(D) = (D^3 + D^5)/(1 + D + D^2).$$

Then

$$\begin{aligned} e_D(r(D)) &= 3 = \operatorname{del} r(D) \\ e_{D^{-1}}(r(D)) &= -3 = -\deg r(D) \\ e_{1+D}(r(D)) &= 2 \\ e_{1+D+D^2}(r(D)) &= -1 \end{aligned}$$

and all other p-valuations are zero. It is easy to verify that the product formula holds. Also, $[r(D)]_D = [r(D)]_{D^{-1}} = [r(D)]_{1+D} = 1$ and $[r(D)]_{1+D+D^2} = D$.

III. Predictable Valuation Property of Rational Matrices

In this section we first extend valuations to vectors of rational functions. Then we see that the concept of "orthogonality" that was introduced into valuation theory by Monna [8] provides a natural generalization of the predictable degree property of rational matrices.

If $\boldsymbol{r}(D) = (r_1(D), \cdots, r_s(D))$ is a vector of rational functions $r_i(D) \in F(D)$, then the p-valuation $e_p(\boldsymbol{r}(D))$ is defined for all $p \in \mathcal{P}^*$ by the "box norm" [8], namely

$$e_p(\boldsymbol{r}(D)) \stackrel{\text{def}}{=} \min_{1 \le i \le s} \{e_p(r_i(D))\}.$$

This generalizes the notion of the "greatest common divisor"; indeed, if $\boldsymbol{r}(D)$ is a set of polynomials, then the greatest common divisor of the set $\boldsymbol{r}(D)$ is

$$\gcd \boldsymbol{r}(D) = \prod_{p \in \mathcal{P}} p^{e_p(\boldsymbol{r}(D))}.$$

Definition 1: A vector $\boldsymbol{r}(D) = (r_1(D), \cdots, r_s(D))$ of rational functions is *zero-free* if $e_p(\boldsymbol{r}(D)) \le 0$ for all $p \in \mathcal{P}^*$.

In Section VIII we will find that zero-freedom implies a certain kind of invertibility.

We may generalize the definition of delay and degree to a vector $\boldsymbol{r}(D)$ as follows:

$$\begin{aligned} \operatorname{del} \boldsymbol{r}(D) &= e_D(\boldsymbol{r}(D)) = \min_i \{\operatorname{del} r_i(D)\} \\ \deg \boldsymbol{r}(D) &= -e_{D^{-1}}(\boldsymbol{r}(D)) = \max_i \{\deg r_i(D)\}. \end{aligned}$$

Properties i)–iii), appropriately generalized, continue to hold:

i) $e_p(\boldsymbol{r}(D)) = \infty$, if and only if $\boldsymbol{r}(D) = \boldsymbol{0}$;
ii) $e_p(k(D)\boldsymbol{r}(D)) = e_p(k(D)) + e_p(\boldsymbol{r}(D))$ for all $k(D) \in F(D)$;
iii) $e_p(\boldsymbol{r}(D) + \boldsymbol{s}(D)) \ge \min\{e_p(\boldsymbol{r}(D)), e_p(\boldsymbol{s}(D))\}$.

However, the product formula becomes an inequality, since for any i

$$\sum_{p \in \mathcal{P}^*} e_p(\boldsymbol{r}(D)) \deg p \le \sum_{p \in \mathcal{P}^*} e_p(r_i(D)) \deg p = 0.$$

We therefore define the *defect* of $\boldsymbol{r}(D)$ as the nonnegative quantity [2]

$$\operatorname{def} \boldsymbol{r}(D) \stackrel{\text{def}}{=} -\sum_{p \in \mathcal{P}^*} e_p(\boldsymbol{r}(D)) \deg p$$

which can also be written as

$$\begin{aligned} \operatorname{def} \boldsymbol{r}(D) &= \deg \boldsymbol{r}(D) - \sum_{p \in \mathcal{P}} e_p(\boldsymbol{r}(D)) \deg p \\ &= \deg \boldsymbol{r}(D) - \deg (\gcd \boldsymbol{r}(D)). \end{aligned}$$

In view of property ii) and the product formula, we have for all nonzero $k(D) \in F(D)$

$$\operatorname{def} k(D)\boldsymbol{r}(D) = \operatorname{def} \boldsymbol{r}(D).$$

Thus every nonzero vector in a one-dimensional rational vector space has the same defect.

The residue vector $[\boldsymbol{r}(D)]_p$ is defined as the vector of residues of $\boldsymbol{r}(D)p^{-e_p(\boldsymbol{r}(D))}$ modulo p. Thus if $e_p(r_i(D)) > e_p(\boldsymbol{r}(D))$, then $[r_i(D)]_p = 0$, even if $r_i(D) \neq 0$. If $\boldsymbol{r}(D)$ is expanded as a vector of formal Laurent series in p, then $[\boldsymbol{r}(D)]_p p^{e_p(\boldsymbol{r}(D))}$ is the first nonzero term in the expansion.

A *generator matrix* $G(D) = (g_{ij}(D))_{1 \le i \le k,\, 1 \le j \le n}$ for a linear, time-invariant convolutional code $\mathcal{C}$ is a $k \times n$ matrix of causal rational functions $g_{ij}(D) \in F(D)$, $1 \le i \le k$, $1 \le j \le n$, where F is a field and $F(D)$ is the field of rational functions in the delay operator D over F. (In behavioral system theory, the requirement that $G(D)$ be causal is sometimes dropped.) Without essential loss of generality, we assume from now on that $G(D)$ has full rank k. The *convolutional code* generated by $G(D)$ is

$$\mathcal{C} = \{\boldsymbol{u}(D)G(D) \mid \boldsymbol{u}(D) \in F((D))^k\}$$

where the input k-tuple $\boldsymbol{u}(D)$ ranges through the vector space of all k-tuples of formal Laurent series in D over F, denoted by $F((D))$. We also write $G(D)$ as

$$G(D) = \begin{pmatrix} \boldsymbol{g}_1(D) \\ \vdots \\ \boldsymbol{g}_k(D) \end{pmatrix}.$$

Then the set of codewords in $\mathcal{C}$ can be written as linear combinations

$$\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D) = \sum_i u_i(D)\boldsymbol{g}_i(D)$$

of the row vectors $\boldsymbol{g}_i(D)$ over $F((D))$. Two generator matrices are *equivalent* if they generate the same code $\mathcal{C}$.

Recall that a code sequence $\boldsymbol{v}(D)$ in $F((D))^n$ is rational if it is the formal Laurent series expansion in powers of D of a rational sequence in $F(D)^n$, which we shall continue to denote as $\boldsymbol{v}(D)$. The rational subcode $\mathcal{C}_r$ of $\mathcal{C}$ is defined to be the set of all rational code sequences $\boldsymbol{v}(D)$ in $\mathcal{C}$ (which includes all finite code sequences). Since we have assumed

that $G(D)$ has full rank and thus has a right inverse $G^{-1}(D)$, it is clear that $\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D)$ is rational if and only if $\boldsymbol{u}(D)$ is rational; i.e.,

$$\mathcal{C}_r = \{\boldsymbol{u}(D)G(D) \mid \boldsymbol{u}(D) \in F(D)^k\}.$$

Thus the rational subcode $\mathcal{C}_r$ is simply a vector space over $F(D)$ of dimension k, a subspace of $F(D)^n$.

We can regard two $k \times n$ equivalent generator matrices $G(D)$ and $G'(D)$ as two different bases for the same rational vector space $\mathcal{C}_r$. Thus they are equivalent if and only if there is a $k \times k$ rational nonsingular matrix $T(D)$ such that $G(D) = T(D)G'(D)$.

Let $G(D) = \{\boldsymbol{g}_i(D),\ 1 \le i \le k\}$ be a set of vectors $\boldsymbol{g}_i(D) \in F(D)^n$. In view of properties ii) and iii), for any vector

$$\boldsymbol{v}(D) = \sum_i u_i(D)\boldsymbol{g}_i(D)$$

and any $p \in \mathcal{P}^*$, we have

$$\begin{aligned} e_p(\boldsymbol{v}(D)) &\ge \min_i\{e_p(u_i(D)\boldsymbol{g}_i(D))\} && \text{(iii)} \\ &= \min_i\{e_p(u_i(D)) + e_p(\boldsymbol{g}_i(D))\}. && \text{(ii)} \end{aligned}$$

Monna [8] defines the set $G(D)$ to be *p-orthogonal* if equality holds for all rational k-tuples $\boldsymbol{u}(D)$; i.e., if for all $\boldsymbol{u}(D)$ in $F(D)^k$ we have

$$e_p(\boldsymbol{v}(D)) = \min_i\{e_p(u_i(D)) + e_p(\boldsymbol{g}_i(D))\}.$$

If $G(D)$ is p-orthogonal for all p in $\mathcal{P}^*$, then the set $G(D)$ is called *globally orthogonal.*

A $k \times n$ polynomial matrix $G(D) = \{\boldsymbol{g}_i(D) \in F[D]^n,\ 1 \le i \le k\}$ was said in [1] to have the *predictable degree property* (PDP) if for all $\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D)$, where $\boldsymbol{u}(D)$ and $\boldsymbol{v}(D)$ are polynomial vectors

$$\deg \boldsymbol{v}(D) = \max_i\{\deg u_i(D) + \deg \boldsymbol{g}_i(D)\}.$$

Equivalently, in the terminology we are using here, $G(D)$ has the PDP if for all $\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D)$

$$e_{D^{-1}}(\boldsymbol{v}(D)) = \min_i\{e_{D^{-1}}(u_i(D)) + e_{D^{-1}}(\boldsymbol{g}_i(D))\}$$

i.e., if $G(D)$ is D^{-1}-orthogonal. Hence, the PDP is naturally generalized as follows:

Definition 2: A rational matrix $G(D) \in F(D)^{k\times n}$ has the *predictable degree property* (PDP) if it is D^{-1}-orthogonal.

Definition 3: For any $p \in \mathcal{P}^*$, a rational matrix $G(D) \in F(D)^{k\times n}$ has the *predictable p-valuation property* (PVP$_p$) if it is p-orthogonal.

Definition 4: A rational matrix $G(D) \in F(D)^{k\times n}$ has the *global predictable valuation property* (GPVP) if it is globally orthogonal.

We shall see in Section VII below that the GPVP is an essential property of canonical generator matrices of convolutional codes.

In [2] the high-order coefficient vectors $[\boldsymbol{g}_i(D)]_h$ were defined as what we would call here the residue vectors $[\boldsymbol{g}_i(D)]_{D^{-1}}$, and it was shown that for the PDP to hold the matrix $[G(D)]_h$ consisting of these vectors must have full rank. The following natural generalization was proposed in [2, Appendix]:

Definition 5: Given a rational matrix $G(D) \in F(D)^{k\times n}$, its *p-residue matrix* $[G(D)]_p \in F[D]_p^{k\times n}$ is the $F[D]_p$-matrix whose ith row is the residue vector $[\boldsymbol{g}_i(D)]_p,\ 1 \le i \le k$.

The following theorem then gives a basic test for p-orthogonality:

Theorem 1: For any $p \in \mathcal{P}^*$, a rational matrix

$$G(D) = \{\boldsymbol{g}_i(D) \in F(D)^n,\ 1 \le i \le k\}$$

has the PVP$_p$ (is p-orthogonal) if and only if its p-residue matrix $[G(D)]_p$ has full rank k over $F[D]_p$.

Proof: In general, if $\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D)$, where $\boldsymbol{u}(D) = (u_1(D), \cdots, u_k(D))$ and $u_i(D) \in F(D),\ 1 \le i \le k$, then

$$e_p(\boldsymbol{v}(D)) \ge d$$

where

$$d = \min_i\{e_p(u_i(D)) + e_p(\boldsymbol{g}_i(D))\}.$$

Let $\mathcal{I}$ be the set of indices such that the minimum is achieved; i.e.,

$$\mathcal{I} = \{i \mid e_p(u_i(D)) + e_p(\boldsymbol{g}_i(D)) = d\}.$$

Then, if $\boldsymbol{v}(D) \ne \boldsymbol{0}$, the formal Laurent series expansion of $\boldsymbol{v}(D)$ in p may be written as

$$\boldsymbol{v}(D) = \boldsymbol{v}_d p^d + \boldsymbol{v}_{d+1}p^{d+1} + \cdots.$$

$G(D)$ is p-orthogonal if and only if for all $\boldsymbol{u}(D) \ne \boldsymbol{0}$ $e_p(\boldsymbol{v}(D)) = d$; i.e., $\boldsymbol{v}_d \ne \boldsymbol{0}$. We may write the formal Laurent series expansions of the nonzero $u_i(D)$ and of the $\boldsymbol{g}_i(D)$ as

$$\begin{aligned} u_i(D) &= [u_i(D)]_p p^{e_p(u_i(D))} + u_i^{(1)}(D), \quad 1 \le i \le k \\ \boldsymbol{g}_i(D) &= [\boldsymbol{g}_i(D)]_p p^{e_p(\boldsymbol{g}_i(D))} + \boldsymbol{g}_i^{(1)}(D), \quad 1 \le i \le k \end{aligned}$$

where for all i $[u_i(D)]_p \ne 0$, $e_p(u_i^{(1)}(D)) > e_p(u_i(D))$, $[\boldsymbol{g}_i(D)]_p \ne \boldsymbol{0}$, and $e_p(\boldsymbol{g}_i^{(1)}(D)) > e_p(\boldsymbol{g}_i(D))$. Then the lowest order coefficient of $\boldsymbol{v}(D)$ is given by

$$\boldsymbol{v}_d = \sum_{i\in\mathcal{I}}[u_i(D)]_p[\boldsymbol{g}_i(D)]_p.$$

If $\boldsymbol{v}_d = \boldsymbol{0}$, then the p-residue vectors $[\boldsymbol{g}_i(D)]_p$ are linearly dependent over $F[D]_p$ and $[G(D)]_p$ does not have full rank. Conversely, if $[G(D)]_p$ does not have full rank over $F[D]_p$, then there exists some nontrivial linear combination of rows that equals zero:

$$\sum_i u_i(D)[\boldsymbol{g}_i(D)]_p = 0$$

where $u_i(D) \in F[D]_p$; therefore, with the input sequence

$$(u_1(D)p^{-e_p(\boldsymbol{g}_1(D))},\ u_2(D)p^{-e_p(\boldsymbol{g}_2(D))}, \cdots, u_k(D)p^{-e_p(\boldsymbol{g}_k(D))})$$

we have $d = 0$ and $\boldsymbol{v}_d = \boldsymbol{0}$, so

$$e_p(\boldsymbol{v}(D)) > 0 = \min_i\{e_p(u_i(D)p^{e_p(\boldsymbol{g}_i(D))}) + e_p(\boldsymbol{g}_i(D))\}$$

which implies that $G(D)$ is not p-orthogonal. This completes the proof. □

This test involves only the determination of the rank of a $k \times n$ matrix $[G(D)]_p$ over the field $F[D]_p$, and is thus easy to carry out for any polynomial p of moderate degree.

Remark: The authors of [13] *define* a rational matrix $G(s)$ over the complex field to be "column-reduced at q" if $G(q)$ has full rank. This is clearly equivalent to the predictable valuation property at $p = s - q$, although this is not explicitly stated in [13].

The following simple result is needed for the proof of our general canonicality theorem (Theorem 13):

Lemma 2: For any p in $\mathcal{P}^*$, any $k \times n$ rational matrix $G(D)$, and any $k \times k$ diagonal rational matrix $A(D)$ with nonzero diagonal elements, $G'(D) = A(D)G(D)$ has the PVP_p if and only if $G(D)$ has the PVP_p.

Proof: Since $[G'(D)]_p = [A(D)]_p[G(D)]_p$, the theorem follows immediately from Theorem 1. □

We will now develop several equivalent conditions for a generator matrix $G(D) = \{\boldsymbol{g}_i(D),\ 1 \le i \le k\}$ to have the GPVP (be globally orthogonal). The essential ideas are as follows.

Each generator $\boldsymbol{g}_i(D)$ generates a one-dimensional rational subspace $\mathcal{C}_i = \{u(D)\boldsymbol{g}_i(D) \mid u(D) \in F(D)\}$ in which all nonzero vectors have a common defect, namely $\operatorname{def} \boldsymbol{g}_i(D)$. We shall define the *external defect* of $G(D)$ as the sum of these generator (subspace) defects

$$\operatorname{extdef} G(D) \stackrel{\text{def}}{=} \sum_{i=1}^{k} \operatorname{def} \boldsymbol{g}_i(D).$$

We shall show (Theorem 5) that $\operatorname{extdef} G(D)$ is minimized among all generator matrices $G(D)$ that generate a given code $\mathcal{C}$ if and only if $G(D)$ has the GPVP. Furthermore, all such matrices $G(D)$ have the same set of generator defects $\{\operatorname{def} \boldsymbol{g}_i(D),\ 1 \le i \le k\}$.

First, we state a technical lemma. Let $\mathcal{M}_k$ denote the set of all $k \times k$ submatrices $M(D)$ of $G(D)$. For all p in $\mathcal{P}^*$, define the p-valuation of $G(D)$ as [3]

$$e_p(G(D)) = \min_{M(D) \in \mathcal{M}_k} \{e_p(\det M(D))\}$$

and then correspondingly define the *internal defect* of $G(D)$ by

$$\operatorname{intdef} G(D) \stackrel{\text{def}}{=} -\sum_{p \in \mathcal{P}^*} e_p(G(D)) \deg p.$$

The following lemma states that $\operatorname{intdef} G(D)$ is invariant over all generator matrices for $\mathcal{C}$ [3]:

Lemma 3: If $G(D)$ and $G'(D)$ are equivalent generator matrices, then $\operatorname{intdef} G(D) = \operatorname{intdef} G'(D)$.

Proof: If $G(D)$ and $G'(D)$ are equivalent, then $G'(D) = T(D)G(D)$ for some $k \times k$ nonsingular rational matrix $T(D)$. The set $\mathcal{M}'_k$ of $k \times k$ submatrices of $G'(D)$ is then

$$\mathcal{M}'_k = \{T(D)M(D) \mid M(D) \in \mathcal{M}_k\}.$$

Since

$$\begin{aligned} e_p(\det(T(D)M(D))) &= e_p(\det T(D) \det M(D)) \\ &= e_p(\det T(D)) + e_p(\det M(D)) \end{aligned}$$

we have

$$e_p(G'(D)) = e_p(\det T(D)) + e_p(G(D))$$

and therefore, using the product formula (since $\det T(D) \neq 0$)

$$\begin{aligned} \operatorname{intdef} G'(D) &= -\sum_{p \in \mathcal{P}^*} e_p(\det T(D)) \deg p + \operatorname{intdef} G(D) \\ &= \operatorname{intdef} G(D). \end{aligned}$$ □

The next lemma shows that a linear combination of globally orthogonal vectors has at least as great a defect as any vector involved in the combination.

Lemma 4: If a rational matrix $G(D) = \{\boldsymbol{g}_i(D),\ 1 \le i \le k\}$ has the GPVP, then for every

$$\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D) = \sum_{i=1}^{k} u_i(D)\boldsymbol{g}_i(D)$$

$$\operatorname{def} \boldsymbol{v}(D) \ge \operatorname{def} \boldsymbol{g}_j(D)$$

for all j such that $u_j(D) \neq 0$.

Proof: If $G(D)$ has the GPVP, then we have

$$\begin{aligned} \operatorname{def} \boldsymbol{v}(D) &= -\sum_{p \in \mathcal{P}^*} e_p(\boldsymbol{v}(D)) \deg p \\ &= -\sum_{p \in \mathcal{P}^*} \min_i \{e_p(u_i(D)) + e_p(\boldsymbol{g}_i(D))\} \deg p \\ &\qquad \text{(GPVP)} \\ &\ge -\sum_{p \in \mathcal{P}^*} (e_p(u_j(D)) + e_p(\boldsymbol{g}_j(D))) \deg p \\ &\qquad \text{(for any } j, 1 \le j \le k) \\ &= -\sum_{p \in \mathcal{P}^*} e_p(\boldsymbol{g}_j(D)) \deg p \qquad \text{(if } u_j(D) \neq 0) \\ &= \operatorname{def} \boldsymbol{g}_j(D). \end{aligned}$$ □

Remark: This lemma suggests that a globally orthogonal generator matrix for $\mathcal{C}$ may be constructed by the following greedy algorithm. For the first generator $\boldsymbol{g}_1(D)$, choose any nonzero rational code sequence in $\mathcal{C}$ that has minimum defect. For the second generator $\boldsymbol{g}_2(D)$, choose any rational code sequence in $\mathcal{C}$ not in the subcode $\mathcal{C}_1$ generated by $\boldsymbol{g}_1(D)$ that has minimum defect. For $\boldsymbol{g}_3(D)$, choose any rational code sequence not in the space spanned by the sets $\mathcal{C}_1$ and $\mathcal{C}_2$, where $\mathcal{C}_2$ is the subcode generated by $\boldsymbol{g}_2(D)$ that has minimum defect, and so forth until k generators are chosen. (This algorithm in fact does work; this is straightforward to show from the fact that the similar algorithm given for polynomial matrices in Section IV works, by substituting "defect" for "degree.") This algorithm may have first appeared in the system theory literature in [13], although, as already mentioned, its roots have been traced back as far as Wedderburn [15] and Gantmacher [17].

This remark suggests that the generator defects $\operatorname{def} \boldsymbol{g}_i(D)$ of a globally orthogonal generator matrix for $\mathcal{C}$ are invariants of $\mathcal{C}$. The next theorem states this fundamental result and other related results.

Theorem 5: Let $G(D) = \{\boldsymbol{g}_i(D), 1 \le i \le k\}$ be a $k \times n$ rational matrix of rank k and $\mathcal{C}$ be the convolutional code generated by $G(D)$ over $F((D))$. Then the following statements are equivalent:

a) $G(D)$ has the GPVP.
b) For every $p \in \mathcal{P}^*$, $[G(D)]_p$ is of full rank k over $F[D]_p$.
c) For every $p \in \mathcal{P}^*$, $e_p([G(D)]_p) = 0$.
d) For every $p \in \mathcal{P}^*$, $e_p(G(D)) = \sum_{i=1}^{k} e_p(\boldsymbol{g}_i(D))$.
e) $\operatorname{extdef} G(D) = \operatorname{intdef} G(D)$.
f) $\operatorname{extdef} G(D)$ is minimal among all generator matrices which generate $\mathcal{C}$.
g) The set of defects $\{\operatorname{def}\boldsymbol{g}_1(D), \cdots, \operatorname{def}\boldsymbol{g}_k(D)\}$ is equal to the set of defects of any other generator matrix for C that has the GPVP.

Proof:

a) $\Leftrightarrow$ b) follows from Theorem 1.

b) $\Leftrightarrow$ c) is clear.

c) $\Leftrightarrow$ d). We again write the formal Laurent series expansions of the $\boldsymbol{g}_i(D)$'s as

$$\boldsymbol{g}_i(D) = p^{e_p(\boldsymbol{g}_i(D))}[\boldsymbol{g}_i(D)]_p + \boldsymbol{g}_i^{(1)}(D), \quad i = 1, \cdots, k$$

where $[\boldsymbol{g}_i(D)]_p \neq 0$ and $e_p(\boldsymbol{g}_i^{(1)}(D)) > e_p(\boldsymbol{g}_i(D))$ for all i. Then

$$G(D) = A(D)[G(D)]_p + G_1(D)$$

where $A(D) = \operatorname{diag}\{p^{e_p(\boldsymbol{g}_1(D))}, \cdots, p^{e_p(\boldsymbol{g}_k(D))}\}$ and the equivalence follows.

d) $\Leftrightarrow$ e). We have

$$\begin{aligned}\operatorname{extdef} G(D) &= \sum_{i=1}^{k} \operatorname{def}\boldsymbol{g}_i(D) \\ &= -\sum_{p\in\mathcal{P}^*}\sum_{i=1}^{k} e_p(\boldsymbol{g}_i(D)) \deg p.\end{aligned}$$

But for all p in $\mathcal{P}^*$, extending the proof of c) $\Leftrightarrow$ d), we have

$$\sum_{i=1}^{k} e_p(\boldsymbol{g}_i(D)) \le e_p(G(D))$$

with equality if and only if c) holds. Thus

$$\begin{aligned}\operatorname{extdef} G(D) &\ge -\sum_{p\in\mathcal{P}^*} e_p(G(D)) \deg p \\ &= \operatorname{intdef} G(D)\end{aligned}$$

with equality if and only if d) holds.

e) $\Leftrightarrow$ f). Follows immediately.

e) $\Leftrightarrow$ g). Let

$$G'(D) = \{\boldsymbol{g}_i'(D), 1 \le i \le k\}$$

be a generator matrix of $\mathcal{C}$ with the GPVP. Then there exists a nonsingular $k \times k$ matrix $A(D)$ such that $G(D) = A(D)G'(D)$. Order the generators $\{\boldsymbol{g}_i'(D), 1 \le i \le k\}$ so that $\operatorname{def}\boldsymbol{g}_j'(D) \ge \operatorname{def}\boldsymbol{g}_i'(D)$ if $j \ge i$. For each $\boldsymbol{g}_i(D)$ there is a unique expression

$$\boldsymbol{g}_i(D) = \sum_{j=1}^{n} a_{ij}\boldsymbol{g}_j'(D).$$

Let $l(i)$ be the largest j such that $a_{ij}(D) \neq 0$. Order $\boldsymbol{g}_1(D), \cdots, \boldsymbol{g}_n(D)$ such that $l(1) \le l(2) \le \cdots \le l(k)$. Thus $A(D)$ is in echelon form. Since $A(D)$ is nonsingular, $l(j) = j$, $j = 1, \cdots, k$. By Lemma 4

$$\operatorname{def}\boldsymbol{g}_i(D) \ge \operatorname{def}\boldsymbol{g}_{l(i)}'(D) = \operatorname{def}\boldsymbol{g}_i'(D).$$

Thus

$$\begin{aligned}\operatorname{extdef} G(D) &= \sum_{i=1}^{k} \operatorname{def}\boldsymbol{g}_i(D) \ge \sum_{i=1}^{k} \operatorname{def}\boldsymbol{g}_i'(D) \\ &= \operatorname{extdef} G'(D).\end{aligned}$$

Therefore

$$\operatorname{extdef} G(D) = \operatorname{extdef} G'(D) \Leftrightarrow \operatorname{def}\boldsymbol{g}_i(D) = \operatorname{def}\boldsymbol{g}_i'(D), \quad i = 1, \cdots, k$$

$\Leftrightarrow \{\operatorname{def}\boldsymbol{g}_1(D), \cdots, \operatorname{def}\boldsymbol{g}_k(D)\}$ is an invariant of $\mathcal{C}$. □

In summary, Theorem 5 shows that the external defect of $G(D)$ is lower-bounded by the internal defect of $G(D)$ (which by Lemma 3 is an invariant of $\mathcal{C}$), that the lower bound is met if and only if $G(D)$ has the GPVP, and that the set of generator defects $\{\operatorname{def}\boldsymbol{g}_i(D)\}$ of a globally orthogonal generator matrix $G(D)$ is an invariant of $\mathcal{C}$. Thus for any code $\mathcal{C}$ there exists a set of parameters $\{\nu_i(\mathcal{C}), 1 \le i \le k\}$ and $\nu(\mathcal{C}) = \sum_i \nu_i(\mathcal{C})$ such that $\nu_i(\mathcal{C}) = \operatorname{def}\boldsymbol{g}_i(D)$, $i = 1, \cdots, k$ and $\nu(\mathcal{C}) = \operatorname{extdef} G(D)$ if and only if $G(D)$ is a globally orthogonal generator matrix for $\mathcal{C}$. In the next section we shall see that $\{\nu_i(\mathcal{C}), 1 \le i \le k\}$ is the set of *constraint lengths* (*controller indices*) of the convolutional code (behavioral system) $\mathcal{C}$, and that $\nu(\mathcal{C})$ is the *overall constraint length* (minimal state space dimension) of $\mathcal{C}$.

Remark: The facts that the external defect is lower-bounded by the internal defect with equality iff the GPVP holds and that the indices $\nu_i(\mathcal{C})$ are invariants of $\mathcal{C}$ are the main results of [13].

IV. Canonical Polynomial Generator Matrices

After discussing realizations and defining minimality and canonicality, we construct a canonical polynomial generator matrix $G(D)$ for a convolutional code $\mathcal{C}$ using the elementary behavior-theoretic construction of [9].

A realization of a $k \times n$ generator matrix $G(D)$ (a rate $R = k/n$ *encoder*) is a k-input, n-output linear (over F) sequential circuit whose input/output (I/O) transfer function is $G(D)$. The complexity of an encoder is measured by the number of its memory elements, which is the dimension of its physical state space.

From the general principles of realization theory (see, e.g., [9]), the dimension of the physical state space of any encoder for $\mathcal{C}$ is lower-bounded by the dimension

$$\nu(\mathcal{C}) = \dim \Sigma(\mathcal{C})$$

of the *minimal state space* $\Sigma(\mathcal{C})$, which is the quotient space

$$\Sigma(\mathcal{C}) = \mathcal{C}/(\mathcal{C}_+ + \mathcal{C}_-)$$

where $\mathcal{C}_+ = \{\boldsymbol{v}(D) \text{ in } \mathcal{C} \mid \operatorname{del} \boldsymbol{v}(D) \geq 0\}$ is the subspace of all code sequences that "start" at time 0 or later, and $\mathcal{C}_- = \{\boldsymbol{v}(D) \text{ in } \mathcal{C} \mid \boldsymbol{v}(D) \text{ finite and } \deg \boldsymbol{v}(D) < 0\}$ is the subspace of all code sequences that "end" before time 0. We therefore define a *minimal encoder* as a generator matrix $G(D)$ that can be realized with $\nu(\mathcal{C})$ memory elements. Every code $\mathcal{C}$ has a minimal encoder [9].

The dimension $\mu(G(D))$ of a minimal realization of a given generator matrix $G(D)$ may be determined by a straightforward extension of this result, using ideas from behavioral system theory [19], [20]. Define the input/output (I/O) code generated by $G(D)$ to be the set

$$C^b = \{(\boldsymbol{u}(D), \boldsymbol{u}(D)G(D)) \mid \boldsymbol{u}(D) \text{ in } F((D))^k\}$$

of I/O pairs $\{\boldsymbol{u}(D), \boldsymbol{u}(D)G(D)\}$ generated by $G(D)$. Then $\mu(G(D))$ (sometimes called the McMillan degree of $G(D)$ [22]) is the dimension of the minimal state space of $\mathcal{C}^b$

$$\mu(G(D)) = \dim \Sigma(\mathcal{C}^b).$$

Then there exists a realization of $G(D)$ with physical state-space dimension $\mu(G(D))$, and no realization of lesser dimension.

Since $\mathcal{C}$ is embedded in $\mathcal{C}^b$, it is clear that

$$\mu(G(D)) \geq \nu(\mathcal{C})$$

with equality if and only if $G(D)$ is a minimal encoder for $\mathcal{C}$.

One method of realizing a rate $R = k/n$ generator matrix $G(D)$ is to realize each of its generators $\boldsymbol{g}_i(D)$ independently as a rate $R = 1/n$ encoder, with the output simply being the sum of the k component outputs. Such a realization is said to be in *controller canonical form*. If each generator $\boldsymbol{g}_i(D)$ is realized minimally with $\mu(\boldsymbol{g}_i(D))$ memory elements, then a controller canonical form realization of $G(D)$ requires a total of

$$\mu_{\text{ccf}}(G(D)) = \sum_{1 \leq i \leq k} \mu(\boldsymbol{g}_i(D))$$

memory elements. Thus

$$\mu_{\text{ccf}}(G(D)) \geq \mu(G(D)) \geq \nu(\mathcal{C}).$$

A generator matrix $G(D)$ will be called *canonical* if $\mu_{\text{ccf}}(G(D)) = \nu(\mathcal{C})$, which of course implies that $G(D)$ is minimal ($\mu(G(D)) = \nu(\mathcal{C})$).

If $G(D)$ is polynomial, then each generator $\boldsymbol{g}_i(D)$ has an obvious minimal realization with $\mu(\boldsymbol{g}_i(D)) = \deg \boldsymbol{g}_i(D)$ memory elements arranged in a feedbackfree shift register. Therefore

$$\mu_{\text{ccf}}(G(D)) = \sum_i \deg \boldsymbol{g}_i(D).$$

Hence, a polynomial generator matrix $G(D)$ is canonical if

$$\sum_{1 \leq i \leq k} \deg \boldsymbol{g}_i(D) = \nu(\mathcal{C}).$$

We now briefly describe a greedy construction of a canonical polynomial generator matrix $G(D)$ for a convolutional code $\mathcal{C}$, following the construction of [9] of a minimal encoder in controller canonical form for a general group code. (This construction generalizes straightforwardly to time-varying codes.) The construction produces a set $\{\boldsymbol{g}_i(D), 1 \leq i \leq k\}$ of "shortest independent polynomial generators" for $\mathcal{C}$ such that

$$\sum_{1 \leq i \leq k} \deg \boldsymbol{g}_i(D) = \nu(\mathcal{C}).$$

(The first construction of this kind was apparently that of Roos [21].)

First define $\mathcal{C}_{[j,j']}$ for $j' \geq j$ as the subcode of $\mathcal{C}$ consisting of all code sequences that are zero outside the interval $[j, j']$. $\mathcal{C}_{[j,j']}$ is a vector space over F with dimension denoted by $\dim \mathcal{C}_{[j,j']}$. If $\mathcal{C}$ is time-invariant, then $\mathcal{C}_{[j,j']}$ is a time shift of $\mathcal{C}_{[0,j'-j]}$, so we may restrict our attention to the subcodes $\mathcal{C}_{[0,j]}$, $j \geq 0$.

The construction goes as follows. Choose the first generator $\boldsymbol{g}_1(D)$ as a nonzero polynomial code sequence of least degree. Choose $\boldsymbol{g}_2(D)$ as a nonzero polynomial code sequence of least degree not in the rate $R = 1/n$ code $\mathcal{C}_1$ generated by $\boldsymbol{g}_1(D)$; choose $\boldsymbol{g}_3(D)$ as a nonzero polynomial code sequence of least degree not in the rate $R = 2/n$ code $\mathcal{C}_2$ generated by $\boldsymbol{g}_1(D)$ and $\boldsymbol{g}_2(D)$, and so forth, until a set $G(D) = \{\boldsymbol{g}_i(D), 1 \leq i \leq k\}$ of k generators has been chosen that generates $\mathcal{C}$.

It is easy to see that the degrees $\deg \boldsymbol{g}_i(D)$ are uniquely defined by $\mathcal{C}$ (cf. [2], [9]); in fact, they are the *constraint lengths* (or *controller indices*) of $\mathcal{C}$, denoted by

$$\nu_i(\mathcal{C}) \stackrel{\text{def}}{=} \deg \boldsymbol{g}_i(D), \quad 1 \leq i \leq k.$$

It is also clear (cf. [2], [9]) that

a) $\nu_1(\mathcal{C}) \leq \nu_2(\mathcal{C}) \leq \cdots \leq \nu_k(\mathcal{C})$;
b) $e_p(\boldsymbol{g}_i(D)) = 0$ for all p in $\mathcal{P}$ (else $\boldsymbol{g}_i(D)$ would not be the shortest polynomial generator not in $\mathcal{C}_{i-1}$), $1 \leq i \leq k$;
c) the dimension of $\mathcal{C}_{[0,j]}$ is the total number of time shifts $D^m \boldsymbol{g}_i(D)$ of generators $\boldsymbol{g}_i(D)$ such that $m \geq 0$ and $\deg D^m \boldsymbol{g}_i(D) = m + \nu_i(\mathcal{C}) \leq j$; i.e.,

$$\dim \mathcal{C}_{[0,j]} = \sum_{1 \leq i \leq k} \max\{0, j - \nu_i(\mathcal{C}) + 1\}.$$

Given the dimensions $\dim \mathcal{C}_{[0,j]}$ for all $j \geq 0$, this defines a system of equations that has a unique solution, viz., the degrees $\{\nu_i(\mathcal{C}), 1 \leq i \leq k\}$.

These results are summarized in the following theorem [2]:

Theorem 6: The degrees $\{\nu_1(\mathcal{C}) \leq \nu_2(\mathcal{C}) \leq \cdots \leq \nu_k(\mathcal{C})\}$ of a rate $R = k/n$, linear, time-invariant convolutional code (behavioral system) $\mathcal{C}$ are the unique integers that satisfy the following equations for all $j \geq 0$:

$$\dim \mathcal{C}_{[0,j]} = \sum_{1 \leq i \leq k} \max\{0, j - \nu_i(\mathcal{C}) + 1\}.$$

Furthermore, the minimal state-space dimension of $\mathcal{C}$ is

$$\nu(\mathcal{C}) = \sum_{1 \leq i \leq k} \nu_i(\mathcal{C}).$$

Finally, there exists a canonical polynomial generator matrix $G(D)$ for $\mathcal{C}$ with

$$\deg \boldsymbol{g}_i(D) = \nu_i(\mathcal{C}), \quad 1 \le i \le k$$

and a polynomial generator matrix for $\mathcal{C}$ is canonical if and only if

$$\deg \boldsymbol{g}_i(D) = \nu_i(\mathcal{C}), \quad 1 \le i \le k.$$

Moreover, the components of each $\boldsymbol{g}_i(D)$ constructed above are relatively prime. Therefore, $\operatorname{def} \boldsymbol{g}_i(D) = \deg \boldsymbol{g}_i(D)$. It follows immediately from Theorem 5 that for any generator matrix $G'(D) = \{\boldsymbol{g}'_i(D), 1 \le i \le k\}$ of $\mathcal{C}$ having the GPVP, $\operatorname{def} \boldsymbol{g}'_i(D) = \nu_i(\mathcal{C})$, $1 \le i \le k$.

Example 3 [4]: Let $\mathcal{C}$ be the rate $R = 2/3$ convolutional code over $\mathbb{F}_2$ generated by $G(D) = \{\boldsymbol{g}_1(D), \boldsymbol{g}_2(D)\}$, where

$$\boldsymbol{g}_1(D) = (1 + D,\ D,\ 1)$$
$$\boldsymbol{g}_2(D) = (1 + D^2 + D^3,\ 1 + D + D^2 + D^3,\ 0).$$

The shortest nonzero polynomial code sequence is $\boldsymbol{g}_1(D)$, so $\nu_1(\mathcal{C}) = 1$. The next shortest code sequence not dependent on $\boldsymbol{g}_1(D)$ has degree 2; e.g.,

$$\boldsymbol{g}'_2(D) = D^2\boldsymbol{g}_1(D) + \boldsymbol{g}_2(D) = (1,\ 1 + D + D^2,\ D^2)$$

so $\nu_2(\mathcal{C}) = 2$. The minimal state space of $\mathcal{C}$ thus has dimension $\nu(\mathcal{C}) = 3$. A canonical polynomial generator matrix for $\mathcal{C}$ is $G'(D) = \{\boldsymbol{g}_1(D), \boldsymbol{g}'_2(D)\}$.

The following theorem shows that canonical polynomial generator matrices must have the global predictable valuation property. (Example 4 of the next section shows that the GPVP alone does not assure that $G(D)$ is canonical.)

Theorem 7: A polynomial generator matrix $G(D) = \{\boldsymbol{g}_i(D), 1 \le i \le k\}$ is canonical if and only if $G(D)$ has the global predictable valuation property (GPVP) and $e_p(\boldsymbol{g}_i(D)) = 0$, $1 \le i \le k$, for all $p \in \mathcal{P}$.

Proof: Assume that $G(D)$ is canonical. Since $\boldsymbol{g}_i(D)$ is polynomial, $e_p(\boldsymbol{g}_i(D)) \ge 0$ for all $p \in \mathcal{P}$. Suppose that $e_p(\boldsymbol{g}_i(D)) > 0$ for some $p \in \mathcal{P}$. Then $\boldsymbol{g}_i(D)/p$ is a polynomial code sequence with degree less than $\deg \boldsymbol{g}_i(D) = \nu_i(\mathcal{C})$. Replacing $\boldsymbol{g}_i(D)$ in $G(D)$ by $\boldsymbol{g}_i(D)/p$ we obtain a generator matrix $G'(D)$ with $\mu_{\text{ccf}}(G'(D)) < \mu_{\text{ccf}}(G(D)) = \nu(\mathcal{C})$, a contradiction. Hence, $e_p(\boldsymbol{g}_i(D)) = 0$ for all $p \in \mathcal{P}$.

Suppose that the p-residue matrix $[G(D)]_p$ does not have full rank for some $p \in \mathcal{P}^*$. Then, as in the proof of Theorem 1, there is some nontrivial linear combination

$$\sum_i u_i(D)[\boldsymbol{g}_i(D)]_p = \mathbf{0}$$

where $u_i(D) \in F[D]_p$. By the proof of Theorem 1, $G(D)$ is not p-orthogonal. We conclude that if $G(D)$ is canonical, then $G(D)$ has the GPVP.

Conversely, let $G(D)$ be a polynomial generator matrix $G(D) = \{\boldsymbol{g}_i(D), 1 \le i \le k\}$ for $\mathcal{C}$ such that $e_p(\boldsymbol{g}_i(D)) = 0$ for all p in $\mathcal{P}$ and $G(D)$ has the GPVP. Let $\nu_i = \deg \boldsymbol{g}_i(D)$, $1 \le i \le k$. We shall show that the dimensions of the subcodes $\mathcal{C}_{[0,j]}$, $j \ge 0$, are determined by the parameters $\{\nu_i, 1 \le i \le k\}$ as in Theorem 6, which will imply that $\deg \boldsymbol{g}_i(D) = \nu_i(\mathcal{C})$, $1 \le i \le k$, and, thus, that $G(D)$ is canonical.

The subcode $\mathcal{C}_{[0,j]}$ is the set of all polynomial $\boldsymbol{v}(D)$ in $\mathcal{C}$ with $\deg \boldsymbol{v}(D) \le j$. Now, let $\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D)$. By the GPVP and the fact that $e_p(\boldsymbol{g}_i(D)) = 0$ for all p in $\mathcal{P}$, we have for all p in $\mathcal{P}$

$$\begin{aligned} e_p(\boldsymbol{v}(D)) &= \min_i \{e_p(u_i(D)) + e_p(\boldsymbol{g}_i(D))\} \\ &= \min_i \{e_p(u_i(D))\} \end{aligned}$$

so $e_p(\boldsymbol{v}(D)) \ge 0$ for all p in $\mathcal{P}$ if and only if for all i $e_p(u_i(D)) \ge 0$ for all p in $\mathcal{P}$; i.e., $\boldsymbol{v}(D)$ is polynomial if and only if for all i, $u_i(D)$ is polynomial. Furthermore, by the PDP, the degree of $\boldsymbol{v}(D)$ is given by

$$\deg \boldsymbol{v}(D) = \max_i \{\deg u_i(D) + \deg \boldsymbol{g}_i(D)\}.$$

Thus $\deg \boldsymbol{v}(D) \le j$ if and only if $\deg u_i(D) \le j - \nu_i$, $1 \le i \le k$. It follows that

$$\dim \mathcal{C}_{[0,j]} = \sum_{1 \le i \le k} \max\{0, j - \nu_i + 1\}.$$

Therefore, by Theorem 6, $G(D)$ is canonical. □

Remark: Theorem 7 foreshadows the general criterion for canonicality to be given in Theorem 13 below. To see that Theorem 7 is the specialization of Theorem 13 to polynomial generator matrices, note that if $\boldsymbol{g}_i(D)$ is polynomial then $e_p(\boldsymbol{g}_i(D)) \ge 0$ for all p in $\mathcal{P}$ and $e_{D^{-1}}(\boldsymbol{g}_i(D)) \le 0$. Therefore, the following statement is equivalent to Theorem 7: "A polynomial generator matrix $G(D)$ is canonical if and only if $G(D)$ has the GPVP and $e_p(\boldsymbol{g}_i(D)) \le 0$ for all p in $\mathcal{P}^*$."

Example 3 (cont.): Let $G(D)$ be as before. Since

$$[\boldsymbol{g}_1(D)]_{D^{-1}} = [\boldsymbol{g}_2(D)]_{D^{-1}} = (1, 1, 0)$$

$[G(D)]_{D^{-1}}$ does not have full rank. It follows that $G(D)$ does not have the predictable degree property and, hence, $G(D)$ is not canonical.

V. Minimal Generator Matrices

We have defined a rational generator matrix $G(D)$ to be minimal if the minimal state-space dimension $\mu(G(D))$ is equal to $\nu(\mathcal{C})$. In this section we give necessary and sufficient conditions for a rational generator matrix $G(D)$ to be minimal.

From the definition of minimality, $G(D)$ is minimal if and only if a minimal realization for the I/O code $\mathcal{C}^b = \{(\boldsymbol{u}(D), \boldsymbol{u}(D)G(D)\}$ has the same dimension as a minimal realization for the code $\mathcal{C} = \{\boldsymbol{u}(D)G(D)\}$.

From the construction of the previous section, a minimal realization for the I/O code $\mathcal{C}^b$ can be constructed from a set of "shortest independent generators" $\boldsymbol{g}^b_i(D)$ for $\mathcal{C}^b$. The degrees of such a set of generators must sum to $\mu(G(D))$, the dimension of a minimal realization of $\mathcal{C}^b$.

The idea of the following theorem is that a realization for $\mathcal{C}$ can be obtained from a minimal realization for $\mathcal{C}^b$, by projecting the "output" $(\boldsymbol{u}(D), \boldsymbol{u}(D)G(D))$ of $\mathcal{C}^b$ onto its second component, $\boldsymbol{u}(D)G(D)$. We denote this projection by the map P: $\mathcal{C}^b \to \mathcal{C}$. We shall show that this yields a minimal realization of $\mathcal{C}$ if and only if the restrictions of the

generators $\{\boldsymbol{g}_i^b(D)\}$ are a set $\{\boldsymbol{g}_i(D)\}$ of "shortest independent generators" for $\mathcal{C}$.

Define the *span* of a set of $f(D)$ of formal Laurent series in D as the interval from index of the first nonzero component of $f(D)$ to the index of the last nonzero component, if there is one, or to infinity otherwise. In other words, if $f(D)$ is rational, then

$$\operatorname{span} f(D) = \begin{cases} [\operatorname{del} f(D), \deg f(D)], & \text{if } f(D) \text{ is finite} \\ [\operatorname{del} f(D), \infty], & \text{if } f(D) \text{ is infinite.} \end{cases}$$

Then observe that

$$\operatorname{span} \boldsymbol{u}(D)G(D) \subseteq \operatorname{span} (\boldsymbol{u}(D), \boldsymbol{u}(D)G(D))$$

with equality if and only if

$$\operatorname{span} \boldsymbol{u}(D) \subseteq \operatorname{span} \boldsymbol{u}(D)G(D).$$

The following theorem shows that $G(D)$ is minimal if and only if these span inclusions hold with equality for all rational input sequences $\boldsymbol{u}(D)$.

Theorem 8: A rational generator matrix $G(D)$ is minimal $(\mu(G(D)) = \nu(\mathcal{C}))$ if and only if for all rational input sequences $\boldsymbol{u}(D)$, $\operatorname{span} \boldsymbol{u}(D) \subseteq \operatorname{span} \boldsymbol{u}(D)G(D)$.

Proof: Since we assumed that generator matrices have full rank, the projection map $P : \mathcal{C}^b \to \mathcal{C}$ is bijective. It follows that for all $j \geq 0$ the dimensions of the vector spaces $\mathcal{C}^b_{[0,j]}$ and $P(\mathcal{C}^b_{[0,j]})$ are equal. Now $P(\mathcal{C}^b_{[0,j]})$ is a subspace of $\mathcal{C}_{[0,j]}$, so

$$\dim \mathcal{C}^b_{[0,j]} = \dim P\left(\mathcal{C}^b_{[0,j]}\right) \leq \dim \mathcal{C}_{[0,j]}$$

for all $j \geq 0$, which from Theorem 6 implies that $\nu_i(\mathcal{C}) \leq \nu_i(\mathcal{C}^b)$, $1 \leq i \leq k$, with equality if and only if $\dim \mathcal{C}^b_{[0,j]} = \dim \mathcal{C}_{[0,j]}$ for all $j \geq 0$.

Now $\dim \mathcal{C}^b_{[0,j]} = \dim \mathcal{C}_{[0,j]}$ if and only if $P(\mathcal{C}^b_{[0,j]}) = \mathcal{C}_{[0,j]}$, or, equivalently, if and only if $\operatorname{span} \boldsymbol{u}(D)G(D) \subseteq [0,j]$ implies $\operatorname{span} (\boldsymbol{u}(D), \boldsymbol{u}(D)G(D)) \subseteq [0,j]$, or, again equivalently, if and only if $\operatorname{span} \boldsymbol{u}(D)G(D) \subseteq [0,j]$ implies $\operatorname{span} \boldsymbol{u}(D) \subseteq [0,j]$.

Thus if $\operatorname{span} \boldsymbol{u}(D) \subseteq \operatorname{span} \boldsymbol{u}(D)G(D)$ for all rational input sequences $\boldsymbol{u}(D)$, then $\dim \mathcal{C}^b_{[0,j]} = \dim \mathcal{C}_{[0,j]}$ for all $j \geq 0$, so by Theorem 6 $\nu_i(\mathcal{C}) = \nu_i(\mathcal{C}^b)$, $1 \leq i \leq k$, which implies that $\nu(\mathcal{C}) = \nu(\mathcal{C}^b)$. The realization of $G^b(D)$ in controller canonical form has $\nu(\mathcal{C}^b) = \nu(\mathcal{C})$ memory elements. Considering the k inputs and the last n outputs of this realization we get a realization of $G(D)$ in $\nu(\mathcal{C})$ memory elements, which implies that $\mu(G(D)) = \nu(\mathcal{C})$.

Conversely, suppose that $\operatorname{span} \boldsymbol{u}(D)G(D)$ does not cover $\operatorname{span} \boldsymbol{u}(D)$ for some rational $\boldsymbol{u}(D)$. If $\boldsymbol{u}(D)G(D)$ is infinite, then this can be true only if $\operatorname{del} \boldsymbol{u}(D)G(D) > \operatorname{del} \boldsymbol{u}(D)$. Multiplying through by a common denominator of the components of $\boldsymbol{u}(D)G(D)$, we find a finite $\boldsymbol{u}(D)G(D)$ such that $\operatorname{del} \boldsymbol{u}(D)G(D) > \operatorname{del} \boldsymbol{u}(D)$. If there is a finite $\boldsymbol{u}(D)G(D)$ such that $\operatorname{span} \boldsymbol{u}(D)G(D)$ does not cover $\operatorname{span} \boldsymbol{u}(D)$, then there is a polynomial $\boldsymbol{u}(D)G(D)$ with $\operatorname{del} \boldsymbol{u}(D)G(D) = 0$ and $\deg \boldsymbol{u}(D)G(D) = j$ (say), with the same property. Then, $\boldsymbol{u}(D)G(D) \in \mathcal{C}_{[0,j]}$ but $(\boldsymbol{u}(D), \boldsymbol{u}(D)G(D)) \notin \mathcal{C}^b_{[0,j]}$, so $\boldsymbol{u}(D)G(D) \notin P(\mathcal{C}^b_{[0,j]})$, which implies that $\dim \mathcal{C}^b_{[0,j]} < \dim \mathcal{C}_{[0,j]}$. Hence, from Theorem 6, $\nu(\mathcal{C}^b) > \nu(\mathcal{C})$. But $\nu(\mathcal{C}^b) = \mu(G(D))$; therefore, $G(D)$ is not minimal. □

Remark 1: A more general concept of a generator matrix exists in behavioral system theory, in which the input–output pairs $(\boldsymbol{u}(D), \boldsymbol{u}(D)G(D))$ need not be causally related—i.e., it is not required that $\operatorname{del} \boldsymbol{u}(D)G(D) \geq \operatorname{del} \boldsymbol{u}(D)$ for all $\boldsymbol{u}(D)$. Theorem 8 remains valid in this more general setting. If the causality condition $\operatorname{del} \boldsymbol{u}(D)G(D) \geq \operatorname{del} \boldsymbol{u}(D)$ is imposed, then Theorem 8 implies that if $G(D)$ is minimal then $\operatorname{del} \boldsymbol{u}(D)G(D) = \operatorname{del} \boldsymbol{u}(D)$ for all $\boldsymbol{u}(D)$.

Remark 2: Theorem 8 may be straightforwardly extended to the case in which $\mathcal{C}$ is linear but not time-invariant. Then a canonical set of generators for $\mathcal{C}$ consists of sets of shortest independent generators $\boldsymbol{g}_{[j,j']}(D) \in \mathcal{C}$. An input/output system is minimal if and only if it associates with each such generator an input sequence $\boldsymbol{u}_{[j,j']}(D)$ such that $\operatorname{span} \boldsymbol{u}_{[j,j']}(D) \subseteq \operatorname{span} \boldsymbol{g}_{[j,j']}(D) = [j,j']$.

Remark 3: Theorem 8 is closely related to the minimality criterion of [20] (and could have been derived from it). According to [20], a realization of a time-invariant group code can be nonminimal in only three ways: i) if there is an infinite nontrivial state sequence (not the zero state sequence) that produces an all-zero output sequence; ii) if there is a nontrivial transition from the zero state (not to the zero state) that produces a zero output; iii) if there is a nontrivial transition to the zero state (not from the zero state) that produces a zero output. Condition i) corresponds to a case in which there is an infinite input $\boldsymbol{u}(D)$ that produces a finite output $\boldsymbol{u}(D)G(D)$ (the "catastrophic" case); condition ii) corresponds to a case in which there is an input $\boldsymbol{u}(D)$ that produces an output $\boldsymbol{u}(D)G(D)$ with $\operatorname{del} \boldsymbol{u}(D)G(D) > \operatorname{del} \boldsymbol{u}(D)$; and condition iii) corresponds to a case in which there is a finite input $\boldsymbol{u}(D)$ that produces a finite output $\boldsymbol{u}(D)G(D)$ with $\deg \boldsymbol{u}(D)G(D) < \deg \boldsymbol{u}(D)$. It is easy to see that $\operatorname{span} \boldsymbol{u}(D)G(D)$ does not cover $\operatorname{span} \boldsymbol{u}(D)$ if and only if one of these three conditions is satisfied.

Example 4: The 1×1 generator matrix

$$G(D) = D$$

clearly has the GPVP but does not satisfy the condition of Theorem 8; hence, it is not minimal and thus not canonical. Indeed, if $G(D) = r(D)$ for any nonzero rational $r(D)$, then by the product formula $G(D)$ is minimal if and only if $r(D)$ is a unit.

Theorem 9: A rational generator matrix $G(D)$ is minimal if and only if any one of the following conditions holds:

a) For all $\boldsymbol{u}(D) \in F(D)^k$, if $\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D)$ then $\operatorname{span} \boldsymbol{u}(D) \subseteq \operatorname{span} \boldsymbol{v}(D)$.

b) Given a canonical polynomial generator matrix $G_{\text{cp}}(D) = \{\boldsymbol{g}_i(D), 1 \leq i \leq k\}$ for the code $\mathcal{C}$ generated by $G(D)$, if $\{\boldsymbol{u}_i(D), 1 \leq i \leq k\}$ are the input k-tuples such that $\boldsymbol{u}_i(D)G(D) = \boldsymbol{g}_i(D)$, $1 \leq i \leq k$, then $\boldsymbol{u}_i(D)$ is polynomial with

$$\deg \boldsymbol{u}_i(D) \leq \deg \boldsymbol{g}_i(D) = \nu_i(\mathcal{C}), \quad 1 \leq i \leq k.$$

c) For all $\boldsymbol{u}(D) \in F(D)^k$, if $\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D)$, then $e_p(\boldsymbol{v}(D)) \leq e_p(\boldsymbol{u}(D))$ for all $p \in \mathcal{P}^*$.

Proof: We prove the theorem by first observing that a) is only a reformulation of Theorem 8. Then we show that a) $\Rightarrow$ b) $\Rightarrow$ c) $\Rightarrow$ a).

a) $\Rightarrow$ b). If $(\boldsymbol{u}_i(D), \boldsymbol{g}_i(D))$ is an input/output pair, then by assumption span $(\boldsymbol{u}_i(D)) \subseteq$ span $(\boldsymbol{g}_i(D))$, which implies that $\boldsymbol{u}_i(D)$ is polynomial (so $e_p(\boldsymbol{u}_i(D)) \geq 0$ for all $p \in \mathcal{P}$) and

$$\deg \boldsymbol{u}_i(D) \leq \deg \boldsymbol{g}_i(D) = \nu_i(\mathcal{C}).$$

b) $\Rightarrow$ c). Since $G_{\rm cp}(D)$ is canonical polynomial it follows from Theorem 7 that $e_p(\boldsymbol{g}_i(D)) = 0$, $1 \leq i \leq k$, for all $p \in \mathcal{P}$. Since $\boldsymbol{u}_i(D)$, $1 \leq i \leq k$, are all polynomial, $e_p(\boldsymbol{u}_i(D)) \geq 0$ for all $p \in \mathcal{P}$. Therefore

$$e_p(\boldsymbol{u}_i(D)) \geq e_p(\boldsymbol{g}_i(D)), \quad 1 \leq i \leq k, \text{ for all } p \in \mathcal{P}.$$

For $p = D^{-1}$

$$\begin{aligned} e_{D^{-1}}(\boldsymbol{u}_i(D)) &= -\deg \boldsymbol{u}_i(D) \geq -\deg \boldsymbol{g}_i(D) \\ &= e_{D^{-1}}(\boldsymbol{g}_i(D)), \quad 1 \leq i \leq k. \end{aligned}$$

Hence

$$e_p(\boldsymbol{u}_i(D)) \geq e_p(\boldsymbol{g}_i(D)), \quad 1 \leq i \leq k, \text{ for all } p \in \mathcal{P}^*.$$

Let $\boldsymbol{u}(D) \in F(D)^k$ and $\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D)$. We can express

$$\boldsymbol{v}(D) = \sum_{i=1}^{k} u_i(D)\boldsymbol{g}_i(D)$$

where $u_i(D) \in F(D)$. Then

$$\boldsymbol{v}(D) = \sum_{i=1}^{k} u_i(D)\boldsymbol{u}_i(D)G(D) = \boldsymbol{u}(D)G(D).$$

Since $G(D)$ is of full rank k

$$\sum_{i=1}^{k} u_i(D)\boldsymbol{u}_i(D) = \boldsymbol{u}(D).$$

Therefore, for all $p \in \mathcal{P}^*$ we have

$$\begin{aligned} e_p(\boldsymbol{u}(D)) &\geq \min_{1 \leq i \leq k} \{e_p(u_i(D)) + e_p(\boldsymbol{u}_i(D))\} \\ &\geq \min_{1 \leq i \leq k} \{e_p(u_i(D)) + e_p(\boldsymbol{g}_i(D))\} = e_p(\boldsymbol{v}(D)) \end{aligned}$$

where the last equality follows from the fact that $G_{\rm cp}(D)$ is canonical and thus has the GPVP.

c) $\Rightarrow$ a). Now

$$\operatorname{del} \boldsymbol{u}(D) = e_D(\boldsymbol{u}(D)) \geq e_D(\boldsymbol{v}(D)) = \operatorname{del} \boldsymbol{v}(D)$$

so span $(\boldsymbol{u}(D))$ begins not earlier than span $(\boldsymbol{v}(D))$. If $\boldsymbol{v}(D)$ is infinite, then clearly span $(\boldsymbol{u}(D)) \subset$ span $(\boldsymbol{v}(D))$. If $\boldsymbol{v}(D)$ is finite, then for some $l \geq 0$, we can assume that $\boldsymbol{v}(D)D^l \in F[D]^n$. Therefore, $e_p(\boldsymbol{v}(D)D^l) \geq 0$ for all $p \in \mathcal{P}$. We have $\boldsymbol{v}(D)D^l = \boldsymbol{u}(D)D^lG(D)$. From b), we deduce that $e_p(\boldsymbol{u}(D)D^l) \geq 0$ for all $p \in \mathcal{P}$. Hence, $\boldsymbol{u}(D)D^l \in F[D]^k$. Then

$$\begin{aligned} \deg(\boldsymbol{u}(D)D^l) &= -e_{D^{-1}}(\boldsymbol{u}(D)D^l) \\ &\leq -e_{D^{-1}}(\boldsymbol{v}(D)D^l) = \deg(\boldsymbol{v}(D)D^l) \end{aligned}$$

so $\deg \boldsymbol{u}(D) \leq \deg \boldsymbol{v}(D)$. Hence, span $(\boldsymbol{u}(D)) \subseteq$ span $(\boldsymbol{v}(D))$, and the proof is completed. □

Example 3 (cont.): Let $G(D) = \{\boldsymbol{g}_1(D), \boldsymbol{g}_2(D)\}$ and $G'(D) = \{\boldsymbol{g}_1(D), \boldsymbol{g}'_2(D)\}$ be as before, with $G'(D)$ canonical. Then the input sequences $\boldsymbol{u}_1(D)$ and $\boldsymbol{u}_2(D)$ such that $\boldsymbol{u}_1(D)G(D) = \boldsymbol{g}_1(D)$ and $\boldsymbol{u}_2(D)G(D) = \boldsymbol{g}'_2(D)$ are

$$\begin{aligned} \boldsymbol{u}_1(D) &= (1, 0) \\ \boldsymbol{u}_2(D) &= (D^2, 1). \end{aligned}$$

Since $\boldsymbol{u}_1(D)$ and $\boldsymbol{u}_2(D)$ are polynomial, and

$$\begin{aligned} \deg \boldsymbol{u}_1(D) &= 0 < \nu_1(\mathcal{C}) = 1 \\ \deg \boldsymbol{u}_2(D) &= 2 = \nu_2(\mathcal{C}) \end{aligned}$$

it follows that $G(D)$ is minimal, although not canonical.

Remark 1: As we shall see in Section VIII, condition b) is essentially a condition that $G(D)$ must be globally invertible.

Remark 2: Condition c) can be characterized as a zero-free property; the p-valuation of an output sequence cannot be greater than the p-valuation of the input sequence that generated it, for any p in $\mathcal{P}^*$.

Remark 3: Theorem 9 yields a simple proof of the well-known fact [1] that a systematic generator matrix $G(D)$ is minimal. A systematic generator matrix is one that embeds the input k-tuple $\boldsymbol{u}(D)$ in the output n-tuple $\boldsymbol{v}(D)$, so that we may write $\boldsymbol{v}(D) = (\boldsymbol{u}(D), \boldsymbol{p}(D))$, where $\boldsymbol{p}(D)$ is a "parity check" $(n-k)$-tuple. It is obvious that span $\boldsymbol{u}(D) \subseteq$ span $\boldsymbol{v}(D)$; or, equally, that $e_p(\boldsymbol{v}(D)) \leq e_p(\boldsymbol{u}(D))$, since $e_p(\boldsymbol{v}(D)) = \min\{e_p(\boldsymbol{u}(D)), e_p(\boldsymbol{p}(D))\} \leq e_p(\boldsymbol{u}(D))$.

VI. Canonical Rational Rate $R = 1/n$ Generator Matrices

We now derive the simple and well-known conditions for a rate $R = 1/n$ generator matrix to be minimal [1], [22].

A generator matrix for a rate $R = 1/n$ convolutional code $\mathcal{C}$ consists of a single rational generator $\boldsymbol{g}(D)$; then

$$\mathcal{C} = \{u(D)\boldsymbol{g}(D) \mid u(D) \in F((D))\}.$$

By our definition of canonicality, a rate $R = 1/n$ generator matrix $\boldsymbol{g}(D)$ is canonical if and only if it is minimal.

We may write $\boldsymbol{g}(D)$ uniquely as

$$\boldsymbol{g}(D) = \left(\prod_{p \in \mathcal{P}} p^{e_p(\boldsymbol{g}(D))} \right) \boldsymbol{g}'(D) = (n(D)/d(D))\boldsymbol{g}'(D)$$

where

a) $n(D)$ is the polynomial product $\prod_p p^{e_p(\boldsymbol{g}(D))}$ over those $p \in \mathcal{P}$ such that $e_p(\boldsymbol{g}(D)) > 0$;

b) $d(D)$ is the polynomial product $\prod_p p^{-e_p(\boldsymbol{g}(D))}$ over those $p \in \mathcal{P}$ such that $e_p(\boldsymbol{g}(D)) < 0$;

c) $\boldsymbol{g}'(D)$ is a polynomial n-tuple with $e_p(\boldsymbol{g}'(D)) = 0$ for all $p \in \mathcal{P}$.

Clearly, $\boldsymbol{g}'(D)$ is uniquely defined up to multiplication by units.

Theorem 7 implies the following well-known theorem [1]:

Theorem 10: The minimal state-space dimension of a rate $R = 1/n$ code $\mathcal{C}$ generated by a single generator $\boldsymbol{g}(D)$ is $\nu(\mathcal{C}) = \operatorname{def} \boldsymbol{g}(D)$. The generator $\boldsymbol{g}'(D) = (d(D)/n(D))\boldsymbol{g}(D)$ defined above is a canonical polynomial generator matrix for $\mathcal{C}$ with $\deg \boldsymbol{g}'(D) = \nu(\mathcal{C})$.

Proof: By Theorem 7, $\boldsymbol{g}'(D)$ is a canonical polynomial generator matrix for $\mathcal{C}$, since $e_p(\boldsymbol{g}'(D)) = 0$ for all p in $\mathcal{P}$ and the GPVP holds trivially. Therefore, $\nu(\mathcal{C}) = \deg \boldsymbol{g}'(D)$. Furthermore

$$\deg \boldsymbol{g}'(D) = \operatorname{def} \boldsymbol{g}'(D) = -e_{D^{-1}}(\boldsymbol{g}'(D))$$

since $e_p(\boldsymbol{g}'(D)) = 0$ for all p in $\mathcal{P}$, and

$$\begin{aligned}\operatorname{def} \boldsymbol{g}(D) &= \operatorname{def}(n(D)/d(D)) + \operatorname{def} \boldsymbol{g}'(D) \\ &= \operatorname{def} \boldsymbol{g}'(D) = \nu(\mathcal{C}).\end{aligned}$$

□

It then follows from our minimality test that:

Theorem 11: A rational rate $R = 1/n$ generator matrix $\boldsymbol{g}(D) = (n(D)/d(D))\boldsymbol{g}'(D)$ is minimal (and thus canonical) if and only if $n(D) = 1$ and $\deg d(D) \leq \deg \boldsymbol{g}'(D)$.

Proof: The input to $\boldsymbol{g}(D)$ that generates $\boldsymbol{g}'(D)$ is $u(D) = d(D)/n(D)$; $u(D)$ is polynomial if and only if $n(D) = 1$ and, then, $\operatorname{span} u(D) \subseteq \operatorname{span} \boldsymbol{g}'(D)$ if and only if $\deg d(D) \leq \deg \boldsymbol{g}'(D)$. □

Corollary 12: A rational rate $R = 1/n$ generator matrix $\boldsymbol{g}(D)$ is minimal if and only if $e_p(\boldsymbol{g}(D)) \leq 0$ for all $p \in \mathcal{P}^*$. □

Proof: Again write $\boldsymbol{g}(D) = (n(D)/d(D))\boldsymbol{g}'(D)$, where $e_p(\boldsymbol{g}'(D)) = 0$ for $p \in \mathcal{P}$. Assume that $\boldsymbol{g}(D)$ is minimal, then by Theorem 11 $n(D) = 1$ and $\deg d(D) \leq \deg \boldsymbol{g}'(D)$. Since $e_{D^{-1}}(\boldsymbol{g}'(D)) = -\deg \boldsymbol{g}'(D)$, we have

$$\begin{aligned}e_p(\boldsymbol{g}(D)) &= e_p(\boldsymbol{g}'(D)) + e_p(n(D)) - e_p(d(D) \\ &\leq 0 \text{ for } p \in \mathcal{P} \\ e_{D^{-1}}(\boldsymbol{g}(D)) &= e_{D^{-1}}(\boldsymbol{g}'(D)) + e_{D^{-1}}(n(D)) - e_{D^{-1}}(d(D)) \\ &\leq 0.\end{aligned}$$

Conversely, if $e_p(\boldsymbol{g}(D)) \leq 0$ for all $p \in \mathcal{P}^*$, then

$$e_p(n(D)/d(D)) = e_p(\boldsymbol{g}(D)) - e_p(\boldsymbol{g}'(D)) \leq -e_p(\boldsymbol{g}'(D)).$$

Hence

$$e_p(n(D)/d(D)) \leq 0 \text{ for } p \in \mathcal{P}$$

which implies $n(D) = 1$, and

$$e_{D^{-1}}(n(D)/d(D)) \leq \deg \boldsymbol{g}'(D)$$

which implies $\deg d(D) \leq \deg \boldsymbol{g}'(D)$. □

Corollary 12 foreshadows Theorem 13 in the next section.

VII. Canonical Rational Generator Matrices

We recall that a rational generator matrix $G(D) = \{\boldsymbol{g}_i(D), 1 \leq i \leq k\}$ is canonical if $\mu_{\text{ccf}}(G(D)) = \nu(\mathcal{C})$, where

$$\mu_{\text{ccf}}(G(D)) = \sum_{1 \leq i \leq k} \mu(\boldsymbol{g}_i(D))$$

is the sum of the dimensions $\mu(\boldsymbol{g}_i(D))$ of minimal independent realizations of each of the generators $\boldsymbol{g}_i(D)$, and $\nu(\mathcal{C})$ is the minimal state-space dimension of the code $\mathcal{C}$. A canonical generator matrix thus specifies a minimal realization in controller canonical form.

We shall give canonicality tests for a rational generator matrix $G(D)$ that generalizes Theorem 7 (for polynomial generator matrices) and Corollary 12 (for rate $R = 1/n$ rational generator matrices).

The equivalence of statements a) and c) in the following theorem, first stated in [5], corrects the assertion of [2] and [3] that $G(D)$ is canonical if and only if it has the GPVP.

Theorem 13: Let $G(D)$ be a $k \times n$ rational matrix of rank k and $\mathcal{C}$ be the convolutional code generated by $G(D)$ over $F((D))$. Then the following statements are equivalent:

a) $G(D)$ is canonical.

b) $e_p(\boldsymbol{g}_i(D)) \leq 0$, $1 \leq i \leq k$, for all $p \in \mathcal{P}^*$, and $\operatorname{extdef} G(D) = \nu(\mathcal{C})$.

c) $e_p(\boldsymbol{g}_i(D)) \leq 0$, $1 \leq i \leq k$, for all $p \in \mathcal{P}^*$, and $G(D)$ has the GPVP.

Proof:

a) $\Leftrightarrow$ b). Clearly, $G(D)$ is not canonical if a generator $\boldsymbol{g}_i(D)$ can be replaced by a generator $\boldsymbol{g}_i'(D)$ such that $\mu(\boldsymbol{g}_i'(D)) < \mu(\boldsymbol{g}_i(D))$, while generating the same code. If $e_p(\boldsymbol{g}_i(D)) > 0$ for some i and for some p in $\mathcal{P}^*$, then by Corollary 12 $\boldsymbol{g}_i(D)$ is not minimal for the rate $R = 1/n$ code that it generates, and can be replaced by a minimal generator $\boldsymbol{g}_i'(D)$ for the same code, so $G(D)$ cannot be canonical. Hence, if $G(D)$ is canonical, then $\boldsymbol{g}_i(D)$ is minimal as a rate $R = 1/n$ generator, and by Theorem 10

$$\mu(\boldsymbol{g}_i(D)) = \operatorname{def} \boldsymbol{g}_i(D).$$

It follows that if $G(D)$ is canonical, then

$$\nu(\mathcal{C}) = \mu_{\text{ccf}}(G(D)) = \sum_{1 \leq i \leq k} \operatorname{def} \boldsymbol{g}_i(D).$$

Conversely, if $e_p(\boldsymbol{g}_i(D)) \leq 0$, $1 \leq i \leq k$, for all p in $\mathcal{P}^*$, then by Corollary 12 and Theorem 10 each generator $\boldsymbol{g}_i(D)$ can be realized with $\operatorname{def} \boldsymbol{g}_i(D)$ memory elements. If moreover

$$\sum_{i=1}^{k} \operatorname{def} \boldsymbol{g}_i(D) = \nu(\mathcal{C})$$

then $G(D)$ can be realized in controller canonical form with $\nu(\mathcal{C})$ memory elements, so $G(D)$ is canonical.

b) $\Leftrightarrow$ c). Let the rows of $G(D)$ be $\boldsymbol{g}_1(D), \cdots, \boldsymbol{g}_k(D)$. As in Section VI, we may write each generator uniquely as

$$\boldsymbol{g}_i(D) = (n_i(D)/d_i(D))\boldsymbol{g}_i'(D)$$

where $\boldsymbol{g}_i'(D)$ is polynomial with $e_p(\boldsymbol{g}_i'(D)) = 0$ for p in $\mathcal{P}$ and

$$\deg \boldsymbol{g}_i'(D) = \operatorname{def} \boldsymbol{g}_i'(D) = \operatorname{def} \boldsymbol{g}_i(D).$$

If $A(D)$ is the $k \times k$ diagonal matrix with nonzero diagonal elements $d_i(D)/n_i(D)$, then

$$G'(D) = A(D)G(D).$$

It follows from Lemma 2 that $G'(D)$ has the GPVP if and only if $G(D)$ has the GPVP. Furthermore, the external defect of $G'(D)$, namely the sum of the degrees of its generators, is equal to the external defect of $G(D)$.

Suppose that $e_p(\boldsymbol{g}_i(D)) \leq 0$, $1 \leq i \leq k$, for all p in $\mathcal{P}^*$, and $\operatorname{extdef} G(D) = \nu(\mathcal{C})$. The polynomial matrix $G'(D) = A(D)G(D)$ then has $e_p(\boldsymbol{g}'_i(D)) = 0$, $1 \leq i \leq k$, for all p in $\mathcal{P}$, and $\deg \boldsymbol{g}'_i(D) = \operatorname{def} \boldsymbol{g}_i(D)$. It follows that

$$\begin{aligned}\operatorname{extdef} G'(D) &= \sum_{1\leq i\leq k} \operatorname{def} \boldsymbol{g}'_i(D) \\ &= \sum_{1\leq i\leq k} \deg \boldsymbol{g}_i(D) = \nu(\mathcal{C})\end{aligned}$$

and therefore $G'(D)$ is canonical. By Theorem 7, $G'(D)$ has the GPVP, and therefore $G(D)$ must have the GPVP.

Conversely, if $e_p(\boldsymbol{g}_i(D)) \leq 0$, $1 \leq i \leq k$, for all p in $\mathcal{P}^*$, and $G(D)$ has the GPVP, then $G'(D)$ has the GPVP by Lemma 2, so by Theorem 7 $G'(D)$ is canonical, and

$$\nu(\mathcal{C}) = \sum_{1\leq i\leq k} \deg \boldsymbol{g}'_i(D).$$

Thus the external defect of $G(D)$ is equal to $\nu(\mathcal{C})$, so by c) $G(D)$ is canonical. □

Remark 1: Since a canonical generator matrix $G(D)$ has defects $\{\operatorname{def} \boldsymbol{g}_i(D),\ 1 \leq i \leq k\}$ equal to those of an equivalent canonical polynomial generator matrix $G_{\mathrm{cp}}(D)$, it is clear that $G_{\mathrm{cp}}(D)$ can be obtained from $G(D)$ simply by multiplying each generator $\boldsymbol{g}_i(D)$ by

$$d_i(D) = \prod_{p\in\mathcal{P}} p^{-e_p(\boldsymbol{g}_i(D))}$$

(the least common multiple of the denominators of $\{g_{ij}(D),\ 1 \leq j \leq n\}$). More generally, since any generator matrix $G(D)$ with the GPVP has defects equal to those of an equivalent canonical polynomial generator matrix $G_{\mathrm{cp}}(D)$, $G_{\mathrm{cp}}(D)$ can be obtained from $G(D)$ by multiplying each generator $\boldsymbol{g}_i(D)$ by

$$d_i(D)/n_i(D) = \prod_{p\in\mathcal{P}} p^{-e_p(\boldsymbol{g}_i(D))}$$

since if $\boldsymbol{g}'_i(D) = (d_i(D)/n_i(D))\boldsymbol{g}_i(D)$, then $e_p(\boldsymbol{g}'_i(D)) = 0$ for all $p \in \mathcal{P}$ and

$$\deg \boldsymbol{g}'_i(D) = \operatorname{def} \boldsymbol{g}'_i(D) = \operatorname{def} \boldsymbol{g}_i(D) = \nu_i(\mathcal{C}).$$

Remark 2: Theorem 13 shows that canonicality is the intersection of two independent properties, global orthogonality (the GPVP), and minimality. Global orthogonality ensures

$$\sum_i \operatorname{def} \boldsymbol{g}_i(D) = \nu(\mathcal{C})$$

but as Example 4 ($G(D) = D$) shows, it does not ensure that each $\boldsymbol{g}_i(D)$ can be realized with $\operatorname{def} \boldsymbol{g}_i(D)$ memory elements. The minimality condition of Corollary 12, $e_p(\boldsymbol{g}_i(D)) \leq 0$ for all p in $\mathcal{P}^*$, ensures that each generator $\boldsymbol{g}_i(D)$ can be realized with $\operatorname{def} \boldsymbol{g}_i(D)$ memory elements and thus is a minimal encoder for the one-dimensional code that it generates.

VIII. Minimality and Invertibility via the IFT

In this section we will use the invariant factor theorem (IFT) [23] with respect to both $F[D]$ and $F[D^{-1}]$ to show the equivalence of minimality and global invertibility.

First we state the extended invariant factor theorem (IFT), sometimes called the Smith–McMillan canonical form [22]:

Theorem 14 (Extended Invariant Factor Theorem): Let $G(D)$ be a full-rank $k \times n$ rational matrix, where $k \leq n$. Then $G(D)$ may be written as follows:

$$G(D) = A(D)\Gamma(D)B(D)$$

where $A(D)$ and $B(D)$ are, respectively, $k \times k$ and $n \times n$ matrices with unit determinants, and where $\Gamma(D)$ is a diagonal matrix with diagonal elements $\gamma_i(D)$, $1 \leq i \leq k$, called the *invariant factors* of $G(D)$ relative to the ring $F[D]$. The invariant factors are uniquely determined by $G(D)$ as follows:

$$\gamma_i(D) = \Delta_i(D)/\Delta_{i-1}(D)$$

where $\Delta_0(D) = 1$ by convention and

$$\Delta_i(D) = \prod_{p\in\mathcal{P}} p^{\min\{e_p(\det M_i(D))|M_i(D)\in\mathcal{M}_i\}}$$

where $\mathcal{M}_i$ is the set of $i \times i$ submatrices of $G(D)$, $1 \leq i \leq k$. Consequently

$$\prod_{i=1}^{k} \gamma_i(D) = \Delta_k(D)$$

$$\sum_{i=1}^{k} e_p(\gamma_i(D)) = e_p(\Delta_k(D)),\ p \in \mathcal{P}.$$

For all p in $\mathcal{P}$, the invariant factors satisfy the divisibility property

$$e_p(\gamma_i(D)) \leq e_p(\gamma_{i+1}(D)), \quad 1 \leq i < k.$$

It is easy to show that if $G(D)$ is regarded as a matrix over $F(D^{-1})$, then the invariant factors $\widetilde{\gamma}_i(D^{-1})$ of $G(D)$ with respect to $F[D^{-1}]$ have the same p-valuations as the invariant factors $\gamma_i(D)$ for all p in $\mathcal{P}$ except for D. Therefore, it makes sense to define the p-valuations of the invariant factors of $G(D)$ for all p in $\mathcal{P}^*$ and all i by

$$\begin{aligned}\gamma_{D,i} &= e_D(\gamma_i(D)), && \text{if } p = D \\ \gamma_{D^{-1},i} &= e_{D^{-1}}(\widetilde{\gamma}_i(D^{-1})), && \text{if } p = D^{-1} \\ \gamma_{p,i} &= e_p(\gamma_i(D)) = e_p(\widetilde{\gamma}_i(D^{-1})), && \text{otherwise.}\end{aligned}$$

If we define $\widetilde{\Delta}_0(D^{-1}) = 1$ and

$$\widetilde{\Delta}_i(D^{-1}) = \prod_{p\in\mathcal{P}^*\setminus\{D\}} p^{\min\{e_p(\det M_i(D^{-1}))|M_i(D^{-1})\in\mathcal{M}_i\}}$$

then

$$\widetilde{\gamma}_i(D^{-1}) = \widetilde{\Delta}_i(D^{-1})/\widetilde{\Delta}_{i-1}(D^{-1}).$$

To simplify the computation of these p-valuations for small generator matrices we define $\delta_{p,0} = 0$ for all p in $\mathcal{P}^*$ and

$$\begin{aligned}\delta_{p,i} &= e_p(\Delta_i(D)), \quad p \in \mathcal{P} \\ \delta_{D^{-1},i} &= e_{D^{-1}}(\widetilde{\Delta}_i(D^{-1}))\end{aligned}$$

and then we have for all p in $\mathcal{P}^*$ and $1 \le i \le k$

$$\gamma_{p,i} = \delta_{p,i} - \delta_{p,i-1},$$

which implies

$$\delta_{p,k} = \sum_{i=1}^{k} \gamma_{p,i}.$$

Remark: We can now recognize that for $p \in \mathcal{P}$

$$\begin{aligned} e_p(G(D)) &= \min\{e_p(\det M_k(D)) \mid M_k(D) \in \mathcal{M}_k\} \\ &= e_p(\Delta_k(D)) = \delta_{p,k} \end{aligned}$$

and

$$\begin{aligned} e_{D^{-1}}(G(D)) &= \min\{e_{D^{-1}}(\det M_k(D)) \mid M_k(D) \in \mathcal{M}_k\} \\ &= e_{D^{-1}}(\widetilde{\Delta}_k(D^{-1})) = \delta_{D^{-1},k}. \end{aligned}$$

Thus the internal defect can be computed directly from the p-valuations of the invariant factors of $G(D)$ by

$$\begin{aligned} \operatorname{intdef} G(D) &= -\sum_{p \in \mathcal{P}^*} \delta_{p,k} \deg p \\ &= -\sum_{p \in \mathcal{P}^*} \left(\sum_{i=1}^{k} \gamma_{p,i} \right) \deg p. \end{aligned}$$

Now we have a theorem that shows the equivalence of a criterion for minimality in terms of invertibility [24], [4] to the global zero-freedom of the invariant factors of $G(D)$.

Theorem 15: For a full-rank rational $k \times n$ matrix $G(D)$, the following are equivalent:

a) $G(D)$ is minimal.
b) $\gamma_{p,k} \le 0$ for all p in $\mathcal{P}^*$.
c) $G(D)$ has an $F[D]$-inverse and an $F[D^{-1}]$-inverse.

Proof: We prove a) $\Rightarrow$ b) $\Rightarrow$ c) $\Rightarrow$ a).

a) $\Rightarrow$ b). If $G(D)$ is minimal, then Theorem 8 implies that a polynomial output sequence $\boldsymbol{u}(D)G(D)$ must be generated by a polynomial input sequence $\boldsymbol{u}(D)$, and an antipolynomial output sequence $\boldsymbol{u}(D^{-1})G(D^{-1})$ must be generated by an antipolynomial input sequence $\boldsymbol{u}(D^{-1})$. Let $G(D) = A(D)\Gamma(D)B(D)$ be an invariant factor decomposition of $G(D)$; then $G^{-1}(D) = B^{-1}(D)\Gamma^{-1}(D)A^{-1}(D)$ is a right inverse of $G(D)$, where $A^{-1}(D)$ and $B^{-1}(D)$ are polynomial since $A(D)$ and $B(D)$ have unit determinants. Suppose $\gamma_{p,k} > 0$ for some p in $\mathcal{P}$. Then the input $\boldsymbol{u}(D) = (0, 0, \cdots, 1/p)A^{-1}(D)$ is nonpolynomial (since $\boldsymbol{u}(D)A(D)$ is nonpolynomial), but $\boldsymbol{u}(D)G(D)$ is polynomial, contradiction. Using the IFT with respect to $F[D^{-1}]$, we can show a similar contradiction if $\gamma_{D^{-1},k} > 0$.

b) $\Rightarrow$ c). For p in $\mathcal{P}$, if $\gamma_{p,k} \le 0$, then $\gamma_{p,i} \le 0$, since by the IFT $\gamma_{p,i} \le \gamma_{p,k}$ for $i \le k$. Hence, if $\gamma_{p,k} \le 0$ for all p in $\mathcal{P}$, $\Gamma^{-1}(D)$ is polynomial, and then $G^{-1}(D) = B^{-1}(D)\Gamma^{-1}(D)A^{-1}(D)$ is the desired polynomial right inverse of $G(D)$. Similarly, if also $\gamma_{D^{-1},k} \le 0$, then $\Gamma^{-1}(D^{-1})$ is antipolynomial, and the IFT with respect to $F[D^{-1}]$ yields an $F[D^{-1}]$-inverse of $G(D)$.

c) $\Rightarrow$ a). Assume that $G(D)$ has an $F[D]$-inverse $G^{-1}(D)$, and let $\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D)$ be any sequence in $\mathcal{C}$. Then $\boldsymbol{u}(D) = \boldsymbol{v}(D)G^{-1}(D)$, so

i) $\boldsymbol{u}(D)$ is finite if $\boldsymbol{v}(D)$ is finite;
ii) $\operatorname{del} \boldsymbol{u}(D) \ge \operatorname{del} \boldsymbol{v}(D)$.

Similarly, if $G(D)$ has an $F[D^{-1}]$-inverse, then

iii) $\deg \boldsymbol{u}(D) \le \deg \boldsymbol{v}(D)$.

Conditions i)– iii) imply that $\operatorname{span} \boldsymbol{u}(D) \subseteq \operatorname{span} \boldsymbol{v}(D)$, which by Theorem 8 implies that $G(D)$ is minimal. □

Remark 1: The conditions i)– iii) of the last part of the proof may be related to the conditions of the minimality test of [20] as follows. Condition ii) holds for all $\boldsymbol{v}(D)$ in $\mathcal{C}$ if and only if $\gamma_{p,k} \le 0$ for $p = D$, and is equivalent to the condition that no nontrivial zero-output state transition starting from the zero state occurs in the minimal state realization of $G(D)$ when it is used as a state realization of $\mathcal{C}$. Similarly, condition iii) holds for all $\boldsymbol{v}(D)$ if and only if $\gamma_{p,k} \le 0$ for $p = D^{-1}$, and is equivalent to the condition that there is no nontrivial zero-output state transition ending in the zero state. Finally, condition i) holds for all $\boldsymbol{v}(D)$ if and only if $\gamma_{p,k} \le 0$ for all other $p \in \mathcal{P}^*$, and is equivalent to the condition that there is no nontrivial infinite zero-output state path (the "noncatastrophic" condition).

Remark 2: Statement b) in Theorem 15 may be stated as: "all invariant factors of $G(D)$ are zero-free" (see Definition 1). This test requires computing only the sets of $k \times k$ and $(k-1) \times (k-1)$ minors of $G(D)$.

Remark 3: The fact that a rational matrix is zero-free at p if and only if its inverse is pole-free at p is identified as a "generalized Bezout test" in [13]. So this theorem might be called a "global generalized Bezout test."

Example 5: Let

$$G(D) = \begin{pmatrix} 1 & 0 \\ D & 1 \end{pmatrix}$$

be a generator matrix over $\mathbb{F}_2$. Then the 1×1 minors of $G(D)$ are $\{1, 0, D, 1\}$, and the 2×2 minor is the determinant $\det G(D) = 1$. The greatest common polynomial divisor of the 1×1 minors is 1, so $\delta_{p,1} = 0$ for all $p \in \mathcal{P}$. However, the maximum degree of the 1×1 minors is 1, so it follows that $\delta_{D^{-1},1} = -1$. Since $\det G(D) = 1$ we have $\delta_{p,2} = 0$ for all $p \in \mathcal{P}^*$. Therefore

$$\gamma_{p,2} = \delta_{p,2} - \delta_{p,1} = \begin{cases} 0, & p \in \mathcal{P} \\ 1, & p = D^{-1} \end{cases}$$

so $G(D)$ is not globally zero-free and not minimal.

Indeed, $G(D)$ does have an $F[D]$-inverse, namely, its unique inverse

$$G^{-1}(D) = \begin{pmatrix} 1 & 0 \\ -D & 1 \end{pmatrix}$$

but $G^{-1}(D)$ is not an $F[D^{-1}]$-matrix, so $G(D)$ has no $F[D^{-1}]$-inverse.

Example 6 [5]: Let

$$G(D) = \begin{pmatrix} 1 & \frac{D}{1+D} & \frac{1}{1+D} \\ \frac{D^2}{1+D+D^2} & \frac{1}{1+D+D^2} & 1 \end{pmatrix}$$

be a generator matrix over $\mathbb{F}_2$. The greatest common divisor of the 1×1 minors of $G(D)$ is $\Delta_1(D) = 1/(1 + D^3)$ and that of the 2×2 minors is $\Delta_2(D) = 1/(1 + D^3)$. Therefore, $\gamma_2(D) = \Delta_2(D)/\Delta_1(D) = 1$.

$G(D)$ can also be written as a rational matrix in D^{-1}, viz.,

$$G(D) = \begin{pmatrix} 1 & \frac{1}{1+D^{-1}} & \frac{D^{-1}}{1+D^{-1}} \\ \frac{1}{1+D^{-1}+D^{-2}} & \frac{D^{-2}}{1+D^{-1}+D^{-2}} & 1 \end{pmatrix}.$$

We have, similarly,

$$\widetilde{\Delta}_1(D^{-1}) = 1/(1 + D^{-3})$$

and

$$\widetilde{\Delta}_2(D^{-1}) = 1/(1 + D^{-3}).$$

Therefore, we also have $\gamma_2(D^{-1}) = 1$. Thus $\gamma_{p,2} = 0$ for all $p \in \mathcal{P}^*$. By Theorem 15, $G(D)$ is minimal.

We shall now see that conditions b) and c) of Theorem 15 are equivalent to a more general global invertibility condition, as follows. Let $\mathcal{P}'$ be any proper subset of $\mathcal{P}^*$, and define the ring $R(\mathcal{P}')$ as

$$R(\mathcal{P}') \stackrel{\text{def}}{=} \{g(D) \in F(D) \mid e_p(g(D)) \geq 0 \quad \text{for all } p \in \mathcal{P}'\}.$$

(We exclude $\mathcal{P}' = \mathcal{P}^*$, since then by the product formula $R(\mathcal{P}^*)$ is merely the set of degree-zero polynomials and 0: $R(\mathcal{P}^*) = \{g(D) = g_0 \mid g_0 \in F\}$.) Then $R(\mathcal{P}')$ is a principal ideal domain whose primes are the elements of $\mathcal{P}'$, and $F(D)$ is the field of quotients of $\mathcal{P}'$. Examples are:

a) If $\mathcal{P}' = \mathcal{P}^* \backslash \{D^{-1}\}$, then $R(\mathcal{P}') = F[D]$.
b) If $\mathcal{P}' = \mathcal{P}^* \backslash \{D\}$, then $R(\mathcal{P}') = F[D^{-1}]$.
c) If $\mathcal{P}' = \mathcal{P}^* \backslash \{D, D^{-1}\}$, then $R(\mathcal{P}')$ is the ring of finite sequences, sometimes called the set of Laurent polynomials, denoted by $F[D, D^{-1}]$.
d) If $\mathcal{P}' = \{D\}$, then $R(\mathcal{P}')$ is the ring of causal rational functions; i.e., $R(\mathcal{P}')$ is the set of rational functions which when expanded as formal Laurent series in D are in the set $F[[D]]$ of formal power series in D.
e) If $\mathcal{P}' = \{p\}$ for any $p \in \mathcal{P}^*$, then $R(\mathcal{P}')$ is the set of rational functions which when expanded as formal Laurent series in p over $F[D]_p$ are in the set of formal power series in p over $F[D]_p$, which we denote simply by $F[[p]]$.

The invariant factor theorem holds over any such ring $R(\mathcal{P}')$, and the invariant factors $\gamma_{p,i}$ for every prime $p \in \mathcal{P}'$ are unchanged from those defined above. It follows from Theorem 15 that $G(D)$ is minimal if and only if $G(D)$ has an $R(\mathcal{P}')$-inverse for every such proper subset $\mathcal{P}' \subsetneq \mathcal{P}^*$: furthermore, if $\{\mathcal{P}^{(n)}\}$ is any collection of subsets of $\mathcal{P}^*$ whose union is $\mathcal{P}^*$, then $G(D)$ is minimal if and only if $G(D)$ has an $R(\mathcal{P}^{(n)})$-inverse for every $\mathcal{P}^{(n)}$ in that collection. For example:

Corollary 16: A rational $k \times n$ generator matrix $G(D)$ is minimal if and only if

a) $G(D)$ has both an $F[D]$-inverse and an $R(\{D^{-1}\})$-inverse (i.e., a rational $F[[D^{-1}]]$-inverse).
b) $G(D)$ has an $F[D, D^{-1}]$-inverse, an $F[[D]]$-inverse, and an $F[[D^{-1}]]$-inverse.
c) For all $p \in \mathcal{P}^*$, $G(D)$ has an $R(\{p\})$-inverse (i.e., a rational $F[[p]]$-inverse). □

Thus we may say that $G(D)$ is minimal if and only if $G(D)$ is globally invertible.

Remark: It is easy to see that $G(D)$ is noncatastrophic (the output $\boldsymbol{u}(D)G(D)$ is finite if and only if the input $\boldsymbol{u}(D)$ is finite) if and only if $G(D)$ has a finite right inverse (i.e., an $F[D, D^{-1}]$-inverse). Corollary 16 shows that noncatastrophicity is necessary but not sufficient for minimality (e.g., $G(D) = D$ is noncatastrophic but nonminimal). Compare the "wrapping input property" of Fornasini and Valcher [25], which is equivalent to noncatastrophicity for one-dimensional Laurent polynomial generator matrices.

Example 7: Let $G(D)$ be the 2×2 matrix

$$G(D) = \begin{pmatrix} 1 & 0 \\ 1 + D^2 & D \end{pmatrix}.$$

Then the sets $\boldsymbol{\delta}_k$ of p-valuations $\delta_{p,k}$ of $k \times k$ minors for $p = D^{-1}, D, 1 + D, \ldots$ are

$$\boldsymbol{\delta}_1 = \{-2, 0, 0, \cdots\}$$
$$\boldsymbol{\delta}_2 = \{-1, 1, 0, \cdots\}$$

which implies that the sets $\boldsymbol{\gamma}_k$ of invariant factors $\gamma_{p,k}$ are

$$\boldsymbol{\gamma}_1 = \{-2, 0, 0, \cdots\}$$
$$\boldsymbol{\gamma}_2 = \{1, 1, 0, \cdots\}.$$

Therefore, $G(D)$ has a finite $F[D, D^{-1}]$-inverse and is thus noncatastrophic, but $G(D)$ does not have either an $F[D]$- or an $F[D^{-1}]$-inverse. Indeed, the unique inverse of $G(D)$ is

$$G^{-1}(D) = \begin{pmatrix} 1 & 0 \\ D^{-1} + D & D^{-1} \end{pmatrix}.$$

IX. Discussion

This paper shows that canonicality is the intersection of two independent properties: minimality and the global predictable valuation property.

Minimality is identified with the absence of any zeros, including at D^{-1}; with global invertibility; or with the property that the output always "covers" the input, in the sense that $e_p(\boldsymbol{u}(D)G(D)) \leq e_p(\boldsymbol{u}(D))$ for all $p \in \mathcal{P}^*$ and thus $\operatorname{span} \boldsymbol{u}(D) \subseteq \operatorname{span} \boldsymbol{u}(D)G(D)$.

The GPVP is identified with the property that the residue matrix $[G(D)]_p$ has full rank for all $p \in \mathcal{P}^*$, and with the property that the external defect of $G(D)$ is equal to the internal defect of $G(D)$, which is the minimal state-space dimension $\nu(\mathcal{C})$. A minimal generator matrix need not have the GPVP, and a matrix with the GPVP need not be minimal.

From a vector space viewpoint, a globally orthogonal basis $G(D) = \{\boldsymbol{g}_i(D), 1 \leq i \leq k\}$ corresponds to a canonical decomposition of a rational vector space $\mathcal{C}_r$ into a direct sum $\mathcal{C}_r = \mathcal{C}_1 + \mathcal{C}_2 + \cdots + \mathcal{C}_k$ of one-dimensional subspaces $\mathcal{C}_i$ of minimal defect, where we regard the defect $\operatorname{def} \mathcal{C}_i$ of a one-dimensional space $\mathcal{C}_i$ as the common defect of its nonzero vectors. To obtain a canonical basis, we select any vector $\boldsymbol{g}_i(D)$ from each subspace $\mathcal{C}_i$ that can be realized with $\operatorname{def} \mathcal{C}_i$ memory elements. To obtain a canonical polynomial basis, we select the essentially unique (up to units) polynomial vector $\boldsymbol{g}'_i(D)$ in each $\mathcal{C}_i$ whose degree is equal to $\operatorname{def} \mathcal{C}_i$.

ACKNOWLEDGMENT

The authors are grateful for the detailed comments of P. Rapisarda, M. E. Valcher, and the reviewers on earlier versions of the manuscript.

REFERENCES

[1] G. D. Forney Jr., "Convolutional codes I: Algebraic structure," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 720–738, 1970.
[2] ——, "Minimal bases of rational vector spaces, with applications to multivariable linear systems," *SIAM J. Contr.*, vol. 18, pp. 499–520, 1975.
[3] ——, "Algebraic structure of convolutional codes, and algebraic system theory," in *Mathematical System Theory*, A. C. Antoulas, Ed. New York: Springer-Verlag, 1991, pp. 527–558.
[4] R. Johannesson and Z.-x. Wan, "A linear algebra approach to minimal convolutional encoders," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1219–1233, 1993.
[5] ——, "On canonical encoding matrices and the generalized constraint lengths of convolutional codes," in *Communications and Cryptography*, R. E. Blahut *et al.*, Eds. Boston, MA: Kluwer, 1994, pp. 187–199.
[6] ——, "A generalization of the predictable degree property to rational convolutional encoding matrices," in *Proc. 1994 IEEE Int. Symp. on Information Theory* (Trondheim, Norway, June 27–July 1, 1994), p. 17.
[7] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes (1)," in *Proc. IEEE Int. Conf. on Communications* (Geneva, Switzerland, May 1983), pp. 1064–1070.
[8] A. F. Monna, *Analyse Non-Archimedienne*. Berlin: Springer, 1970.
[9] G. D. Forney, Jr., and M. D. Trott, "The dynamics of group codes: State spaces, trellis diagrams, and canonical encoders," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1491–1513, Sept. 1993.
[10] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.
[11] S. Kung and T. Kailath, "Some notes on valuation theory in linear systems," in *Proc. IEEE Conf. on Decision and Control*, 1978, pp. 515–517.
[12] B. Lévy, Ph.D. dissertation, Elec. Eng. Dept., Stanford Univ., Stanford, CA, 1978.
[13] G. C. Verghese and T. Kailath, "Rational matrix structure," *IEEE Trans. Automat. Contr.*, vol. 26, pp. 434–439, Apr. 1981.
[14] B. F. Wyman and M. K. Sain, "A unified pole-zero module for linear transfer functions," *Syst. Contr. Lett.*, vol. 5, pp. 117–120, 1984.
[15] J. H. M. Wedderburn, *Lectures on Matrices* (Amer. Math. Soc. Colloquium Lectures). New York: Dover, 1934 (reprint).
[16] N. P. Vekua, *Systems of Singular Integral Equations*. Groningen, The Netherlands: Nordhoff, 1967; Russian original, 1950.
[17] F. R. Gantmacher, *Theory of Matrices*, vols. I and II. New York: Chelsea, 1959.
[18] N. Jacobson, *Basic Algebra II*, 2nd ed. New York: Freeman, 1989.
[19] J. C. Willems, "Models for dynamics," in *Dynamics Reported*, vol. 2, U. Kirchgraber and H. O. Walther, Eds. New York: Wiley, 1989, pp. 171–269.
[20] H.-A. Loeliger, G. D. Forney, Jr., T. Mittelholzer, and M. D. Trott, "Minimality and observability of group systems," *Linear Algebra and Appl.*, vol. 205–206, pp. 937–963, July 1994.
[21] C. Roos, "On the structure of convolutional and cyclic convolutional codes," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 676–683, Sept. 1979.
[22] R. E. Kalman, P. L. Falb, and M. A. Arbib, *Topics in Mathematical System Theory*. New York: McGraw-Hill, 1969.
[23] B. Hartley and T. O. Hawkes, *Rings, Modules and Linear Algebra*. London, UK: Chapman and Hall, 1970.
[24] G. D. Forney, Jr., "Correction to 'Convolutional codes I: Algebraic structure'," *IEEE Trans. Inform. Theory*, vol. IT-17, p. 360, 1971.
[25] E. Fornasini and M. E. Valcher, "Multidimensional systems with finite support behaviors: Signal structure, generation and detection," preprint, Feb. 1995.