Reprinted from *Proc. 4th Joint Swedish-Soviet Int. workshop on Info. Th.*, 1989, pp. 14-18.

CONVOLUTIONAL CODES OVER RINGS

James L. Massey and Thomas Mittelholzer
Inst. for Signal and Info. Proc.
Swiss Federal Institute of Technology
8092 Zurich, SWITZERLAND

Abstract. It is shown that the "natural" linear codes for M-ary phase modulation are linear codes over the ring of integers modulo M, Z_M. Examples are given to illustrate the "normal" behavior of convolutional codes over Z_M and the pathologies that can arise. The condition, when $M = p^m$ and p is a prime, for a polynomial encoder to be catastrophic is given. It is shown that a convolutional code over Z_M can have no non-catastrophic polynomial encoder and no minimal polynomial encoder.

Why Convolutional Codes over Rings?

We stress that our introduction of convolutional codes over rings is not the consequence of a desire on our part to employ ever more esoteric algebraic coding theory, but was forced upon us by our investigation of codes for M-ary phase modulation. With the normalization to unity of the energy in each modulation symbol, the signal points for M-ary phase modulation are equally-spaced around the unit circle in the appropriate two-dimensional Euclidean signal space. Hence, we may consider the i-th of these signal points to be represented by the complex number W_M^i for $i = 0, 1, \ldots, M-1$ where $W_M = e^{j2\pi/M}$. The squared Euclidean distance between signal points i and j is thus

$$d_E^2 = \mathrm{abs}^2(W_M^j - W_M^i) = \mathrm{abs}^2(1 - W_M^{j-i}). \quad (1)$$

The key point to notice here is that the difference j-i in (1) can be treated as a modulo-M difference because W_M is a primitive M-th root of unity in the complex plane.

It follows that if we consider i and j to be elements of the ring of integers modulo M, Z_M, if we define the phase weight of the element i by

$$w(i) = \mathrm{abs}^2(1 - W_M^i) \quad (2)$$

and the phase distance between the elements i and j by

$$d(i) = w(j - i), \quad (3)$$

and if we define the phase weight of a sequence $\underline{x}$ (resp. phase distance between two sequences $\underline{x}$ and $\underline{y}$) of elements of Z_M as the sum of the weights (resp. distances) in each component, then

$$d(\underline{x}, \underline{y}) = w(\underline{x} - \underline{y}). \quad (4)$$

More importantly, $d(\underline{x}, \underline{y})$ is exactly the squared Euclidean distance between the corresponding sequences of modulation symbols. It follows that if we employ linear codes over the ring of integers modulo M, then phase weight and phase distance will enjoy the same key property (4) as does Hamming distance for codes over finite fields, but with the advantage that this phase distance exactly equals the squared Euclidean distance between phase modulation sequences. Moreover, Z_M is essentially the unique algebraic system with these most desirable features. One is forced to consider codes over Z_M if he wishes to employ linear codes in a natural way for phase-modulated signals.

By an (n,k) "linear block code over Z_M", we mean a rank-k free submodule of the free R-module R^n where $R = Z_M$ (cf. [1, p. 171] for the appropriate algebraic definitions). By an (n,k) "linear convolutional code over Z_M", we mean the same except that R is now the ring of fractions whose numerators are polyomials with coefficents in Z_M, as are the denominators that are further restricted to have 1 as their trailing coefficient. These are the natural generalizations of the corresponding definitions for codes over fields.

It follows immediately from these definitions that minimum weight and minimum distance coincide for these ring linear codes just as they do for linear codes over fields. In particular, for ring convolutional codes, it follows that

$$d_{free} = w_{free}, \quad (5)$$

where d_{free} is the minimum phase distance between two encoded sequences that differ in their initial corresponding information digits and where w_{free} is the minimum phase weight of the non-

zero encoded sequences in the convolutional code. Rather than to compile further algebraic properties of linear ring codes, we give instead a sequence of examples that illustrate the most important aspects of ring convolutional codes.

A Well-Behaved Example.

(This and the following example will both utilize the ring Z_4.) Consider the (2,1) convolutional encoder with encoding matrix

$$G(D) = [1 \quad 1+2D]. \tag{6}$$

This encoding matrix is systematic and hence is non-catastrophic (i.e., there is no infinite weight input sequence that produces a finite weight encoded sequence). A realization of this encoder is shown in Fig. 1(a) and the corresponding state transition diagram is shown in Fig. 1(b). Each transition in the latter is labelled with the input information digit and the resulting two output digits. This encoder has only two states, although the coding alphabet has size 4. Observing from (2) that $w(1) = w(3) = 2$ and $w(2) = 4$ for Z_4 (and of course $w(0) = 0$), one easily checks from Fig. 1(b) that this code has $w_{free} = d_{free} = 8$, which one can show is the largest squared Euclidean distance attainable with any two-state trellis encoder for quadrature phase modulation.

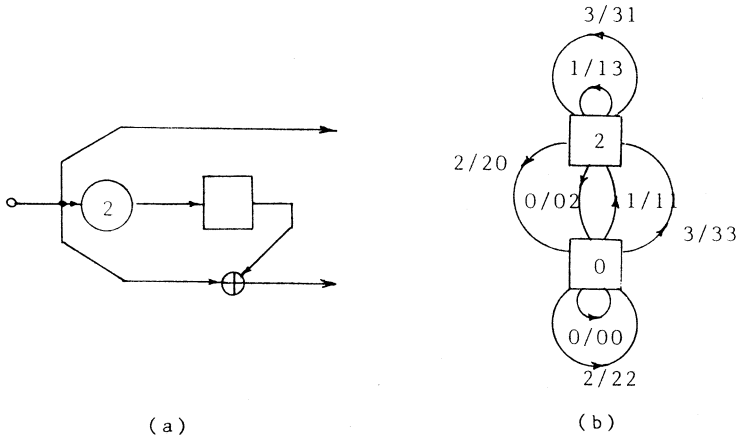

Fig. 1: (a) An encoder realizing the encoding matrix of (6) and (b) its state-transition diagram.

A Pathological Example.

The following example will show that many things can happen with ring convolutional codes that cannot occur in their field counterparts. Consider the (2,1) convolutional encoder with encoding matrix

$$G(D) = [1+D \quad 1+3D]. \tag{7}$$

The obvious realization of this encoder with one delay cell has 4 states, all of which are reachable from the zero state and no two of which are equivalent. No polynomial encoder for this code can be realized with fewer than 4 states. Yet, this encoder is catastrophic as the reader can verify by drawing its state transtion diagram, or by noting that the infinite weight (all-two) input $I(D) = [2/(1-D)] = [2/(1+3D)]$ gives the output $T(D) = I(D)G(D) = [2 \quad 2]$, as follows from the fact that $2(1+D)/(1-D) = 2(1-D+2D)/(1-D) = 2(1-D)/(1-D) = 2$. Yet there is no way to take out a common factor from the two polynomials in G(D) so as to reduce their degree (as one would be able to do for a (2,1) catastrophic encoder over a field). In fact, every polynomial encoder for this particular convolutional code is catastrophic, as can be checked without much difficulty.

How does one test a polynomial encoder for catastrophicity? The answer, whose proof will appear in a forthcoming paper [2], is somewhat surprising:

Theorem: A polynomial encoder G(D) over the ring Z_M, where $M = p^m$ and p is a prime, is catastrophic if and only if, when the coefficients of the polynomials in G(D) are each reduced modulo p, the resulting polynomial encoder over the finite field GF(p) is catastrophic.

The systematic encoder for the convolutional code encoded by the encoding matrix in (7) has the encoding matrix

$$G(D) = [1 \quad (1+3D)/(1+D)] = [1 \quad 1 + 2D/(1+D)] \tag{8}$$

which is realized by the encoder in Fig. 2 that has only two states. This means that this convolutional code has no polynomial encoder that is minimum in the sense of having the smallest possible number of states for any encoder for that code. We have the suspicion that there is always a minimal systematic encoder (appropriately defined), but we cannot yet prove this.

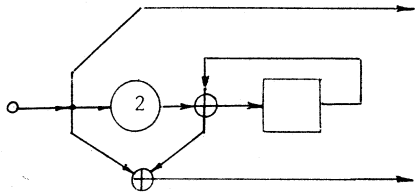

Fig. 2: A realization of the encoding matrix in (8).

How Good Are Ring Codes?

If one compares two convolutional encoders of the same rate (measured in bits of information per modulation symbol) and having the same number of encoder states (so that the complexity of Viterbi decoding is comparable for both) by the Euclidean distance achieved for M-ary phase modulation (and by the smallest number of occurrences of this distance in case of ties), then the ring codes appear to win hands down for M = 4 and M = 8 phase modulation. Our forthcoming paper [2] will give a list of ring codes for these cases showing that generally the ring codes give at least as large a Euclidean distance as any known field code or nonlinear trellis code even when the ring codes are constrained also to be "phase-invariant", i.e., closed (in an appropriate sense) under a phase shif of 360/M degrees, which is a most desirable practical property that cannot be attained with the linear field codes. Without the requirement to be phase-invariant, the ring convolutional codes are of course even better. We look for the ring codes to become the practical choice for coding of phase-modulated signals.

References.

[1] N. Jacobson, **Basic Algebra I (2nd Ed.).** New York: Freeman, 1985.

[2] J.L. Massey, T. Mittelholzer, T. Riedel and M. Vollenweider, "Ring Convolutional Codes for Phase Modulation" (manuscript in preparation).