Greedy Loop Transversal Codes, Metrics, and Lexicodes

F.A. Hummer and J.D.H. Smith
Department of Mathematics
Iowa State University
Ames, IA 50011, U.S.A.

ABSTRACT. In a group channel, codes correcting and detecting arbitrary patterns of errors (not necessarily "white noise") are described metrically. This yields sphere-packing and Gilbert bounds on the sizes of all and of maximal codes respectively. The loop transversal approach builds linear codes correcting arbitrary error patterns. In the binary case, the greedy loop transversal algorithm builds lexicodes.

The loop transversal approach to coding theory [S2] focuses on the set of errors to be corrected by a linear code, rather than on the code itself. It is particularly suited to the construction of codes correcting errors that do not have a "white noise" distribution. The present paper is part of a continuing programme developing the theory of such codes. The basic loop transversal approach is described in the first section. The second section concentrates on the binary case, considering loop transversal codes built by a greedy algorithm. The setting for these codes is the linearly ordered set of natural numbers within Conway's characteristic 2 field $\mathbf{On}_2$ [Co]. The third section introduces a metric such that codes in a group channel correcting and detecting general error patterns are characterized (Theorem 3.3) by a minimum distance between codewords, much as in the traditional white noise case. This yields a sphere-packing bound (Corollary 3.4) for all codes and a Gilbert bound (Corollary 3.5) for maximal codes. The form of these bounds is especially suited for application to loop transversal codes.

Loop transversal codes are always linear (although, for example, if they are linear over $\mathbb{Z}_4$, they may map to non-linear binary codes via the Gray map as in [CH]). Over arbitrary alphabets, the use of a greedy algorithm to build a so-called lexicode may yield a non-linear code. In the binary case, however, lexicodes are linear [BP], [CS], [Le]. The fourth section presents

JCMCC 22 (1996), pp. 143-155

an adaptation of the Brualdi-Pless proof of the linearity of white-noise lexicodes: it is this proof which lends itself to the comparison with greedy loop transferal codes. The result (Theorem 4.1) is formulated quite generally, showing how a binary greedy algorithm foliates vector spaces over GF(2). The general result is specialized in the fifth section to yield the linearity of white-noise lexicodes (Thrown 5.1) and lexicodes correcting and detecting arbitrary error patterns (Theorem 5.2). Corollary 5.3 derives the linearity of the Brualdi-Pless "B-greedy" codes in a novel way. Instead of being built greedily with respect to a non-standard ordering to correct a white-noise error pattern, they are considered as being built lexicographically (greedily with respect to the standard ordering) but correcting a distorted error pattern. The final section demonstrates the coincidence of binary lexicodes and greedy loop transversal codes for arbitrary error patterns (Theorem 6.1).

1 Loop transversal codes

A *transversal* T to a subgroup C of a group $(V, +, 0)$ is a subset of V with

$$V = \bigcup_{t \in T} (C + t). \tag{1.1}$$

Thus each element x of V can be expressed uniquely as

$$x = x\delta + x\varepsilon \tag{1.2}$$

with $x\delta$ in C and $x\varepsilon$ in T. If C is a linear code in the channel V, then (1.2) is interpreted as the decoding of a received word x to a codeword $x\delta$ with presumed error $x\varepsilon$. A binary operation $*$ is defined on T by

$$t * u = (t + u)\varepsilon. \tag{1.3}$$

For any t, u in T, the equation $v * t = u$ has a unique solution v [S1,§2.2]. If the equation $t * y = u$ also has a unique solution, then T is said to be a *loop transversal*. Equivalently, the algebra $(T, *, 0\varepsilon)$ is a loop. If V is abelian (as usual in coding theory), then each transversal is automatically a loop transversal, and the loop $(T, *, 0\varepsilon)$ is an abelian group. For x_i in T, it is convenient to use the notation $\prod_{i=1}^{r} x_i$ defined inductively by $\prod_{i=1}^{0} x_i = 0\varepsilon$ and $\prod_{i=1}^{r} x_i = \left[\prod_{i=1}^{r-1} x_i\right] * x_r$ for $r > 0$. In compound expressions, $*$ and $\prod$ will bind more strongly than $+$ and $\sum$.

Now specialize to the usual coding theory case that V is a finite-dimensional vector space and C a subspace over a field F. Define $\lambda \times t = (\lambda t)\varepsilon$ for λ in F and t in T. This makes $(T, *, F)$ a vector space over F. Induction on r

extends (1.3) to

$$\left(\sum_{i=1}^{r} \lambda_i t_i\right)\varepsilon = \prod_{i=1}^{r}(\lambda_i \times t_i) \tag{1.4}$$

for t_i in T. Assume that T contains a basis $\{e_1, \ldots, e_n\}$ for V, e.g. $V = F^n$ and each e_i has 1 in the i-th place as its only non-zero coordinate. Then knowledge of the vector space $(T, *, F)$ is sufficient to determine the code C. Indeed $C = \{v\delta \mid v \in V\} = \{v - v\varepsilon \mid v \in V\} = \{\sum_{i=1}^{r} \lambda_i e_i - (\sum_{i=1}^{r} \lambda_i e_i)\varepsilon \mid \lambda_i \in f\}$. Using (1.4), one then has

$$C = \left\{\sum_{i=1}^{r} \lambda_i e_i - \prod_{i=1}^{r}(\lambda_i \times e_i) \mid \lambda_i \in F\right\}. \tag{1.5}$$

This expression is useful for recovering C.

As an abstract vector space, the transversal $(T, *, F)$ is isomorphic to the dual (i.e. orthogonal complement) of the code C. Knowing a few elements of the usual dual merely restricts the possibilities for codewords, but does not specify any (non-zero ones) exactly. By contrast, knowing small parts of the transversal $(T, *, F)$ is sufficient to identify specific codewords; conversely knowing specific codewords determines part of the structure of $(T, *, F)$. In particular, for t_i in T,

$$\sum t_i - \prod t_i \in C. \tag{1.6}$$

The relationship (1.6) between C and T is such that small parts of T determine small parts of C. The relationship is describes as *local duality.*

Normally, local duality is the most efficient way of passing between the code and the transversal. However, an alternative route is available. Note that $(x+y)\varepsilon = x\varepsilon * y\varepsilon$ and $(\lambda x)\varepsilon = \lambda \times (x\varepsilon)$ for λ in F and x, y in V. Thus the *parity map*

$$\varepsilon\colon (V, +, F) \to (T, *, F) \tag{1.7}$$

is a linear transformation. Since the code is the kernel of the parity map, matrices of ε with respect to appropriate bases are parity-check matrices.

2 Greedy loop transversal codes

For the sake of simplicity, attention will generally be restricted to the binary case $F = GF(2)$ from now on. If ρ is a relation on a set X, and x is an element of X, the notation $x^\rho = \{y \in X \mid x\rho y\}$ will be used. Consider the ordered set $(\mathbb{N}, \leq)$ of natural numbers (including 0). The n-dimensional vector space over F may be realized concretely as the set $V_n = 2^{n>} =$

$\{0, 1, \ldots, 2^n - 1\}$ of natural numbers less than 2^n under the exponent-2 abelian group operations $+_2$ or $\sum$ of nim sum [Co, Chs. 6,11]. The subset $\{2^r \mid r < n\}$ forms a basis of weight-1 vectors for $V_n = 2^{n>}$. The advantage of this notation is that it provides a natural way to nest the various vector spaces $2^{n>}$ as $(2^{0>}, +_2) < (2^{1>}, +_2) < \cdots < (2^{n>}, +_2) < \cdots < (\mathbb{N}, +_2)$, with a corresponding nesting $0^> < \{1\} < \{1, 2\} < \cdots < \{2^r \mid r < n\} < \cdots < 2^{\mathbb{N}}$ of their standard bases of weight-1 vectors. A natural number m appears as an element in each vector space $2^{n>}$ with $n > \log_2 m$. Such a possibility of referring to a vector "locally", without leaving to mention explicitly a "global" vector space in which it appears, is crucial to exploiting the local duality afforded by the loop transversal approach.

The order $(\mathbb{N}, \leq)$ is total: any two elements are comparable. One property of the structure $(\mathbb{N}, +_2, \leq)$ is that it contains no chains of the form $m < n < n +_2 2^k < m +_2 2^k$ or $m < 2^k +_2 n < n < 2^k +_2 m$ [BP, Lemma 2.1]. The set $\mathbb{N}$ of natural numbers also carries a partial order. Each natural number m has a unique (binary) expansion as a sum

$$m = \sum_{i=0}^{\infty} m(i) 2^i \tag{2.1}$$

with $m(i) < 2$. A partial order $\subseteq$ is then defined by

$$m \subseteq m' \text{ iff } \forall i \in \mathbb{N}, m.(i) \leq m'(i). \tag{2.2}$$

For a subset X of a poset $(Y, \sqsubseteq)$, define the *subordinate set* $X^{\sqsupseteq}$ of X to be $\bigcup_{x \in X} x^{\sqsupseteq}$. The set X is *subordinate* or *self-subordinate* if $X = X^{\sqsupseteq}$.

An *error pattern* E is a self-subordinate subset of $(\mathbb{N}, \subseteq)$ containing $2^{\mathbb{N}}$, i.e.

$$2^{\mathbb{N}} \subset E = E^{\supseteq}. \tag{2.3}$$

An error pattern models a set of possible errors in the channels $2^{0>}, 2^{1>}, \ldots, 2^{n>}, \ldots$ that codes could be designed to correct. For example, the error pattern describing white noise double errors is

$$E = \{2^p + 2^q \mid \{p, q\} \subset \mathbb{N} \cup \{-\infty\}\}, \tag{2.4}$$

using the convention $2^{-\infty} = 0$. Codes correcting (2.4) are those of minimum distance at least 5. The error pattern describing burst errors of length at most 2 is (2.5)

$$E = \{0\} \cup 2^{\mathbb{N}} \cup 3.2^{\mathbb{N}}. \tag{2.5}$$

Linear block codes correcting (2.5) do not appear to have a good characterization in terms of traditional coding theory concepts such as distance. Of

course, codes correcting (2.4) will also correct (2.5). Indeed, error patterns are ordered by containment, and codes correcting an error pattern E will correct any error pattern E' contained E. The ordered set of error patterns has a least or "bottom" element

$$\perp = \{0\} \cup 2^{\mathbb{N}} \tag{2.6}$$

representing single errors.

Under the operation of nim sum, error patterns form partial algebras. For example, in the double-error pattern E of (2.4), nim sums of pairs $2^p, 2^q$ are defined, namely as $2^p +_2 2^q = 0$ for $p = q$ and $2^p +_2 2^q = 2^p + 2^q$ for $p \neq q$. In (2.5), these nim sums are only defined if $|p - q| \leq 1$.

Fix an error pattern E. Then an *E-syndrome*, or just *syndrome*, is a partial function $s\colon E \to \mathbb{N}$ which:

(a) injects,

(b) is a partial nim-sun homomorphism,

(c) has a domain self-subordinate in $(E, \leq)$, and (2.7)

(d) satisfies: $\forall n \in \mathbb{N}$, $\exists r \in \mathbb{N}$. $\perp \cap s(V_n \cap E)$ spans V_r.

Condition (b) means that $xs +_2 ys = zs$ for $\{x, y, z\} \subset E$ with $x +_2 y = z$. Condition (d) implies that $0s = 0$, $1s = 1$, $2s = 2$, and more generally $2^r s \leq 2^r$. The syndrome is said to be *proper* if s is a properly partial functioll. In view of (c), this is equivalent to finiteness of the domain of s. For a proper syndrome, the *length* is defined to be

$$n = \max\{1 + \lfloor \log_2 m \rfloor \mid m \in \operatorname{dom} s\}. \tag{2.8}$$

The *redundancy* is defined to be

$$r = \max\{1 + \lfloor \log_2(ms) \rfloor \mid m \in \operatorname{dom} s\}. \tag{2.9}$$

A proper syndrome $s\colon E \rightarrow \mathbb{N}$ defines a parity map

$$\epsilon_s\colon 2^{n>} \to 2^{r>} \tag{2.10}$$

by linearity and $2^i \epsilon_s = s^i s$ for $i < n$. By (c), these values $2^i s$ are defined. Condition (b) guarantees that s agrees with ϵ_s on $V_n \cap E$. Condition (d) guarantees that ϵ_s surjects. Condition (a) guarantees that dom s imbeds into $2^{r>}$ under ϵ_s. Thus the partial algebra $(\operatorname{dom} s, +_2)$ may be extended to a total loop transversal $(T, *)$ isomorphic via ϵ_s to $(2^{r>}, +_2)$. Independently of the extension chosen, a code C of redundancy r in the channel V_n, of length n, correcting the set dom s of errors, is specified either as the kernel of ϵ_s or, more efficiently, by local duality.

The greedy algorithm for building the syndrome $s\colon E \to \mathbb{N}$ is

$$2^n s = \min(\mathbb{N} - \{ms +_2 m's \mid m, m' \in V_n \cap E \text{ and } 2^n +_2 m \in E\}). \tag{2.11}$$

This algorithm constructs the code

$$C_n = \ker(\varepsilon\colon V_n \to \mathbb{N}) \tag{2.12}$$

with the parity map of (2.10). The code C_n is called the *greedy loop-transversal code* of length n correcting the error pattern E. The successive dimensions of the codes C_n are collected by the following relationship.

Proposition 2.1.

$$\dim C_n \leq \dim C_{n+1} \leq 1 + \dim C_n. \tag{2.13}$$

Proof: Since $\varepsilon\colon V_{n+1} \to \mathbb{N}$ extends $\varepsilon\colon V_n \to \mathbb{N}$, one has $C_n = \ker(\varepsilon\colon V_n \to \mathbb{N}) \leq \ker(\varepsilon\colon V_{n+1} \to \mathbb{N}) = C_{n+1}$ and $\varepsilon(V_n) \leq \varepsilon(V_{n+1})$. Thus $\dim C_n \leq \dim C_{n+1} = 1 + n - \dim \varepsilon(V_{n+1}) \leq 1 + n - \dim \varepsilon(V_n) = 1 + \dim C_n$. □

3 Metrics and error control

Let $(F, +_F, -_F, 0)$ be an abelian group. Let $V = F^n$ be the n-dimensional channel over the alphabet F. Let D and E be error sets in V that are closed under negation, with 0 lying in E. The channel V decomposes as the disjoint union

$$V = \bigcup_{i<4} (N_i - N_{i-1}), \tag{3.1}$$

where the *neighborhoods* of zero N_i are

$$\begin{aligned} N_{-1} &= \varnothing,\ N_0 = \{0\},\ N_1 = E, \\ N_2 &= (D \cup E) +_F E \text{ and} \\ N_3 &= V. \end{aligned}$$

Define a natural-number valued norm on V by

$$\|x\| = i \Leftrightarrow x \in N_i - N_{i-1}. \tag{3.2}$$

Proposition 3.1. *The norm (3.2) satisfies the triangle inequality*

$$\|x +_F y\| \leq \|x\| + \|y\| \tag{3.3}$$

for x, y in V.

Proof: The equality is trivially satisfied if the left hand side is less than 3. Otherwise, $x +_F y \notin (D \cup E) +_F E$, but then x and y cannot both lie in E. □

Corollary 3.2. *Under the distance function*

$$(x, y) \mapsto \|x -_f y\|, \tag{3.4}$$

the channel V becomes a metric space. □

Corollary 3.2 provides a metric approach to code design in channels with asymmetric error sets analogous to the convential metric approach used for white noise channels. The set E comprises the errors to be corrected by a code, while D comprises the (possibly empty set of) errors to be detected l)ut not necessarily corrected.

Theorem 3.3. *Let C be a subset of V. Suppose $0 \in E = -_F E$ and $D = -_F D$. Then C is an E-correcting, D-detecting code iff the minimum distance between its elements under the metric (3.4) is 3. In particular, suppose that D is empty. Then C is an E-correcting code iff it has minimum distance 3 under (3.1).*

Proof: If C is not E-correcting, there is a pair of distinct codewords c_1, c_2 and a pair of E-errors e_1, e_2 such that $c_1 +_F e_1 = c_2 +_F e_2$. Then $c_1 -_F c_2 = e_2 -_F e_1 = e_2 +_F (-_F e_1) \in E +_F E$, so that $\|c_1 -_F c_2\| < 3$. If C is not D-detecting, there is a pair of distinct codewords c_1, c_2, a D-error d and all E-error e such that $c_1 +_F e = c_2 +_F d$. Then $c_1 -_F c_2 = d +_F (-_F e) \in (D \cup E) +_F E$, so again $\|c_1 -_F c_2\| < 3$.

Conversely, suppose that there is a pair of distinct codewords c_1, c_2 with $\|c_1 -_F c_2\| < 3$, i.e. with $c_1 -_F c_2 = d +_F e$, $d \in D \cup E$, $e \in E$. Then $c_1 +_F (-_F d) = c_2 +_F e$. If d, and thus $-_F d$, lies in D, then C does not detect the D-error $-_F d$. If d, and thus $-_F d$, lies in E, then C cannot correct the E-errors $-_F d$ or e. □

Let $\mathcal{C}(V)$ denote the subposet of the power set $\mathcal{P}(V)$ comprising all codes correcting errors from E. For a finite alphabet F, Theorem 3.3 then yields versions of the sphere-packing and Gilbert bounds.

Corollary 3.4 (Sphere-packing bound). *For any element C of $\mathcal{C}(V)$,*

$$|C| \leq |V|/|N_1| = |V|/|E|. \tag{3.5}$$

Proof: The first neighborhoods $c +_F N_1$ of the codewords are disjoint. Indeed, $(c_1 +_F N_1) \cap (c_2 +_F N_1) \neq 0$ for distinct codewords c_1, c_2, say $c_1 +_F e_1 = c_2 +_F e_2$ with $e_i \in E$, would imply $\|c_1 -_F c_2\| < 3$. □

Corollary 3.5 (Gilbert bound). *For a maximal element c of $\mathcal{C}(V)$,*

$$|C| \geq |V|/|N_2| = |V|/|E +_F E|. \tag{3.6}$$

Proof: If a code C is such that $|C| < |C|/|N_2|$, then the second neighborhoods $c +_F N_2$ of the codewords do not cover V. An uncovered element v could then be adjoined to C to produce a larger element $C \cup \{v\}$ of $\mathcal{C}(V)$. □

Maximal codes C thus satisfy

$$\frac{|V|}{|N_2|} \leq |C| \leq \frac{|V|}{|N_1|}. \tag{3.7}$$

4 Linearity of greedy foliations

Let $(V, +_2)$ be the n-dimensional initial segment $2^{n>}$ of the natural numbers, considered as a vector space over $GF(2)$. Let $\chi\colon V \to \{0,1\}$ satisfy $0_\chi = 1$. Define a function

$$g\colon (V, +_2) \to (\mathbb{N}, +_2) \tag{4.1}$$

inductively by the following greedy algorithm:

$$\begin{array}{l} \textbf{begin } 0g := 0 \\ \textbf{for} \quad 1 \leq x \leq 2^n \textbf{ do} \\ \qquad \textbf{if} \quad \forall t \in g(x^>),\ \chi(x +_2 (g\lceil x^>)^{-1}\{t\}) \neq \{1\} \\ \qquad \textbf{then } xg := \min(\mathbb{N} - g(x^>)) \\ \qquad \textbf{else} \quad xg := \min\{t \mid \chi(x +_2 (g\lceil x^>)^{-1}\{t\}) = \{1\}\} \\ \textbf{end.} \end{array} \tag{4.2}$$

Thus the algorithm assigns the smallest possible numerical value to a vector x consistent with the requirement that no two vectors y, z in any pre-image $g^{-1}\{t\}$ have $(y +_2 z)\chi = 0$.

Theorem 4.1. *The function (4.1) defined by the greedy algorithm (4.2) is linear.*

Proof: Filter the vector space V as

$$\{0\} = V_0 < V_1 < \cdots < V_n = V \tag{4.3}$$

with subspaces $V_i = 2^{i>}$ of dimension i. Induction on i will be used to prove the linearity of the restriction $g_i\colon (V_i, +_2) \to (\mathbb{N}, +_2)$ of g to V_i. Certainly g_0 is linear. Assume that $g_i\colon (V_i, +_2) \to (\mathbb{N}, +_2)$ is linear for fixed i. To prove the linearity of g_{i+1}, it suffices to verify

$$(2^i +_2 z)g_{i+1} = 2^i g_{i+1} +_2 z g_i \tag{4.4}$$

for each z in V_i. There are two cases to deal with.

Case I: $\forall t \in g_i(V_i),\ \chi(2^i +_2 g_i^{-1}\{t\}) \neq \{1\}$.

By the greediness and linearity of g_i, the image $g_i(V_i)$ is an initial subspace $2^{j>}$ of $(\mathbb{N}, +_2, \leq)$. Thus $\min(\mathbb{N} - g_i(V_i)) = 2^j$ and $2^i g_{i+1} = 2^j$, so that (4.4) reduces to

$$(2^i +_2 z)g_{i+1} = 2^j +_2 zg_i. \tag{4.5}$$

Note that (4.5) holds for $z = 0$. To prove (4.5) by induction on z, assume that (4.5) holds for all z less than some non-zero y in V_i. Consider the computation of $(2^i +_2 y)g_{i+1}$ by (4.2). For t in $2^{j>}$, one has $\chi(2^i +_2 y +_2 g_i^{-1}\{t\}) = \chi(2^i +_2 g_i^{-1}\{yg_i +_2 t\}) \neq \{1\}$ and $\chi(2^i +_2 y +_2 g_{i+1}^{-1}\{2^j +_2 t\}) = \chi(2^i +_2 y +_2 +_2(g_i\lceil y^{>})^{-1}\{t\}) = \chi(y +_2 (g_i\lceil y^{>})^{-1}\{t\})$. The circuit

$$\begin{array}{ll} \textbf{if} & \forall t \in g((2^i +_2 y)^{>}),\ \chi(2^i +_2 y +_2 (g\lceil(2^i +_2 y)^{>})^{-1}\{t\}) \neq \{1\} \\ \textbf{then} & (2^i +_2 y)g := \min(\mathbb{N} \mathbin{\dot{-}} g((2^i +_2 y)^{>})) \\ \textbf{else} & (2^i +_2 y)g := \min\{t \mid \chi(2^i +_2 y +_2 (g\lceil(2^i +_2 y)^{>})^{-1}\{t\}) = \{1\} \end{array} \tag{4.6}$$

of the loop of the algorithm (4.2) thus reduces to

$$\begin{array}{ll} \textbf{if} & \forall t \in g(y^{>}),\ \chi(y +_2 (g_i\lceil y^{>})^{-1}\{t\}) \neq \{1\} \\ \textbf{then} & (2^i +_2 y)g := \min(2^j +_2 (\mathbb{N} - g(y^{>})) \\ \textbf{else} & (2^i +_2 y)g := \min\{2^j +_2 t \mid \chi(y +_2 (g_i\lceil y^{>})^{-1}\{t\}) = \{1\}\}, \end{array} \tag{4.7}$$

whence (4.5) holds also for $z = y$.

Case II: $\exists t \in g_i(V_i).\ \chi(2^i +_2 g_i^{-1}\{t\}) = \{1\}$.

Choose v minimal in $g_i(V_i)$ with $\chi(2^i +_2 g_i^{-1}\{v\}) = \{1\}$. Then $2^i g_{i+1} = v$, so that (4.4) reduces to

$$(2^i +_2 z)g_{i+1} = v + 2_z g_i \tag{4.8}$$

for z in V_i. Now for z in V_i, t in $g_i(V_i)$ and $cg_i = 0$, one has $\chi(2^i +_2 z +_2 c +_2 g_i^{-1}\{t\}) = \chi(2^i +_2 z +_2 g_i^{-1}\{t\})$ and $\chi(2^i +_2 z +_2 c +_2 2^i +_2 g_i^{-1}\{t\}) = \chi(2^i +_2 z +_2 2^i +_2 g_i^{-1}\{t\}) = \chi(z +_2 g_i^{-1}\{t\})$. Thus as z runs from 1 to $2^i - 1$, the algorithm (4.2) assigns values $(2^i +_2 z)g_{i+1}$ that only depend on the coset $g_i^{-1}\{zg_i\}$, not on z itself. In other words, $g_{i+1}\colon V_{i+1} \to \mathbb{N}$ factors through the natural projection:

$$\begin{array}{ccc} V_{i+1} & \longrightarrow & V_{i+1}/g_i^{-1}\{0\} \\ {\scriptstyle g_{i+1}}\big\downarrow & & \big\downarrow \\ \mathbb{N} & = & \mathbb{N} \end{array} \qquad \gamma. \tag{4.9}$$

The requirements (4.4) and (4.8) then reduce to

$$(2^i +_2 g_i^{-1}\{t\})\gamma = t +_2 v \tag{4.10}$$

for t in $g_i(V_i)$. This requirement already holds for $t = 0$. It will be proved by induction. Assume that it holds for all t in $u^>$, so that in particular

$$(2^i +_2 g_i^{-1}\{t\})\gamma \neq u +_2 v \tag{4.11}$$

for $t < u$. Set $s = (2^i +_2 g_i^{-1}\{u\})\gamma$. Consider $2^i +_2 z = \min(2^i +_2 g_i^{-1}\{u\})$. By (4.11), $u +_2 v \notin g_{i+1}((2^i +_2 z)^> - V_i)$. Thus

$$s \leq u +_2 v = 2^k +_2 w, \tag{4.12}$$

with $k = \lfloor \log_2 u \rfloor$ (so that k is minimal with $2^k +_2 u < u$) and $w = 2^k +_2 u + 2_v = (2^i +_2 g_i^{-1}\{2^k +_2 u\})\gamma$.

Now suppose that $2^k +_2 s < w$ were to hold. Since $s = (2^i +_2 g_i^{-1}\{u\})\gamma$, one has $\{1\} = \chi(2^i +_2 g_i^{-1}\{u\} +_2 g_i^{-1}\{s\}) = \chi(2^i +_2 g_i^{-1}\{2^k +_2 u\} +_2 g_i^{-1}\{2^k +_2 s\})$. There would thus be some r strictly less than $2^k +_2 u$ with $(2^i +_2 g_i^{-1}\{r\})\gamma = 2^k +_2 s$. By the induction hypothesis, $r +_2 v = 2^k +_2 s$. Now $r < 2^k +_2 u < u$. Since $(\mathbb{N}, \leq)$ contains no chain of the form $r < 2^k +_2 u < 2^k +_2 (2^k +_2 u) < 2^k +_2 r$, it follows that $2^k +_2 r < u$. But then $(2^i +_2 g_i^{-1}\{2^k +_2 r\})\gamma = 2^k +_2 r +_2 v = s$ by the induction hypothesis, contradicting $(2^i +_2 g_i^{-1}\{u\})\gamma = s$. Thus in fact

$$w \leq 2^k +_2 s. \tag{4.13}$$

The induction step proving (4.10) will be complete if it can be shown that $s +_2 u +_2 v = 0$, i.e. $s = 2^k +_2 w$. Now by the choice of k, one has $u +_2 2^k +_2 (2^k)^> \subseteq u^>$. By the induction hypothesis, $t +_2 v = (2^i +_2 g_i^{-1}\{t\})\gamma \neq (2^i +_2 g_i^{-1}\{u\})\gamma = s$ for $t \in u +_2 2^k +_2 (2^k)^>$, whence $s \notin v +_2 u +_2 2^k +_2 (2^k)^> = w +_2 (2^k)^>$. Suppose that $s \neq 2^k +_2 w$. Then in view of (4.12) and (4.13), the only possible order relationships for the four-element subset $\{s, w, 2^k +_2 s, 2^k +_2 w\}$ of $(\mathbb{N}, \leq)$ would be

$$w < s < 2^k +_2 w < 2^k +_2 s \tag{4.14}$$

or

$$s < w < 2^k +_2 s < 2^k +_2 w. \tag{4.15}$$

First suppose that (4.14) holds. Then $s \in w +_2 2^k +_2 (2^k)^>$, i.e. $s +_2 2^k \in w +_2 (2^k)^>$. With (4.14), this gives the contradiction $2^k \geq |(2^k +_2 s)^> - w^>| > |(2^k +_2 s)^> - s^>| = 2^k$. On the other hand, suppose that (4.15) holds. Then again $s +_2 2^k \in w +_2 (2^k)^>$, so that $w +_2 2^k \in s +_2 (2^k)^>$. With (4.15), this gives the contradiction $2^k \geq |(2^k +_2 w)^> - s^>| > |(2^k +_2 w)^> - w^>| = 2^k$. Hence $s = 2^k +_2 w$, as required. □

5 Linearity of binary greedy codes

As a first application of Theorem 4.1, consider the n-dimensional binary channel V with Hamming weight w. For $0 \leq d \leq n$, let χ be the characteristic function of the complement of the shell

$$\{x \in V \mid 0 < w(x) < d\}. \tag{5.1}$$

In this case, the algorithm (4.2) builds the syndrome function of the lexicode of minimum distance d. Thus Theorem 4.1 recovers (cf. [CS])

Theorem 5.1. *The n-dimensional binary lexicode of minimum distance d is linear.* □

Now let E and D be arbitrary error sets in V, with 0 in E. Note that E and D are trivially closed under negation. Consider the neighborhoods of zero defined under (3.1). Let χ be the characteristic function of $N_0 \cup N_3$, i.e. the complement of the "shell" $N_1 \cup N_2$. Consider the function $g: V \to \mathbb{N}$ built by the algorithm (4.2) in this case. Let C be the kernel of g. By construction, C has minimum distance 3 under (3.4). By Theorem 3.3, the code C corrects E-errors and detects D-errors. Indeed, C is the greedy E-correcting and D-detecting code. Thus:

Theorem 5.2. *Let E and D be arbitrary error sets in the binary channel V, with 0 in E. Then the greedy E-correcting, D-detecting code C is linear.* □

Brualdi and Pless [BP] generalized lexicodes in V by considering an arbitrary ordered basis B of V, and ordering V by lexicographic order on the coordinate vectors of elements of V with respect to the basis B. This order was called the *B-order* on V. The code C in V constructed greedily with respect to the B-order subject to maintenance of minimum Hamming distance d was called the *B-greedy code of designed distance d.* Brualdi and Pless [BP, Th. 2.2] showed that B-greedy codes are linear. The proof of Theorem 4.1 above is modeled on the proof of [BP, Th. 2.2]. On the other hand, the Brualdi-Pless result may be recovered from Theorem 5.2.

Corollary 5.3. *For an arbitrary ordered basis B of a n-dimensional binary channel W, the B-greedy code of designed distance d is liner.*

Proof: Identify vectors from the channel W with their coordinate vectors with respect to the ordered basis B. Let V be the set of coordinate vectors. Let E be the set of coordinate vectors of W-vectors of Hamming weight at most $\lfloor (d-1)/2 \rfloor$. If d is odd, let D be empty. If d is even, let D be the set of coordinate vectors of W-vectors of Hamming weight $d/2$. The B-order on W corresponds to lexicographic order on V. The greedy E-correcting, D-detecting code in V is the set of coordinate vectors of the B-greedy code of designed distance d in W. By Theorem 5.2, the formel code is linear. Thus so is the latter. □

6 Binary loop transversal codes and lexicodes

Let E be an error pattern in $\mathbb{N}$, i.e. a self-subordinate subset of $(\mathbb{N}, \subseteq)$ containing $2^{\mathbb{N}}$. For each n, let V_n be the initial segment $2^{n>}$ considered as an n-dimensional binary channel. Let L_n be the lexicode correcting $E \cap V_n$. In other words, let $\chi: V_n \to 2^{>}$ be the characteristic function of $(E \cap V_n) \cup ((E \cap V_n) +_2 (E \cap V_n))$. Then if $g: V_n \to \mathbb{N}$ is built by the algorithm (4.2) using this function χ, the lexicode L_n is the kernel of g. Let C_n be the loop transversal code obtained from the greedy syndrome $s: E \cap V_n \to \mathbb{N}$.

Theorem 6.1. *For each natural number n, the lexicode L_n and greedy loop transversal code C_n coincide. Moreover, $s = g\lceil(E \cap V_n)$.*

Proof: By induction on n. Note $L_0 = \{0\} = C_0$. Suppose that the theorem holds up to dimension n. It will be verified for dimension $n+1$. First, suppose that $\dim L_{n+1} = \dim L_n$. Thus $L_{n+1} = L_n = C_n \subseteq C_{n+1}$. The maximality of L_{n+1} in $\mathcal{C}(V_{n+1})$ then shows that $L_{n+1} = C_{n+1}$. Since $\dim g(V_{n+1}) > \dim g(V_n)$, Case I of the proof of Theorem 4.1 describes the construction of $2^n g$. Thus $2^n g = \min(\mathbb{N} - g(V_n)) = 2^n s$.

Now suppose that $\dim L_{n+1} > \dim L_n$. If $\dim L_{n+1} > 1 + \dim L_n$, pick a basis for L_n and extend it to a basis for L_{n+1}. There are then two distinct elements x, y of the extension. By the maximality of L_n in $\mathcal{C}(V_n)$, neither x nor y lies in V_n. But $\dim(V_{n+1}/V_n) = 1$, so the element $x +_2 y$ of L_{n+1} lies in V_n. This contradicts the maximality of L_n in $\mathcal{C}(V_n)$. Hence $\dim L_{n+1} = 1 + \dim L_n$ in this case. Now $\dim L_{n+1} > \dim L_n$ implies $\dim C_{n+1} > \dim C_n$. Indeed, the increase in dimension in the lexicode construction implies that there is a possible choice of syndrome in the greedy loop transversal construction that will increase the dimension of the loop transversal code. Then by Proposition 2.1, $\dim C_{n+1} = 1 + \dim C_n$. Thus $\dim L_{n+1} = \dim C_{n+1}$ in this case.

Let $l = \min(L_{n+1} - L_n)$ and $c = \min(C_{n+1} - C_n)$. Then $L_{n+1} = C_{n+1}$ iff $l = c$. By the lexicode construction, $l \leq c$. Suppose $2^n s = \sum_{i \in S} 2^i$ and $2^n g = \sum_{i \in G} 2^i$. Suppose $l = 2^n + \sum_{j \in L} 2^j$. Then $lg = 0 = \sum_{i \in G} 2^i +_2 \sum_{j \in L} 2^j g = \sum_{i \in G} 2^i +_2 \sum_{j \in L} 2^j s$, so that $l = 2^n + \sum_{i \in G} 2^i s^{-1}$. Similarly (but using s in place of g), $c = 2^n + \sum_{i \in S} 2^i s^{-1}$. Then $2^n + \sum_{i \in G} 2^i s^{-1} \leq 2^n + \sum_{i \in S} 2^i s^{-1}$. But by the greediness of the loop transversal construction, and the monotonicity of s on $s^{-1}(2^{\mathbb{N}})$, one has $\sum_{i \in S} 2^i s^{-1} \leq \sum_{i \in G} 2^i s^{-1}$. Thus $l = c$ and $2^n s = 2^n g$, as required. □

References

[BP] R.A Brualdi and V. Pless, Greedy codes, *J. Comb. Th. (A)* **64** (1993), 10–30.

[CH] A.R. Calderbank, A.R. Hammons Jr., P.V. Kumar, N.J.A. Sloane and P. Solé, The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals and related codes, *IEEE Trans. Inform. Th.* **40** (1993), 301–319.

[Co] J.H. Conway, *On Numbers and Games*, Cambridge University Press, Cambridge, 1975.

[CS] J.H. Conway and N.J.A. Sloane, Lexicographic codes: error-correcting codes from game theory, *I.E.E.E. Trans.. Info. Th.* **IT-32** (1986), 337–348.

[Le] V.I. Levenshtein, Ob odnom klasse sistematicheskikh kodov, *Dokl. A.N. S.S.S.R.* **131** (1960), 1011–1014, translated as: A class of systematic codes, *Soviet Math. Dokl.* **1** (1960), 368–371.

[S1] J.D.H. Smith, *Representations of Infinite Groups and Finite Quasigroups*, Les Presses de l'Université de Montréal, Montreal, 1986.

[S2] J.D.H. Smith, Loop transversals to linear codes, *J. Comb., Info. and Syst. Sci.* **17** (1992), 1–8.