NON-LINEAR CODES FROM POINTS OF BOUNDED HEIGHT

Chris Hurlburt and Jeffrey Lin Thunder

Abstract. This paper generalizes Elkies' construction of error-correcting nonlinear codes found in [E]. The generalization produces a precise average code size over codes in the new construction. The result is a larger family of codes with similar transmission rates and error detection rates to the nonlinear codes found in [E]. Moreover, we exhibit a connection between these nonlinear codes and solutions to simple homogeneous linear equations defined over the function field of a curve.

Introduction

In this paper we give a generalization of Elkies' construction of error-correcting nonlinear codes found in [E]. Elkies' construction is as follows. Let C be a curve over a finite field $\mathbb{F}_q$ and let D be a divisor of degree zero on C. Elkies constructs a code by evaluating at the rational points on C all rational functions of degree less than a fixed bound in the line bundle associated to D. The resulting code has alphabet $\mathbb{F}_q \cup \{\infty\}$. For a large class of curves these codes are more efficient than Goppa codes over the same curve with the same designed minimal distance. To determine effiency, i.e., the transmission rate plus error detection rate of his codes, Elkies must estimate the average number of rational functions of bounded degree in the line bundle.

Our new approach uses methods and ideas from Diophantine geometry and adelic geometry of numbers to construct codes. Whereas Elkies works with a curve C and divisor D of degree zero, we work with the corresponding function field K and a matrix $B \in \mathrm{GL}_2(K_{\mathbb{A}})$, where $K_{\mathbb{A}}$ is the adele ring of K. This matrix gives rise to a twisted height on projective space over K. We consider all points in projective space over K of twisted height less than a fixed bound, and we construct our code by evaluating these points at all places of degree one. Elkies' codes are a proper subset of the set of codes obtained from our construction. Moreover, by using our larger collection of codes, we

Research of the first author partially supported by NSA grant MDA904-03-1-0031
Research of the second author partially supported by NSF grant DMS-0100791

Typeset by $\mathcal{AMS}$-TEX

are able to establish quite precisely an average transmission rate. Our designed minimal distance is the same as in Elkies' construction; hence Elkies' arguments for higher efficiency than Goppa codes apply equally well to this larger collection of codes.

Another benefit of our construction is the ability to relate each code to particular solutions of a homogeneous linear equation defined over K. In general, determining the points in projective space of twisted height less than a fixed bound corresponds to finding the solutions of height less than a fixed bound to a system of homogeneous linear equations. In our case each twisted height from a matrix $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ corresponds to a single homogeneous linear equation. Through this connection we are able to reformulate each of our codes in terms of the solution set to a homogeneous linear equation.

This paper is structured as follows. The first section establishes our notation and recalls the notion of twisted heights. In the next two sections we describe our construction of error-correcting codes, prove lower bounds for the distance between codewords, and prove how many codewords we get on average. The final section describes precisely how our codes arise from homogeneous linear equations and ends with some concluding remarks on possibilities for further development.

1. Notation and Definitions

Throughout the remainder of this paper, K will be a fixed finitely generated extension of a finite prime field $\mathbb{F}_p$, of transcendence degree 1 over $\mathbb{F}_p$. In other words, K will be a fixed finite algebraic extension of $\mathbb{F}_p[T]$, where T is transcendental over $\mathbb{F}_p$. We denote the cardinality of the field of constants by q. The field K corresponds to a nonsingular projective curve C over $\mathbb{F}_q$. We let $K_{\mathbb{A}}$, $K_{\mathbb{A}}^{\times}$, $M(K)$, and ζ_K denote the adele ring, idele group, set of places, and Dedekind zeta function of K, respectively. Let $J(K)$ denote the number of divisor classes of degree zero, i.e., the cardinality of the jacobian of the curve C.

For each place $v \in M(K)$, let K_v denote the completion of K at the place v and write $\mathrm{ord}_v(x)$ for the order of $x \in K_v$. Here ord_v is normalized so that its image is $\mathbb{Z} \cup \{\infty\}$. Let $\mathfrak{O}_v$ denote the maximal compact subring of K_v (the "v-adic integers"); then $\mathfrak{O}_v$ consists of all $x \in K_v$ with $\mathrm{ord}_v(x) \geq 0$, with the usual convention that $\infty > 0$. The field of constants $\mathbb{F}_q$ consists of 0 together

with all elements $x \in K$ with $\mathrm{ord}_v(x) = 0$ at all places v. For $\mathbf{x} = (x_1, \ldots, x_n) \in K_v^n$ we let

$$\mathrm{ord}_v(\mathbf{x}) = \min_{1 \le i \le n} \{\mathrm{ord}_v(x_i)\}.$$

If $a = (a_v) \in K_{\mathbb{A}}^{\times}$, we get a divisor

$$\mathrm{div}(a) = \sum_{v \in M(K)} \mathrm{ord}_v(a_v) \cdot v.$$

The adelic modulus is defined by $|a|_{\mathbb{A}} = q^{-\deg \mathrm{div}(a)}$. If $\mathbf{x} = (\mathbf{x}_v) \in K_{\mathbb{A}}^n$ is such that $\mathbf{x}_v \neq \mathbf{0}$ for all places v, then we have a divisor

$$\mathrm{div}(\mathbf{x}) = \sum_{v \in M(K)} \mathrm{ord}_v(\mathbf{x}_v) \cdot v,$$

and an adelic length defined by

$$\|\mathbf{x}\|_{\mathbb{A}} = q^{-\deg \mathrm{div}(\mathbf{x})}.$$

Note in particular that if $\mathbf{x}$ is a non-zero element of K^n, then we may view $\mathbf{x}$ via the usual diagonal embedding as such a vector in $K_{\mathbb{A}}^n$.

For $A \in \mathrm{GL}_n(K_{\mathbb{A}})$, we get the following twisted height on $K^n \setminus \{\mathbf{0}\}$:

$$H_A(\mathbf{x}) = \|A(\mathbf{x})\|_{\mathbb{A}}.$$

Note that $\mathrm{div}(aA(\mathbf{x})) = \mathrm{div}(a) + \mathrm{div}(A(\mathbf{x}))$ for any idele a. Thus

$$H_{aA}(\mathbf{x}) = |a|_{\mathbb{A}} H_A(\mathbf{x}). \tag{0}$$

In particular, $H_A(a\mathbf{x}) = H_A(\mathbf{x})$ for any $a \in K^{\times}$. Thus, H_A is really a function on projective $(n-1)$-space $\mathbb{P}^{n-1}(K)$. We let h_A denote the additive height, i.e., $h_A(\mathbf{x}) = \log_q H_A(\mathbf{x})$. Using the additive height, equation (0) becomes

$$h_{aA}(\mathbf{x}) = h_A(\mathbf{x}) - \deg \mathrm{div}(a). \tag{0'}$$

2. Codes

Choose an enumeration $v_1, \ldots, v_N$ of the places of degree 1. These places correspond to the $\mathbb{F}_q$-rational points on the curve C associated with K. Fix an $A \in \mathrm{GL}_2(K_{\mathbb{A}})$ of the form

$$A = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}. \tag{1}$$

Then for any $x \in K$, we associate a codeword of N letters with alphabet $\mathbb{F}_q \cup \{\infty\}$ by setting the i-th letter to be the residue in $\mathbb{F}_q$ of $a_{v_i}x + b_{v_i}$ if $\mathrm{ord}_{v_i}(a_{v_i}x + b_{v_i}) \geq 0$, or ∞ if $\mathrm{ord}_{v_i}(a_{v_i}x + b_{v_i}) < 0$. For a fixed parameter h, our code $C_A(h)$ will consist of the words associated to those x for which $h_A(x, 1) \leq h$. We note that when A is a matrix of the form (1) where $b = 0$ and $|a|_{\mathbb{A}} = 1$, the resulting code $C_A(h)$ is the code Elkies considers in [E] and denotes by $C_{\mathrm{div}(a)}(h)$. Thus, Elkies' nonlinear codes are particular examples of our codes.

Lemma 1. *Fix an A of the form (1) as above an $A \in \mathrm{GL}_2(K_{\mathbb{A}})$ of the form*

$$A = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$$

and an $h \in \mathbb{Z}$. Let x and y be distinct elements of K such that $h_A(x, 1),\ h_A(y, 1) \leq h$. Then the codewords associated to x and y have at least

$$N - 2h - \deg \mathrm{div}(a) = N - 2h - \deg \mathrm{div}(\det(A))$$

coordinates which are distinct.

Proof. Let $D_1 = \mathrm{div}\big(A(x, 1)\big)$ and $D_2 = \mathrm{div}\big(A(y, 1)\big)$. Note that both $-D_1$ and $-D_2$ are effective. Let

$$D = \mathrm{div}\big(A(x, 1) - A(y, 1)\big) = \mathrm{div}(a(x - y)) = \mathrm{div}(a) + \mathrm{div}(x - y).$$

Write D as a difference of effective divisors: $D = D^+ - D^-$. Then

$$\deg D = \deg D^+ - \deg D^- = \deg \mathrm{div}(a), \tag{2}$$

since $\mathrm{div}(x - y)$ is a principal divisor.

If $v \in M(K)$ is in the support of both D_1 and D_2, then the coefficient at v of $-D^-$ is greater than the sum of the coefficients of D_1 and D_2. For all other places, the coefficient at v of $-D^-$ is at least the sum of the coefficients of D_1 and D_2. Thus,

$$-\deg D^- \geq \deg D_1 + \deg D_2 + l, \tag{3}$$

where l is the number of places in the support of both D_1 and D_2.

Consider the set of places of degree one where the coordinates of the codewords associated with x and y match. For such a place v, either the corresponding letter is in $\mathbb{F}_q$, implying that v is in the support of D^+, or the letter is ∞, implying that v is in the support of both D_1 and D_2. Hence, the number of such places is no greater than $\deg D^+ + l$. We thus see by (2) and (3) that the number of coordinates in the codewords associated to x and y which are distinct is at least

$$\begin{aligned} N - (\deg D^+ + l) &= N - \deg \operatorname{div}(a) - \deg D^- - l \\ &\geq N - \deg \operatorname{div}(a) + \deg D_1 + \deg D_2 \\ &= N - \deg \operatorname{div}(a) - h_A(x,1) - h_A(y,1) \\ &\geq N - 2h - \deg \operatorname{div}(a). \end{aligned}$$

Some remarks concerning this distance bound are in order. First, when $\deg \operatorname{div}(a) = 0$ we recapture Elkies' distance bound, $N - 2h$. Also, we obviously must have $2h + \deg \operatorname{div}(a) < N$ to have an error correcting code. This puts an upper bound on how large the parameter h can be. As a final remark, we clearly can't have a distance bound larger than N itself, yet Lemma 1 seems to imply this possibility if $2h + \deg \operatorname{div}(a) < 0$. In fact, this can never occur. To see why, suppose $\mathbf{x}$ and $\mathbf{y}$ are linearly independent elements of K^2 and consider the element B of $\mathrm{GL}_2(K)$ with columns $\mathbf{x}^{tr}$ and $\mathbf{y}^{tr}$. Call this matrix B. Then the columns of the product AB are simply the transposes of $A(\mathbf{x})$ and $A(\mathbf{y})$. By Hadamard's inequality, we see that

$$\|A(\mathbf{x})\|_{\mathbb{A}} \cdot \|A(\mathbf{y})\|_{\mathbb{A}} \geq |\det(AB)|_{\mathbb{A}} = |\det(A)|_{\mathbb{A}} \cdot |\det(B)|_{\mathbb{A}} = |\det(A)|_{\mathbb{A}},$$

since $\operatorname{div}(\det(B))$ is a principal divisor. Thus,

$$h_A(\mathbf{x}) + h_A(\mathbf{y}) \geq -\deg \operatorname{div}(\det(A)) = -\deg \operatorname{div}(a)$$

for any two linearly independent $\mathbf{x}$ and $\mathbf{y}$. In particular, if $2h < -\deg \operatorname{div}(a)$, then Lemma 1 is inapplicable as there can be no two distinct $(x,1)$ and $(y,1)$ with $h_A(x,1),\ h_A(y,1) \leq h$.

3. Codes from Arbitrary Matrices and the Number of Codewords

Given a $B \in \mathrm{GL}_2(K_{\mathbb{A}})$, there is a norm-preserving $U \in \mathrm{GL}_2(K_{\mathbb{A}})$ such that UB is upper triangular. By norm-preserving, we mean that $\|U_v(\mathbf{x}_v)\|_v = \|\mathbf{x}_v\|_v$ for all $\mathbf{x}_v \in K_v^2$ and all places v. (This is equivalent to saying $U_v(\mathfrak{O}_v^2) = \mathfrak{O}_v^2$ for all places v.) To see why this is so, we remark that one can construct an upper triangular $T \in \mathrm{GL}_2(K_{\mathbb{A}})$ such that BT is norm-preserving; this is done via an analog of Gram-Schmidt. Then T^{-1} is the desired upper-triangular element of $\mathrm{GL}_2(K_{\mathbb{A}})$. Say the lower righthand corner entry of UB is $c \in K_{\mathbb{A}}^{\times}$. Then $c^{-1}UB = A$ will be of the form (1). The code we actually associate to B is the code obtained from A as described in the previous section. Though A is not uniquely determined, we will show in Section 4 that any two such As produce equivalent codes.

For a given $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ and $z \in \mathbb{Z}$, let $\mathcal{N}(B, z)$ denote the number of $\xi \in \mathbb{P}^1(K)$ such that $h_B(\xi) \leq z$. Alternately, $\mathcal{N}(B, z)$ is the number of one-dimensional subspaces $K\mathbf{x} \subset K^2$ such that $\deg \mathrm{div}\big(B(\mathbf{x})\big) \geq -z$.

Lemma 2. *Let $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ and $z \in \mathbb{Z}$. Then*

$$\mathcal{N}(B, z) = \mathcal{N}(B\gamma, z) = \mathcal{N}(UB, z)$$

for all $\gamma \in \mathrm{GL}_2(K)$ and norm-preserving $U \in \mathrm{GL}_2(K_{\mathbb{A}})$. Also,

$$\mathcal{N}(cB, z) = \mathcal{N}(B, z + \deg \mathrm{div}(c))$$

for all $c \in K_{\mathbb{A}}^{\times}$.

Proof. The first equality is clear since any $\gamma \in \mathrm{GL}_2(K)$ gives a permutation of the one-dimensional subspaces of K^2, and the second equality follows directly from the definition of height and norm-preserving. Finally, by equation (0'), $h_{cB}(\mathbf{x}) = h_B(\mathbf{x}) - \deg \mathrm{div}(c)$ for all non-zero $\mathbf{x} \in K^2$ and all $c \in K_{\mathbb{A}}^{\times}$.

Lemma 3. *Fix an $h \in \mathbb{Z}$ with $0 \leq h < N/2$ and a $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ such that $\deg \mathrm{div}(\det(B)) = 2m$ for $m \in \mathbb{Z}$. Choose a non-zero $\mathbf{x}_0 \in K^2$ with $h_B(\mathbf{x}_0) > \frac{N}{2} - m$ and a $\gamma \in \mathrm{GL}_2(K)$ with $\gamma(1, 0) = \mathbf{x}_0$. Let $c \in K_{\mathbb{A}}^{\times}$ and let $U \in \mathrm{GL}_2(K_{\mathbb{A}})$ be norm-preserving such that $cUB\gamma = A$ is a matrix of the form*

(1). *Then the code* $C_A(h - m - \deg \operatorname{div}(c))$ *has minimal distance at least* $N - 2h$ *and exactly* $\mathcal{N}(B, h - m)$ *codewords.*

Proof. Since $A = cUB\gamma$, $|\det(U)|_{\mathbb{A}} = 1$, and $|\det(\gamma)|_{\mathbb{A}} = 1$,

$$\deg \operatorname{div}(\det(A)) = \deg \operatorname{div}(\det(cB)) = 2 \deg \operatorname{div}(c) + \deg \operatorname{div}(\det(B)) = 2 \deg \operatorname{div}(c) + 2m.$$

By Lemma 1, a lower bound for the minimal distance is

$$N - 2(h - m - \deg \operatorname{div}(c)) - \deg \operatorname{div}(\det(A)) = N - 2h.$$

By equation (0'), the choice of γ, and the definition of norm-preserving,

$$\begin{aligned} h_A(1,0) = h_{cUB\gamma}(1,0) &= h_B(\mathbf{x}_0) - \deg \operatorname{div}(c) \\ &\geq \frac{N}{2} - m - \deg \operatorname{div}(c) \\ &> h - m - \deg \operatorname{div}(c). \end{aligned}$$

Thus, any $\xi \in \mathbb{P}^1(K)$ with $h_A(\xi) \leq h - m - \deg \operatorname{div}(c)$ has a representative of the form $(x, 1) \in K^2$. Whence, the number of codewords is

$$\begin{aligned} \mathcal{N}(A, h - m - \deg \operatorname{div}(c)) &= \mathcal{N}(cUB\gamma, h - m - \deg \operatorname{div}(c)) \\ &= \mathcal{N}(UB\gamma, h - m) \\ &= \mathcal{N}(B\gamma, h - m) \\ &= \mathcal{N}(B, h - m). \end{aligned}$$

In a similar manner, one can prove

Lemma 3'. *Fix an* $h' \in \mathbb{Z}$ *with* $1/2 \leq h' < (N+1)/2$ *and suppose* $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ *is such that* $\deg \operatorname{div}(\det(B)) = 2m + 1$ *for* $m \in \mathbb{Z}$. *Choose a non-zero* $\mathbf{x}_0 \in K^2$ *with* $h_B(\mathbf{x}_0) \geq \frac{N-1}{2} - m$ *and a* $\gamma \in \mathrm{GL}_2(K)$ *with* $\gamma(1,0) = \mathbf{x}_0$. *Let* $c \in K_{\mathbb{A}}^{\times}$ *and let* $U \in \mathrm{GL}_2(K_{\mathbb{A}})$ *be norm-preserving such that* $cUB\gamma = A$ *is of the form* (1). *Then the code* $C_A(h' - 1 - m - \deg \operatorname{div}(c))$ *has minimal distance at least* $N - 2h' + 1$ *and exactly* $\mathcal{N}(B, h' - 1 - m)$ *codewords.*

In order to establish the transmission rates of our codes, we need to estimate the quantities $\mathcal{N}(B, h - m)$ and $\mathcal{N}(B, h - 1 - m)$ occurring in the above two lemmas. As indicated in the introduction, what we will do is determine the average value (in a precise sense) of these quantities.

Choose an $a_0 \in K_{\mathbb{A}}^{\times}$ with $\deg \operatorname{div}(a_0) = 1$ and let

$$P = \begin{pmatrix} a_0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Define

$$S = \prod_v (\mathfrak{O}_v)^2.$$

For $\mathbf{x} \in (K_{\mathbb{A}})^2$ and $m \in \mathbb{Z}$ we define the "distance function"

$$\chi_m(\mathbf{x}) = \inf_{a \in K_{\mathbb{A}}^{\times}} \{|a|_{\mathbb{A}} \colon \mathbf{x} \in aP^{-m}(S)\}.$$

Note that for $\mathbf{x} = (\mathbf{x}_v)$ of the type $\mathbf{x}_v \neq \mathbf{0}$ for all places v, $\chi_0(\mathbf{x}) = \|\mathbf{x}\|_{\mathbb{A}}$. In particular, for non-zero $\mathbf{x} \in K^2$ and $B \in \mathrm{GL}_2(K_{\mathbb{A}})$, we have $\chi_0(B(\mathbf{x})) = H_B(\mathbf{x})$. More generally, for all $m \in \mathbb{Z}$ we have

$$\chi_m(B(\mathbf{x})) = H_{P^m B}(\mathbf{x}). \tag{4}$$

To ease the notation to follow, let G be the subgroup of $\mathrm{GL}_2(K_{\mathbb{A}})$ consisting of all those B with $|\det(B)|_{\mathbb{A}} = 1$ and let Γ be the discrete subgroup $\mathrm{GL}_2(K)$. There is a Haar measure μ on G for which $\mu(G/\Gamma) = 1$ (see [T, §3]). Let T be the subgroup of Γ consisting of all upper triangular matrices. One may view $\mathbb{P}^1(K)$ as the factor group Γ/T.

Fix a parameter $h \in \mathbb{Z}$ and let

$$f(x) = \begin{cases} 1 & \text{if } x \leq q^h, \\ \frac{q^{h+1}-x}{q^{h+1}-q^h} & \text{if } q^h \leq x \leq q^{h+1}, \\ 0 & \text{if } x \geq q^{h+1}. \end{cases}$$

We note that by equation (4), we may view $f \circ \chi_m(B*)$ as a function on Γ/T. As shown on page 178 of [T], we have

$$\int_{G/\Gamma} \left[\sum_{\xi \in \Gamma/T} f\big(\chi_m(B\xi)\big) \right] d\mu(B) = \frac{q^{2(1-g)+m} J(K)}{(q-1)\zeta_K(2)} \sum_{z \in \mathbb{Z}} q^{2z} f(q^z).$$

Note that this incorporates a correction to Lemma 1 of [T] in the function field case which should read

$$\kappa \frac{\sigma_n(g_n/\gamma_n)}{\mu_n(G_n/\Gamma_n)} = \frac{\alpha^n(S)h(K)}{(1-q^{-n})(q-1)\zeta_K(n)}.$$

A quick calculation gives us

$$\int_{G/\Gamma}\left[\sum_{\xi\in\Gamma/T} f\big(\chi_m(B\xi)\big)\right]d\mu(B)=\frac{q^{2(1-g+h)+m}J(K)}{(1-q^{-2})(q-1)\zeta_K(2)}.$$

In view of (4) and the definition of f, we may rewrite this as

$$\int_{G/\Gamma}\mathcal{N}(P^mB,h)d\mu(B)=\frac{q^{2(1-g+h)+m}J(K)}{(1-q^{-2})(q-1)\zeta_K(2)}. \tag{5}$$

Since the coset $P^mG\subset \mathrm{GL}_2(K_{\mathbb{A}})$ is the subset of those B with $\deg\mathrm{div}(\det(B))=m$, we have the following interpretation of (5).

Lemma 4. *Fix $z,m\in\mathbb{Z}$. Then the mean value $\hat{\mathcal{N}}$ of $\mathcal{N}(B,z)$ over all $B\in\mathrm{GL}_2(K_{\mathbb{A}})/\mathrm{GL}_2(K)$ with $\deg\mathrm{div}(\det(B))=m$ satisfies*

$$\hat{\mathcal{N}}=\frac{q^{2(1-g+z)+m}J(K)}{(1-q^{-2})(q-1)\zeta_K(2)}.$$

Combining Lemma 4 with the Lemmas 3 and 3' yields the following theorem.

Theorem. *Fix an $h\in\mathbb{Z}$ with $0\le h<N/2$ and fix an even integer $2m$. For every $B\in\mathrm{GL}_2(K_{\mathbb{A}})$ with $\deg\mathrm{div}(\det(B))=2m$ we have associated codes as in Lemma 3. All these codes have minimal distance at least $N-2h$. Furthermore, the mean value over all such B (in the sense of Lemma 4) of the number of codewords in such codes is exactly*

$$\frac{q^{2(1-g+h)}J(K)}{(1-q^{-2})(q-1)\zeta_K(2)}.$$

Similarly fix an $h'\in\mathbb{Z}$ with $1/2\le h'<(N+1)/2$ and fix an odd integer $2m+1$. For every $B\in\mathrm{GL}_2(K_{\mathbb{A}})$ with $\deg\mathrm{div}(\det(B))=2m+1$ we have associated codes as in Lemma 3'. All these codes have minimal distance at least $N-2h'+1$. Furthermore, the mean value over all such B (in the sense of Lemma 4) of the number of codewords in such codes is exactly

$$\frac{q^{2(1-g+h')-1}J(K)}{(1-q^{-2})(q-1)\zeta_K(2)}.$$

In a manner analogous to the proof of Proposition 2.3.26 of [TV], it is a straightforward computation to show that

$$\log_q\left(\frac{q^{2(1-g+h)}J(K)}{(1-q^{-2})(q-1)\zeta_K(2)}\right)=2h-g+N\log_q\left(\frac{q+1}{q}\right)-o(g).$$

This is valid for any curve, including any curve in an asymptotically optimal family of curves (these are precisely the curves Elkies uses to construct his codes). This equation shows that codes in this superset have the same number of codewords on average as Elkies' nonlinear codes (cf. [E, equation (9)]). Hence, Elkies' two approaches (see [E,§1.3]) for comparing his nonlinear codes to Goppa codes apply verbatim to comparing codes in this superset to Goppa codes.

4. Equivalent Codes

In this section we make some observations on the codes we get from Lemmas 3 and 3' using the same $B \in \mathrm{GL}_2(K_{\mathbb{A}})$, and how different Bs can give rise to the same codes. We first notice that it suffices to look solely at $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ with $|\det(B)|_{\mathbb{A}} = 1$ or q for our codes. The following is clear from equation (0').

Lemma 5. *Let $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ suppose that $c \in K_{\mathbb{A}}^{\times}$ and $U \in \mathrm{GL}_2(K_{\mathbb{A}})$ is norm-preserving such that $cUB = A$ is of the form* (1)*. Choose an idele a_0 with $|a_0|_{\mathbb{A}} = q$ and let*

$$m = \left[\frac{\deg \operatorname{div}(\det(B))}{2}\right], \qquad B' = a_0^{-m} B.$$

(Here the brackets [ ] denote the greatest integer function.) Then $|\det(B')|_{\mathbb{A}} = 1$ or q, depending on whether $\deg \operatorname{div}(\det(B))$ is even or odd, respectively. Further, $(a_0^m c)UB' = A$ and $h_{B'}(\mathbf{x}) = h_B(\mathbf{x}) + m$ for all non-zero $\mathbf{x} \in K^2$. In particular, the codes obtained from B in Lemma 3 or 3' are exactly the codes obtained from B'.

We now consider the different possible codes one can get from a given $B \in \mathrm{GL}_2(K_{\mathbb{A}})$. Of course, starting with such a B, there are many different norm-preserving matrices U and ideles c for which cUB is of the form (1). Suppose c_1U_1B are c_2U_2B are two such choices. Then $c_1^{-1}c_2U_2U_1^{-1}$ is yet another such matrix of the form (1). Let us write

$$c_1^{-1}c_2U_2U_1^{-1} = \begin{pmatrix} u_1 & u_2 \\ 0 & 1 \end{pmatrix}.$$

Since $U_2U_1^{-1}$ is a norm-preserving upper triangular matrix, its lower diagonal entry, $c_1c_2^{-1}$, is an idele with v-adic modulus 1 at all places v. Thus $|c_{1,v}|_v = |c_{2,v}|_v$ for all $v \in M(K)$. Since the upper diagonal entry of $U_2U_1^{-1}$ is also an idele with v-adic modulus 1 at all places v, we see that $|u_{1,v}|_v = 1$ for all $v \in M(K)$. Similarly, we see that $|u_{2,v}|_v \leq 1$ for all $v \in M(K)$.

In summary, if $C_A(h-m-\deg\operatorname{div}(c))$ and $C_{A'}(h-m-\deg\operatorname{div}(c'))$ are two codes arising from $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ as in Lemma 3, then we have $A' = UA$ for some

$$U = \begin{pmatrix} u_1 & u_2 \\ 0 & 1 \end{pmatrix}, \qquad |u_{1,v}|_v = 1 \text{ and } |u_{2,v}|_v \leq 1 \text{ all } v \in M(K) \tag{6}$$

and we have $|c_v|_v = |c'_v|_v$ for all $v \in M(K)$. Similarly, if $C_A(h-1-m-\deg\operatorname{div}(c))$ and $C_{A'}(h-1-m-\deg\operatorname{div}(c'))$ are two codes arising from $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ as in Lemma 3', then we have $A' = UA$ for U of the form (6) and $|c_v|_v = |c'_v|_v$ for all $v \in M(K)$.

For the following lemma we adopt the usual conventions regarding arithmetic with ∞: $\infty + g = \infty$ for all $g \in \mathbb{F}_q$ and $f\infty = \infty$ for all $f \in \mathbb{F}_q^\times$.

Lemma 6. *Suppose $h \in \mathbb{Z}$ and $A \in \mathrm{GL}_2(K_{\mathbb{A}})$ is of the form (1). If $U \in \mathrm{GL}_2(K_{\mathbb{A}})$ is of the form (6) then there are $f_1, \ldots, f_N \in \mathbb{F}_q^\times$ and $g_1, \ldots, g_N \in \mathbb{F}_q$ such that $(x_1, \ldots, x_N) \in C_A(h)$ if and only if $(x_1 f_1 + g_1, \ldots, x_N f_N + g_N) \in C_{UA}(h)$.*

Conversely, if $f_1, \ldots, f_N \in \mathbb{F}_q^\times$ and $g_1, \ldots, g_N \in \mathbb{F}_q$, then there is a $U \in \mathrm{GL}_2(K_{\mathbb{A}})$ of the form (6) such that $(x_1, \ldots, x_N) \in C_A(h)$ if and only if $(x_1 f_1 + g_1, \ldots, x_N f_N + g_N) \in C_{UA}(h)$.

Proof. Suppose U is a matrix of the form (6). Set f_i to be the residue of u_{1,v_i} and g_i to be the residue of u_{2,v_i} for each $i = 1, \ldots, N$. Then by construction, $(x_1, \ldots, x_N) \in C_A(h)$ if and only if $(x_1 f_1 + g_1, \ldots, x_N f_N + g_N) \in C_{UA}(h)$.

Conversely, suppose the f_is and g_is are given as above. Choose a $u_1 \in K_{\mathbb{A}}^\times$ where the residue of u_{1,v_i} is f_i for each $i = 1, \ldots, N$ and $u_{1,v} = 1$ for all other places v. Choose a u_2 in a similar manner using the g_is. Then

$$U = \begin{pmatrix} u_1 & u_2 \\ 0 & 1 \end{pmatrix}$$

is of the form (6) and $(x_1, \ldots, x_N) \in C_A(h)$ if and only if $(x_1 f_1 + g_1, \ldots, x_N f_N + g_N) \in C_{UA}(h)$.

5. Codes from Linear Equations and Final Remarks

We consider a single homogeneous linear equation in three variables with coefficients in K:

$$\mathbf{c} \cdot \mathbf{Y} = c_1 Y_1 + c_2 Y_2 + c_3 Y_3 = 0, \qquad \mathbf{c} \in K^3 \setminus \{\mathbf{0}\}. \tag{7}$$

This equation defines a two-dimensional subspace of K^3. Take a basis $\mathbf{y}_1,\ \mathbf{y}_2$ of this subspace, so that any solution to (7) may be written uniquely as a linear combination of $\mathbf{y}_1$ and $\mathbf{y}_2$. We need a

basis such that $2h_I(\mathbf{y}_1) > N + h_I(\mathbf{c})$, where $I \in \mathrm{GL}_3(K_{\mathbb{A}})$ denotes the identity. Given any non-zero solution $\mathbf{y} = x_1\mathbf{y}_1 + x_2\mathbf{y}_2$, we may take its height $H_I(\mathbf{y})$ as an element of K^3. By [RT, Proposition 4.2] and [RT, Theorem 1.1], there is a $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ satisfying

$$\deg \mathrm{div}(\det(B)) = -h_I(\mathbf{c}) \tag{8}$$

and

$$h_I(x_1\mathbf{y}_1 + x_2\mathbf{y}_2) = h_B(x_1, x_2) \tag{9}$$

for all non-zero $\mathbf{x} \in K^2$.

This allows us to reformulate Lemmas 3 and 3' from the standpoint of solutions to a homogeneous linear equation.

Lemma 7. *Fix an $h \in \mathbb{Z}$ with $0 \le h < N/2$. Suppose $\mathbf{c} \in K^3 \setminus \{\mathbf{0}\}$ satisfies $h_I(\mathbf{c}) = -2m$ for $m \in \mathbb{Z}$. Let $\mathbf{y}_1,\ \mathbf{y}_2$ is a basis for (7) with $h_I(\mathbf{y}_1) \ge \frac{N}{2} - m$. Get a $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ satisfying conditions (8) and (9) and a matrix $A = cUB$ of the form (1) where $c \in K_{\mathbb{A}}^{\times}$ and $U \in \mathrm{GL}_2(K_{\mathbb{A}})$ is norm-preserving. Then $C_A(h - m - \deg \mathrm{div}(c))$ is an error-correcting code with minimal distance at least $N - 2h$. The number of codewords is exactly the number of one-dimensional subspaces $K\mathbf{y}$ of the solution space with $h_I(\mathbf{y}) \le h - m$.*

Lemma 7'. *Fix an $h \in \mathbb{Z}$ with $1/2 \le h < (N+1)/2$. Suppose $\mathbf{c} \in K^3 \setminus \{\mathbf{0}\}$ satisfies $h_I(\mathbf{c}) = -2m-1$ for $m \in \mathbb{Z}$. Let $\mathbf{y}_1,\ \mathbf{y}_2$ is a basis for (7) with $h_I(\mathbf{y}_1) \ge \frac{N-1}{2} - m$. Get a $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ satisfying conditions (8) and (9) and a matrix $A = cUB$ of the form (1) where $c \in K_{\mathbb{A}}^{\times}$ and $U \in \mathrm{GL}_2(K_{\mathbb{A}})$ is norm-preserving. Then $C_A(h - 1 - m - \deg \mathrm{div}(c))$ is an error-correcting code with minimal distance at least $N - 2h + 1$. The number of codewords is exactly the number of one-dimensional subspaces $K\mathbf{y}$ of the solution space with $h_I(\mathbf{y}) \le h - 1 - m$.*

We said above that every homogeneous linear equation (7) gives rise to such a $B \in \mathrm{GL}_2(K_{\mathbb{A}})$. By [T, Theorem 5], for any $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ there is a $c \in K_{\mathbb{A}}^{\times}$ such that cB arises from such an equation. Note how our choice of basis corresponds to choosing a representative B modulo $\mathrm{GL}_2(K)$. Thus, the codes in Lemmas 7 and 7' are precisely the codes in Lemma 3 and 3'. In particular, Elkies' nonlinear codes can be viewed as coming from linear equations of the form (7). Not only that, but

"on average," in the sense of Lemma 4, the codes generated by equations (7) with $h_I(\mathbf{c})$ even have minimal distance at least $N - 2h$ and

$$\frac{q^{2(1-g+h)}J(K)}{(1-q^{-2})(q-1)\zeta_K(2)}$$

codewords. The equations where $h_I(\mathbf{c})$ is odd will give codes with minimal distance at least $N - 2h + 1$ and

$$\frac{q^{2(1-g+h)-1}J(K)}{(1-q^{-2})(q-1)\zeta_K(2)}$$

codewords, on average.

We end with some final remarks. First, one could well ask if our mean value in Lemma 4 is typical of $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ or whether one can reasonably expect $\mathcal{N}(B, h)$ to be much larger or smaller that the mean. One approach to this problem which has been carried out (to some extent) for the field of rational numbers is to derive higher moments. To our knowledge, this has not been done for function fields. We do have heuristic arguments which indicate that, indeed, the mean value is quite typical.

Finally, for our transmission rate estimates we used only a special case of the machinery in [T]; specifically, we used the "convex body" $S = \prod_v \mathfrak{O}_v^2$. The mean value (Lemma 4) can be computed equally well for any "star convex" S. Perhaps one could construct codes via a different choice of S which would be more efficient.

References

[E] N. Elkies, *Excellent nonlinear codes from modular curves*, STOC'01: Proceedings of the 33rd Annual ACM Symposium on the Theory of Computing, Hersonissos, Crete, Greece (2001), 200-208.

[RT] D. Roy and J. Thunder, *An absolute Siegel's lemma*, J. reine angew. Math. **476** (1996), 1-26.

[T] J. Thunder, *An adelic Minkowski-Hlawka theorem and an application to Siegel's lemma*, J. reine angew. Math. **475** (1996), 167-185.

[TV] M. A. Tsfasman and S. G. Vlăduţ, *Algebraic-Geometric Codes*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.

Department of Mathematics, Northern Illinois University, DeKalb, IL 60115
E-mail address: hurlburt@math.niu.edu

Department of Mathematics, Northern Illinois University, DeKalb, IL 60115
E-mail address: jthunder@math.niu.edu