6.897 Algorithmic Introduction to Coding Theory September 12, 2001

Lecture Notes on Algebra

Lecturer: Madhu Sudan *Scribe: Madhu Sudan*

These notes describe some basic algebraic structures that we will encounter during this course, including:

- Finite fields of all sizes (and shapes).
- (Univariate and multivariate) polynomials over finite fields in one or more variables.
- Vector spaces over finite fields (or Linear algebra).

Unfortunately, there is no simple order in which one can present all these objects — their presentation is interleaved for essential reasons. Polynomials are typically defined with coefficients from fields. Fields are constructed by constructing polynomial rings and then reducing them modulo irreducible polynomials. Linear algebra needs to be based on fields. But it also provides convenient ways of looking at fields. We will try to describe all these connections below. Mostly we are interested in computational and combinatorial consequences. We would like to see how to represent fields so as to perform elementary manipulations efficiently. We would also like to know if some computational problems from linear algebra can be solved efficiently. We are also interested in combinatorial questions such as: How often can a polynomial evaluate to zero? How does one prove that this can not happen too often? The notes below present answers to such questions.

1 Main definition

Since we are interested in polynomials over fields, it would be nice to know the basic algebraic structures which unify both fields and polynomials. Commutative rings are such structures and we define them below.

Definition 1 (Commutative Rings and Fields) *A* commutative ring *is given by a triple $(R, +, \cdot)$, where R is an arbitrary set containing two special elements 0 and 1 and $+, \cdot$ are functions mapping $R \times R$ to R satisfying the following properties for every triple $a, b, c \in R$:*

Associativity: *Both $+$ and $\cdot$ are associative, i.e., $a+(b+c) = (a+b)+c$ and $a\cdot(b\cdot c) = (a\cdot b)\cdot c$.*

Commutativity *Both $+$ and $\cdot$ are commutative, i.e., $a + b = b + a$ and $a \cdot b = b \cdot a$.*

Distributivity: *$\cdot$ distributes over $+$, i.e., $a \cdot (b + c) = a \cdot b + a \cdot c$.*

Identities: *$a + 0 = a$ and $a \cdot 1 = a$.*

Additive Inverses: *For every $a \in R$, there exists an additive inverse $-a \in R$ such that $a + (-a) = 0$.*

If in addition, every non-zero element has a multiplicative inverse, then R is a field. *(I.e., for every $a \in R - \{0\}$, there exists an $a^{-1} \in R$ such that $a \cdot a^{-1} = 1$.)*

Often we will skip the operators $+$ and $\cdot$ and simply refer to the set R as the ring (with addition and multiplication being specified implicitly). Commutative rings form the foundation for much of the elegant results of algebra and algebraic geometry. Within the class of commutative rings, one can get nicer and nicer domains (rings with nicer and nicer properties) and this culminates with the notion of a field. Informally, rings allow the operations of addition, subtraction and multiplication, while a field also allows division. We will see some of the intermediate notions later. Right now we turn to polynomials.

2 Polynomial rings

Given any ring R, and a symbol t (usually referred to as an indeterminate, one can create a ring $R[t]$ of polynomials over R. Such a ring inherits most of the nice properties of the underlying ring. Below is a formal definition of the ring of polynomials.

Definition 2 *Given a commutative ring R and indeterminate t, the set $R[t]$ has as its elements finite sequences of R, with the sequence $f = \langle f_0, \ldots, f_l \rangle$ being interpreted as the formal sum $\sum_{i=0}^{l} f_i t^i$. Addition and multiplication over $R[t]$ are defined accordingly, i.e., if $f = \langle f_0, \ldots, f_l \rangle$ and $g = \langle g_0, \ldots, g_k \rangle$ with $l \leq k$ then $f + g = \langle f_0 + g_0, \ldots, f_l + g_l, g_{l+1}, \ldots, g_k \rangle$ and $f \cdot g = \langle h_0, \ldots, h_{l+k}$ where $h_i = \sum_{j=0}^{i} f_j g_{i-j}$. For a polynomial $f \in R[t]$, given by $f = \sum_{i=0}^{d} f_i t^i$, we define its* degree, *denoted* $\deg(f)$ *to be the largest index d' such that $f_{d'}$ is non-zero.*

Proposition 3 *For every commutative ring R and indeterminate t, $R[t]$ is a commutative ring.*

The most natural ring of polynomials that we will encounter are the ring of polynomials over some (finite) field F, say $F[x]$. Now we can adjoin a new indeterminate y to this ring to another ring $F[x][y]$. We will use the notation $F[x, y]$ to denote such a ring whose elements are simply polynomials in two variables x and y. In particular $F[y][x] = F[x][y] = F[x, y]$. Continuing this way, adjoining m variables $x_1, \ldots, x_m$ to F for some integer m, we get the space of m-variate polynomials $F[x_1, \ldots, x_m]$. It is also possible to define this ring directly and we do so in order to define various notions of degree associated with it.

Definition 4 (Multivariate polynomial rings) *Given a ring R and indeterminates $x_1, \ldots, x_m$ the m-variate polynomial ring over R, denoted $R[x_1, \ldots, x_m]$ has as its elements finite sequences indexed by d-tuple of non-negative integers $f = \langle f_{i_1, \ldots, i_m} \rangle_{0 \leq i_j \leq d_j}$. The element represents the formal sum $\sum_{i_1, \ldots, i_m} f_{i_1, \ldots, i_m} x_1^{i_1} \cdots x_m^{i_m}$. Addition and multiplication are interpreted appropriately. The x_j-*degree *of f, denoted* $\deg_{x_j}(f)$, *is the largest index d'_j such that there exist indices $i_1, \ldots, i_m$ such that $f_{i_1, \ldots, i_{j-1}, d_j, i_{j+1}, \ldots, i_m}$ is non-zero. The* total degree *of f is the largest sum $\sum_{j=1}^{m} i_j$, among tuples $i_1, \ldots, i_m$ for which $f_{i_1, \ldots, i_m}$ is non-zero.*

We will come back to multivariate polynomials later. Right now we move on to descriptions of fields and this will need univariate polynomials.

3 Finite Fields

Fields are the nicest of algebraic structures. that allow all sorts of manipulations efficiently. In particular we can not only define addition and multiplication, but also subtraction $(a - b = a + (-b))$

and division ($a/b = a \cdot b^{-1}$). The most familiar examples of fields are the field of rational numbers $\mathcal{Q}$ and the field of real numbers $\mathbb{R}$. For our purposes fields that have only a finite number of elements are much more important. The following theorem tells us what kind of finite fields exist.

Theorem 5 *For a positive integer q, a field F of cardinality q exists if and only if $q = p^l$ for a prime p and positive integer l.*

We use the notation $\mathbb{F}_q$ to denote the field with q elements. Since we eventually intend to use the fields computationally, we will need to know a little more about such fields. Specifically, given q how can one represent the elements of the field $\mathbb{F}_q$? Given (such representations of elements) $\alpha, \beta \in \mathbb{F}_q$, how can we compute (representation of) $\alpha + \beta$, $\alpha - \beta$, $\alpha \cdot \beta$ and α/β? We answer these questions below:

Prime fields. If $q = p$ for a prime p, then the field $\mathbb{F}_q$ is simply the field of arithmetic modulo p. Thus the natural way to represent the elements of $\mathbb{F}_q$ is using the integers $\{0, \ldots, p-1\}$. It is easy to carry out addition, multiplication, and subtraction in the field can be carried out in time $\text{poly}\log q$. Fermat's little theorem also tell us that if $\beta \neq 0$, then $\beta^{-1} = \beta^{p-2} (\mod p)$ and by using a fast modular exponentiation algorithm, β^{-1} can also be computed in time $\text{poly}\log q$. (Actually addition, multiplication, and subtraction can be computed in time $O(\log q \text{poly}\log\log q)$.)

Before going on to describing fields of cardinality p^l, where $l \geq 2$, we need to define the notion of irreducible polynomials.

Definition 6 (Irreducible polynomials) *Given a ring $F[t]$ of polynomials over a field F, a polynomial $f \in F[t]$ is said to be* reducible *if there exist polynomials g and h in $F[t]$ of degree at least one such that $f = g \cdot h$. f is said to be irreducible if no such polynomial exists.*

We are now ready to describe the remaining finite fields.

Prime power fields. Let $q = p^l$ for prime p and positive integer l. Suppose f is an irreducible polynomial of degree l in $\mathbb{F}_p[t]$. Then $\mathbb{F}_q \cong \mathbb{F}_p[t]/(f)$, i.e., the ring of polynomials in t reduced modulo f. Specifically, the elements of $\mathbb{F}_p[t]/(f)$ are polynomials of degree strictly less than l. (Note that there are exactly p^l such polynomials.) Addition is straightforward polynomial addition. (Note that the degree of the sum is less than l if both polynomials have degree less than l.) Multiplication is performed modulo f, i.e., given g and h we compute $p = gh$ using regular polynomial multiplication and then compute the remainder when p is divided by f. This is a polynomial r of degree less than l and we define $g \cdot h$ to be r in the "field" $\mathbb{F}_p[t]/(f)$. Fermat's little theorem applied to groups shows that $g^{-1} = g^{q-2}$ in this field also.

Exercise: Verify that $\mathbb{F}_q$ as described above satisfies the definitions of a field.

The above construction is would not be very useful, if it weren't for the fact that irreducible polynomials exist and can be found efficiently.

Theorem 7 (cf. [5]) *For every prime p and positive integer l, there exists an irreducible polynomial of degree l over $\mathbb{F}_p[t]$. Furthermore such a polynomial can be found deterministically in time* $\text{poly}(l, p)$ *and probabilistically in expected time* $\text{poly}(l, \log p)$.

Given the above we see that we can pre-compute a representation of a field in expected time poly $\log q$ and then perform all field operations deterministically in time poly $\log q$. In certain scenarios it may be useful to have irreducible polynomial explicitly. In $\mathbb{F}_2[t]$ an infinite sequence of such polynomials is known (cf. [6, Theorem 1.1.28]).

Theorem 8 *For every $l \geq 0$, the polynomial $x^{2\cdot 3^l} + x^{3^l} + 1$ is irreducible over $\mathbb{F}_2[x]$.*

Thus we can construct fields of size $2^{2\cdot 3^l}$ for every integer l totally explicitly. Thus if we were interested in a field of size at least $q = 2^m$, and m is not of the form $2 \cdot 3^l$, we can find an m' of the right form with $m' < 3m$ and the resulting field would be of size less than q^3, which is only polynomially larger than our lower bound.

4 Evaluations of polynomials

We introduced polynomials merely as formal sums — syntactic expressions with no semantics associated with them. Evaluations associate some semantics to them.

Definition 9 *The* evaluation *of a polynomial $f = \sum_{i=0}^{d} f_i t^i \in R[t]$ at the point $\alpha \in R$, denoted $f(\alpha)$, is given by $\sum_{i=0}^{d} f_i \alpha^i$. Evaluations of multivariate polynomials are defined analogously; the evaluation of $f \in R[x_1, \ldots, x_m]$ and $\alpha = \langle \alpha_1, \ldots, \alpha_m \rangle$ is denoted $f(\alpha)$ or $f(\alpha_1, \ldots, \alpha_m)$.*

Evaluations carry natural semantics, i.e., $f(\alpha) + g(\alpha) = (f+g)(\alpha)$ and $f(\alpha) \cdot g(\alpha) = (f \cdot g)(\alpha)$. We are interested in knowing how often a polynomial can evaluate to zero. To answer this question, we first introduce the notion of division of polynomials.

Proposition 10 (Division Algorithm) *Given polynomials $f, g \in F[t]$ for some field F, there exists a unique pair of polynomials q and r (for quotient and remainder) in $F[t]$ satisfying $\deg(r) < \deg(g)$ and $f = g \cdot q + r$. Further the polynomial q satisfies $\deg(q) = \deg(f) - \deg(g)$ if $\deg(f) \geq \deg(g)$.*

The name of the proposition above is due to the fact that the proposition is proved by simply performing long division in the usual manner. Applying the above proposition with $g = t - \alpha$ for some $\alpha \in F$, we get that $f = q \cdot (t - \alpha) + r$ where r has degree zero and hence $r \in F$. Furthermore evaluating the expression above at α yields $f(\alpha) = q(\alpha) \cdot (\alpha - \alpha) + r$, and thus $r = f(\alpha)$. Thus we have $f = q \cdot (t - \alpha) + f(\alpha)$ for some q of degree $\deg(f) - 1$. The following proposition then follows.

Proposition 11 *The polynomial $t - \alpha$ divides f if and only if $f(\alpha) = 0$.*

Thus we get that if distinct elements $\alpha_1 \ldots, \alpha_k$ are all zeroes of a polynomial f (i.e., $f(\alpha_i) = 0$ for $i \in [k]$) then the polynomial $h = \prod_{i=1}^{k} (t - \alpha_i)$ divides f. Since the degree of h is k it follows that the degree of f is at least k. So we get:

Theorem 12 *For a field F, an element $f \in F[t]$ evaluates to zero on at most $\deg(f)$ points in F.*

We now move onto estimating the number of zeroes of multivariate polynomials. To do so, we need a variant of Theorem 12 for multivariate polynomials. We obtain such a result by expanding the scope of the theorem above. We first need a definition.

Definition 13 *A commutative ring R is an* integral domain *if it does not contain any zero divisors, i.e., there do not exist non-zero elements $a, b \in R$ such that $a \cdot b = 0$.*

Note that the rings $F[x_1, \ldots, x_m]$ are integral domains. Integral domains are of interest in that they are almost as nice as fields. Specifically, the following construction gives a field that contains any given integral domain.

Definition 14 *For an integral domain R, its* field of fractions*, denoted $\tilde{R}$, is the ring whose elements are pairs (a, b) with $a \in R$ and $b \in R - \{0\}$, modulo the equivalence $(a, b) \cong (c, d)$ if $a \cdot d = b \cdot c$. The element (a, b) is interpreted as the ratio a/b. Addition and multiplication are defined analogously with $(a, b) + (c, d) = (a \cdot d + b \cdot c, b \cdot d)$ and $(a, b) \cdot (c, d) = (a \cdot c, b \cdot d)$.*

The following proposition is easily verified.

Proposition 15 *For every integral domain R, $\tilde{R}$ is a field. Further, R is contained in $\tilde{R}$.*

The field of fractions of $F[x_1, \ldots, x_m]$ is usually denoted $F(x_1, \ldots, x_m)$ and its elements are the rational functions (ratios of polynomials) in $x_1, \ldots, x_m$.

The following lemma is now an easy consequence of the notions of integral domains, fields of fractions, and Theorem 12.

Lemma 16 *The polynomial $g(\mathbf{x}) \in F[\mathbf{x}]$ is a zero of the polynomial $f(\mathbf{x}, t) \in F[\mathbf{x}][t]$ (i.e., $f(\mathbf{x}, g(\mathbf{x})) = 0$) if and only if $t - g(\mathbf{x})$ divides the polynomial $f(\mathbf{x}, t)$. Hence the polynomial f has at most $\deg_t(f)$ zeroes in $F[x_1, \ldots, x_m]$.*

Proof The proof is simple. We simply view f as a polynomial in $K[t]$, where $K = F(x_1, \ldots, x_m)$ is the field of fractions of $F[x_1, \ldots, x_m]$. Since $g \in K$, it follows that $f(g) = 0$ iff $t - g$ divides f (from Proposition 11). Furthermore, Theorem 12, applied to $f \in K[t]$, says that f has at most $\deg_t(f)$ roots in K, which contains $F[x_1, \ldots, x_m]$. ■

The following theorem is now an easy inductive consequence of Lemma 16.

Theorem 17 *A non-zero polynomial $f \in \mathbb{F}_q[x_1, \ldots, x_m]$ is non-zero on at least $\prod_{i=1}^{m}(q - \deg_{x_i}(f))$ points in $\mathbb{F}_q^m$.*

Proof We start by viewing $\mathbb{F}_q[x_1, \ldots, x_m]$ as a polynomial in $\mathbb{F}_q[x_1, \ldots, x_{m-1}][x_m]$. By Lemma 16, there are at most $\deg_{x_m}(f)$ choices of $\alpha_m \in \mathbb{F}_q[x_1, \ldots, x_{m-1}]$ and hence in $\mathbb{F}_q$ such that $f(x_1, \ldots, x_{m-1}, \alpha_m) = 0$. For an α_m such that $f(x_1, \ldots, x_{m-1}, \alpha_m) \neq 0$ (and there exist $q - \deg_{x_m}(f)$ such α_m's), let $f_{\alpha_m}(x_1, \ldots, x_{m-1}) = f(x_1, \ldots, x_{m-1}, \alpha_m)$. Since f_{α_m} is a polynomial in $F[x_1, \ldots, x_{m-1}]$ with $\deg_{x_i}(f_{\alpha_m}) \leq \deg_{x_i}(f)$, it follows (by induction on m) that f_{α_m} is non-zero on at least $\prod_{i=1}^{m-1}(q - \deg_{x_i}(f))$ points in $\mathbb{F}_q^{m-1}$. The theorem follows. ■

We can derive other variants of the theorem above. One such variant that is quite popular in the CS community, often termed Schwartz's lemma or the DeMillo-Lipton-Schwartz-Zippel lemma [1, 4, 7], is the following:

Theorem 18 *A non-zero polynomial $f \in \mathbb{F}_q[\mathbf{x}]$ of total degree d is zero on at most a d/q fraction of the points in $\mathbb{F}_q^m$.*

Remark: Both Theorems 17 and 18 can also be extended to count the number of zeroes in some space of the form S^m for $S \subseteq \mathbb{F}_q$, but we don't do so here.

Proof The proof again goes by induction, but this time in the reverse order. Let d_m be the degree of f in x_m and let $f_m \in F[x_1, \ldots, x_{m-1}]$ be the coefficient of x^{d_m} in f, where we view f as an element of $F[x_1, \ldots, x_{m-1}][x_m]$. Note that the total degree f_m is at most $d - d_m$. For a random choice of $\alpha_1, \ldots, \alpha_{m-1} \in \mathbb{F}_q$, by induction we have that the probability that $f_m(\alpha_1, \ldots, \alpha_{m-1}) = 0$ is at most $(d - d_m)/q$. If this event happens, then we give up (and assume $f(\alpha_1, \ldots, \alpha_{m-1}, \alpha_m) = 0$ for all α_m). Else we get a polynomial $g(x_m) \stackrel{\text{def}}{=} f(\alpha_1, \ldots, \alpha_{m-1}, x_m)$ in one variable of degree d_m. By Theorem 12 a random choice of α_m is a zero of g with probability at most d_m/q. For $f(\alpha_1, \ldots, \alpha_m)$ to be zero, it must be the case that $f_m(\alpha_1, \ldots, \alpha_{m-1}) = 0$ or $g(\alpha_m) = 0$, Thus by the union bound, we find that $f(\alpha_1, \ldots, \alpha_m) = 0$ with probability at most $(d - d_m)/q + d_m/q = d/q$. ■

Finally we describe yet another variant that was the version used in the classical Reed-Muller codes.

Theorem 19 *If $f \in F_q[x_1, \ldots, x_m]$ has individual degree at most l in each variable and has total degree $d = lk + r$, then it is non-zero on at least $(1 - l/q)^k(1 - r/q)$ fraction of the inputs from $\mathbb{F}_q^m$.*

The proof of the theorem is a simple variant of the two proofs above and so we won't repeat it. It is more interesting to see a consequence. Suppose we have a multilinear polynomial of total degree k over $\mathbb{F}_2$. Then it is non-zero on at least a 2^{-k} fraction of the domain $\mathbb{F}_2^m$. This is exactly the kind of result that was used in the original Reed-Muller codes.

5 Vector spaces over fields

Here we relay some basic definitions and results about linear algebra that form the basis of linear codes.

We will be considering subspaces of the vector space $\mathbb{F}_q^n$, which is endowed with an addition operator "+" and a scalar-vector product "·", where if $\mathbf{x} = \langle x_1, \ldots, x_n \rangle$ and $\mathbf{y} = \langle y_1, \ldots, y_n \rangle$, then $\mathbf{x} + \mathbf{y} = \langle x_1 + y_1, \ldots, x_n + y_n \rangle$ and for $\alpha \in \mathbb{F}_q$, $\alpha \cdot \mathbf{x} = \langle \alpha \cdot x_1, \ldots, \alpha \cdot x_n \rangle$.

Definition 20 (Linear subspace) *A subset $L \subseteq \mathbb{F}_q^n$ is a linear subspace of $\mathbb{F}_q^n$ if for every $\mathbf{x}, \mathbf{y} \in L$ and every $\alpha \in \mathbb{F}_q$ it is the case that $\mathbf{x} + \mathbf{y} \in L$ and $\alpha \cdot \mathbf{x} \in L$.*

Definition 21 (Basis, Dimension) *The* span *of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$, denoted* $\text{span}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$*, is the set $\{\sum_{i=1}^k \alpha_i \cdot \mathbf{x}_i \mid \alpha_1, \ldots, \alpha_k \in \mathbb{F}_q\}$. A set of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$ is* linearly independent *if $\sum_{i=1}^k \alpha_i \cdot \mathbf{x}_i = \mathbf{0}$ implies $\alpha_1 = \cdots = \alpha_k = 0$. For a linear subspace $L \subseteq \mathbb{F}_q^n$, a set of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k \in L$ forms a* basis *if the vectors are linearly independent and their span equals L. The* dimension *of L, denoted* $\dim(L)$*, is the size of the largest basis for L.*

Playing around with the definitions, one can show easily that the span of a set of vectors is a linear subspace, that every linear subspace has a basis, and that all bases for a given subspace have the same size. One way to describe a linear subspace is to give its basis. A different way is to give constraints satisfied by elements of the subspace. We move towards this notion next. To do so, we need the notion of an dot product of vectors. For $\mathbf{x} = \langle x_1, \ldots, x_n \rangle$ and $\mathbf{y} = \langle y_1, \ldots, y_n \rangle$, the dot product of $\mathbf{x}$ and $\mathbf{y}$, denoted $\langle \mathbf{x}, \mathbf{y} \rangle$, equals $\sum_{i=1}^{n} x_i y_i$.

Definition 22 (Null Space) *For a linear subspace $L \subseteq \mathbb{F}_q^n$, it null space, denoted $L^\perp$, is the set $\{\mathbf{y} \in \mathbb{F}_q^n \mid \langle \mathbf{x}, \mathbf{y} \rangle = 0\}$.*

Proposition 23 *The null space of a linear subspace $L \subseteq \mathbb{F}_q^n$ is also a linear subspace of $\mathbb{F}_q^n$ and has dimension $n - \dim(L)$.*

A full proof of this assertion turns out to be somewhat complicated, and seems to involve proving the well-known but non-trivial fact that the row rank of a matrix equals its column rank. Instead of proving this, we will give a sense of how the proof goes, by essentially giving an algorithm to compute the basis of the null space $L^\perp$, given a basis of the space L. The description in the following paragraph is not self-contained — reading this paragraph is not suited for all audiences. A better approach may be to read a chapter on linear algebra from a text on algebra (e.g., [3]).

Suppose $\mathbf{x}_1, \ldots, \mathbf{x}_k$ form a basis for L. Let $\mathbf{G} \in \mathbb{F}_q^{k \times n}$ be the matrix whose ith row is $\mathbf{x}_i$. Since $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are linearly independent, it follows that the rank of $\mathbf{G}$ is k. (Note that we didn't really define the rank of a matrix — there will be other such transgressions in this paragraph.) In particular this means there is an invertible $k \times k$ submatrix in $\mathbf{G}$. By permuting the columns of $\mathbf{G}$, we can write it as $\mathbf{G} = [\mathbf{A}|\mathbf{B}]$ where $\mathbf{A}$ is an invertible square matrix and $\mathbf{B}$ is the rest. Note that we are interested in a basis of the space $L^\perp = \{\mathbf{y} \mid \mathbf{y}\mathbf{G} = \mathbf{0}\}$. Writing all vectors $\mathbf{y} \in \mathbb{F}_q^n$ as $\mathbf{y} = (\mathbf{y}_A, \mathbf{y}_B)$ where $\mathbf{y}_A \in \mathbb{F}_q^k$ and $\mathbf{y}_B \in \mathbb{F}_q^{n-k}$, we get $\mathbf{y}_A\mathbf{A} + \mathbf{y}_B\mathbf{B} = \mathbf{0}$ for all $\mathbf{y} \in L^\perp$. This essentially yields $L^\perp = \{(-\mathbf{y}_B\mathbf{B}\mathbf{A}^{-1}, \mathbf{y}_B) | \mathbf{y}_B \in \mathbb{F}_q^{n-k}\}$, Taking $\mathbf{y}_B$ to be all the unit vectors gives $n - k$ vectors that generate $L^\perp$.

The correct way to think of the null space $L^\perp$ is that its members give linear constraints on the members of L. E.g., the vector $\mathbf{y} \in L^\perp$ enforces the constraint $\sum_{i=1}^{n} y_i x_i = 0$ on the vectors $\mathbf{x}$ in L. Since it suffices to satisfy the constraints given by any basis of $L^\perp$ (the other constraints get satisfied automatically), the basis of $L^\perp$ gives an alternate succinct representation of L.

We now move on to computational versions of the above results: Most of these computational results just build on the essential fact that Gaussian elimination works over any field (and not just rationals or reals or complexes).

Theorem 24 *Given a matrix $\mathbf{G}$ whose rows are the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{F}_q^n$, let $L = \mathrm{span}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$. Then, the following problems can be solved in time $O((n+k)^3)$:*

1. *For $n = k$, compute the determinant of $\mathbf{G}$.*
2. *Compute the rank of $\mathbf{G}$.*
3. *If $\mathrm{rank}\, \mathbf{G} = k$, then a pseudo-inverse matrix $\mathbf{G}^{-1}$ such that $\mathbf{G}\mathbf{G}^{-1} = \mathbf{I}_k$, where $\mathbf{I}_k$ is the $k \times k$ identity matrix.*
4. *Given a vector $\mathbf{b} \in \mathbb{F}_q^n$, compute a vector $\mathbf{x}$ and a matrix $\mathbf{H}$ such that the set $\{\mathbf{z}_\mathbf{y} = \mathbf{x} + \mathbf{y}\mathbf{H} \mid \mathbf{y}\}$ is the set of solutions to $\mathbf{z}\mathbf{G} = \mathbf{b}$ if such a solution exists.*

6 Representing fields by matrices

We have already encountered one representation for element of a field $\mathbb{F}_q$, where $q = p^l$ for some prime (power) p — namely, the elements of $\mathbb{F}_q$ are polynomials over $\mathbb{F}_p$. However it is often to know about other representations. Here we describe two (actually one and a half) representations.

The first representation is only semi-adequate in that it describes how to do addition in $\mathbb{F}_q$, but not how to multiply. But it is useful to get to the second representation. Further it is often useful to think of the two simultaneously. We now start with the simple representation.

Fields as vector spaces. A simple way to think of $\mathbb{F}_{p^l}$ is as $\mathbb{F}_p^l$ - i.e., field elements are just vectors over $\mathbb{F}_p$ and field addition is just vector addition. Formally, there is an invertible transformation $L : \mathbb{F}_{p^l} \to \mathbb{F}_p^l$, that $\alpha \in \mathbb{F}_{p^l}$ is represented by the element $L(\alpha) \in \mathbb{F}_p^l$. such that $L(\alpha+\beta) = L(\alpha) + L(\beta)$. However this representation does not give a clue on how to do field multiplication.

Field elements as linear transformations. One way to think of a field element is that the element α defines a map $\beta \mapsto \alpha \cdot \beta$. Now if we represent β by its linear representation, then we get that α is a *linear* map from $\mathbb{F}_p^l$ to $\mathbb{F}_p^l$. In other words, if we fix the linear representation L, then corresponding to α, we can define a map $M_\alpha : \mathbb{F}_p^l \to \mathbb{F}_p^l$, with $M_\alpha(L(\beta)) = L(\alpha \cdot \beta)$. Note that this map satisfies $M_\alpha(L(\beta) + L(\gamma)) = M_\alpha(L(\beta)) + M_\alpha(L(\gamma))$. Since this is a linear map, this says there is a matrix $\mathbf{M}_\alpha \in \mathbb{F}_p^{k\times k}$ such that $M_\alpha(\mathbf{x}) = \mathbf{x}\mathbf{M}_\alpha$. Furthermore, in this case we have $\mathbf{M}_{\alpha_1\cdot\alpha_2} = \mathbf{M}_{\alpha_1} \cdot \mathbf{M}_{\alpha_2}$. and $\mathbf{M}_{\alpha_1+\alpha_2} = \mathbf{M}_{\alpha_1} + \mathbf{M}_{\alpha_2}$. Thus the transformation $\alpha \mapsto \mathbf{M}_\alpha$ maps $\mathbb{F}_{p^l}$ to $\mathbb{F}_p^{k\times k}$ and has the property that addition and multiplication in the field are just addition and multiplication of matrices! This representation (of α by $\mathbf{M}_\alpha$) can be quite useful at times.

7 Conclusions

Not all the descriptions above were intended to be complete. The idea is to list facts that (a) are true and (b) are assumed to be true in this course. Hopefully we won't use stuff that is not in these notes — but if we do we will try to make explicit note of this later. If you are planning to learn from this class you should either (a) be completely comfortable with assuming the facts stated, or (b) read appropriate algebra texts to review the material on fields and linear algebra. Some recommendations (this may be expanded later) include the text by Lidl and Neidereitter on finite fields [2] and by MacLane and Birkhoff on algebra in general [3].

References

[1] Richard A. Demillo and Richard J. Lipton. A probabilistic remark on algebraic program testing. *Information Processing Letters*, 7(4):193–195, June 1978.

[2] Rudolf Lidl and Harald Niedereitter. *Introduction to Finite Fields and Their Applications.* Cambridge University Press, 2nd edition, 1994.

[3] Saunders MacLane and Garrett Birkhoff. *Algebra.* Chelsea Publishing Company, N.Y., 3rd edition, 1988.

[4] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27(4):701–717, October 1980.

[5] Victor Shoup. New algorithms for finding irreducible polynomials over finite fields. *Mathematics of Computation*, 54:435–447, 1990.

[6] Jacobus H. van Lint. *Introduction to Coding Theory.* Graduate Texts in Mathematics **86**, (Third Edition) Springer-Verlag, Berlin, 1999.

[7] Richard E. Zippel. Probabilistic algorithms for sparse polynomials. In *EUROSAM '79, Lecture Notes in Computer Science*, volume 72, pages 216–225, 1979.