ON DENSITY OF PRIMITIVE ELEMENTS FOR FIELD EXTENSIONS

JOEL V. BRAWLEY AND SHUHONG GAO

ABSTRACT. This paper presents an explicit bound on the number of primitive elements that are linear combinations of generators for field extensions.

It is well known that every finite separable extension of an arbitrary field has a primitive element; that is, there exists a single element in the extension field which generates that field over the ground field. This is a fundamental theorem in algebra which is called the *primitive element theorem* in many textbooks, see for example [3, 6, 7], and it is a useful tool in practical computation of commutative algebra [5]. The existence proofs found in the literature make no attempt at estimating the density of primitive elements. The purpose of this paper is to give an explicit lower bound on the density of primitive elements that are linear combinations of generators. Our derivation uses a blend of Galois theory (specifically the Fundamental Theorem of Galois Theory), basic linear algebra, and a simple form of the principle of inclusion-exclusion from elementary combinatorics. Additionally, for readers familiar with Grobner bases, we show by a geometric example, how to test a linear combination for primitivity without relying on the Galois groups used in deriving our bound.

Let $\mathbb{F}$ be any field and $\mathbb{K}$ a finite algebraic extension of $\mathbb{F}$. An element $\beta \in \mathbb{K}$ is called *primitive* for $\mathbb{K}$ over $\mathbb{F}$ if $\mathbb{K} = \mathbb{F}(\beta)$. Suppose $\mathbb{K}$ is generated by $\alpha_1, \ldots, \alpha_n$, that is,

$$\mathbb{K} = \mathbb{F}(\alpha_1, \ldots, \alpha_n).$$

Consider elements of the form

$$\beta = b_1\alpha_1 + \cdots + b_n\alpha_n$$

where $b_i \in \mathbb{F}$, $1 \leq i \leq n$. We would like to know when and how frequently such elements are primitive for $\mathbb{K}$ over $\mathbb{F}$ when the coefficients are required to come from an arbitrary finite subset of $\mathbb{F}$. An n-tuple $b = (b_1, \ldots, b_n)$ is called *primitive* with respect to $(\alpha_1, \ldots, \alpha_n)$ if the corresponding β is primitive for $\mathbb{K}$ over $\mathbb{F}$.

Date: February 2, 2004.

1991 *Mathematics Subject Classification.* Primary 12Y05; Secondary 12F10.

Key words and phrases. Primitive element theorem, Galois theory, probability.

The second author was supported in part by National Science Foundation (NSF) under Grant DMS0302549, National Security Agency (NSA) under Grant MDA904-02-1-0067, and the DoD Multidisciplinary University Research Initiative (MURI) program administered by the Office of Naval Research (ONR) under Grant N00014-00-1-0565.

Let S be a finite subset of $\mathbb{F}$ and S^n the set of all n-tuples with entries from S. Define

$$\rho(S;\alpha_1,\ldots,\alpha_n) = \frac{1}{|S|^n}\#\{b \in S^n : b_1\alpha_1 + \cdots + b_n\alpha_n \text{ is primitive for } \mathbb{K} \text{ over } \mathbb{F}\}.$$

Thus $\rho(S;\alpha_1,\ldots,\alpha_n)$ represents the density of primitive n-tuples among all the n-tuples over S. Equivalently, $\rho(S;\alpha_1,\ldots,\alpha_n)$ is the probability that β is primitive when $b \in S^n$ is chosen at random. Clearly $\rho(S;\alpha_1,\ldots,\alpha_n) > 0$ implies the existence of primitive elements.

Theorem. *Let $\mathbb{K}$ be a separable extension of degree m of a field $\mathbb{F}$ with $m > 1$. Suppose $\mathbb{K} = \mathbb{F}(\alpha_1,\ldots,\alpha_n)$. Then for any finite subset S of $\mathbb{F}$, we have*

$$\rho(S;\alpha_1,\ldots,\alpha_n) \geq 1 - \frac{m-1}{|S|}.$$

Proof. Our proof will come in steps. We first characterize when β generates $\mathbb{K}$ over $\mathbb{F}$, then we use linear algebra to translate this condition to a problem in matrix theory, and finally we solve the matrix problem as a separate lemma, perhaps of independent interest, using elementary combinatorics. Let $\mathbb{L}$ be a normal closure of $\mathbb{K}$. Since $\mathbb{K}$ is separable over $\mathbb{F}$, $\mathbb{L}$ is Galois over $\mathbb{F}$, say with Galois group G. Let H be the subgroup of G that fixes $\mathbb{K}$, i.e.,

$$H = \{\sigma \in G : \sigma(\beta) = \beta, \forall \beta \in \mathbb{K}\}.$$

We know from Galois theory that $[G : H] = m$. Let $\sigma_0 = 1, \sigma_1, \ldots, \sigma_{m-1}$ be any distinct representatives of (left) cosets of H in G. We claim that, for any $\beta \in \mathbb{K}$,

$$\mathbb{K} = \mathbb{F}(\beta) \quad \text{iff} \quad \sigma_i(\beta) \neq \beta \text{ for } \quad 1 \leq i \leq m-1.$$

In fact, if $\mathbb{K} = \mathbb{F}(\beta)$ and $\sigma_i(\beta) = \beta$ for some $i \geq 1$ then σ_i also fixes $\mathbb{K}$, hence σ_i belongs to H, contradicting our choice of σ_i. On the other hand, suppose $\mathbb{K} \neq \mathbb{F}(\beta)$. Then $[\mathbb{F}(\beta) : \mathbb{F}] < m$. Let H_1 be the subgroup of G that fixes $\mathbb{F}(\beta)$. Then $H \subseteq H_1$ and, by Galois theory again, $[G : H_1] = [\mathbb{F}(\beta) : \mathbb{F}] < m = [G : H]$. This implies that $H_1 \neq H$. Hence H_1 contains a coset of H other than H, say $\sigma_i H$ for some $i \geq 1$. Since σ_i is in H_1, it fixes β. This proves our preliminary claim.

Now let $\beta = b_1\alpha_1 + \cdots + b_n\alpha_n \in \mathbb{K}$ where $b_1,\ldots,b_n \in \mathbb{F}$. Note that the functions $\sigma_j - \sigma_0$, $1 \leq j \leq m-1$, are linear transformations of $\mathbb{L}$ over $\mathbb{F}$, so we have

$$(\sigma_j - \sigma_0)\beta = \sum_{i=1}^{n} b_i(\sigma_j - \sigma_0)\alpha_i.$$

Let A be the $n \times (m-1)$ matrix (a_{ij}) with $a_{ij} = (\sigma_j - \sigma_0)\alpha_i \in \mathbb{L}$ for $1 \leq i \leq n$ and $1 \leq j \leq m-1$. Then

$$((\sigma_1 - \sigma_0)\beta, \ldots, (\sigma_{m-1} - \sigma_0)\beta) = (b_1, \cdots, b_n)A.$$

Note that no column of A is zero, since if otherwise, say j-th column is zero for some $j \geq 1$, then $\sigma_j(\alpha_i) = \alpha_i$ for all $1 \leq i \leq n$, hence $\mathbb{K} = F(\alpha_1, \ldots, \alpha_n)$ would be fixed by σ_j, contradicting the choice that $\sigma_j \notin H$. By the characterization above, $\mathbb{K} = \mathbb{F}(\beta)$ iff $(\sigma_j - \sigma_0)\beta \neq 0$ for all $1 \leq j \leq m-1$. This is equivalent to requiring that each entry of the vector $(b_1, \ldots, b_n)A$ is nonzero. The theorem follows from the following lemma.

Lemma. *Let $\mathbb{L}$ be any field (or any integral domain) and $A = (a_{ij})$ an $n \times m$ matrix over $\mathbb{L}$ with no zero columns. Let S be any subset of $\mathbb{L}$ with k elements. The number of n-tuples $(b_1, \ldots, b_n) \in S^n$ such that the row vector $(b_1, \cdots, b_n)A$ has no zero entry is at least*

$$(1 - \frac{m}{k}) \cdot k^n.$$

Proof. We prove by induction on m and use the inclusion-exclusion principle. For $m = 1$, the number of n-tuples $(b_1, b_2, \ldots, b_n) \in S^n$ such that

$$bA = b_1 a_{11} + b_2 a_{12} + \cdots + b_n a_{1n} = 0$$

is at most k^{n-1}. To see this, note that at least one of the entries a_{1j} is nonzero which, without loss of generality, we assume is a_{11}. For each choice of $(b_2, \ldots, b_n) \in S^{n-1}$, b_1 is uniquely determined by the above equation, and this value of b_1 may or may not lie in S. Hence the total number of solutions $b \in S^n$ is at most k^{n-1}, or equivalently the number of nonsolutions (i.e. the n-tuples $b \in S^n$ such that bA is nonzero) is at least $k^n - k^{n-1} = (n - 1/k)k^{n-1}$. This establishes the inequality for $m = 1$.

Now let $m > 1$ and assume the lemma is true for any $n \times (m-1)$ matrix with no zero columns. Let A be an $n \times m$ matrix with no zero column. We partition A as $A = [A_1, A_2]$ where A_1 has one column and A_2 has $m-1$ columns. Note that $bA = [bA_1, bA_2]$. For $i = 1, 2$, let B_i be the set of $b \in S^n$ so that bA_i has at least one zero component. Then $|B_1| \leq k^{n-1}$ by the base case above, and $|B_2| \leq (m-1)k^{n-1}$ by the induction hypothesis, where $|B|$ denotes the cardinality of a set B. Now by the inclusion-exclusion principle, the number of elements $b \in S^n$ such that bA has no zero components is

$$k^n - |B_1| - |B_2| + |B_1 \cap B_2| \geq k^n - |B_1| - |B_2| \geq k^n - k^{n-1} - (m-1)k^{n-1} = (1 - \frac{m}{k}) \cdot k^n.$$

Hence the proof is complete by induction. □

Remarks. **1.** The above lemma may be of independent interest. It can be restated in terms of probability, namely, if one picks b_i from S randomly then the probability that none of the v_i is zero is at least $1 - m/k$. The proof can also be phrased in the probability language. We mention that this lemma is related to the Separation Probability Lemma in [4], which deals with the probability that all v_i are pairwise distinct (assuming A has no repeated columns), which is a generalization of the well known birthday paradox.

2. Our bound on $\rho(S; \alpha_1, \ldots, \alpha_n)$ works for any collection of generators, even if they contain a lot of linear dependency. For example, when $\mathbb{K} = \mathbb{Q}[\sqrt{2}, \sqrt{3}]$, one may take $\alpha_1 = \sqrt{2}$ and $\alpha_i = i\sqrt{3}$ for $2 \leq i \leq n$ (for any $n > 1$). The bound depends only on the size of the set S and on the degree $m = 4$ of the extension, not on the number of generators. On the other hand, the bound may be weak for special cases. For example $\mathbb{K} = \mathbb{Q}[\sqrt{2}, \sqrt{3}]$ with $\alpha_1 = \sqrt{2}$ and $\alpha_2 = \sqrt{3}$, in this case $a\sqrt{2} + b\sqrt{3}$ is primitive for $\mathbb{K}$ over $\mathbb{Q}$ for all $a, b \in \mathbb{Q} \setminus \{0\}$; hence $\rho(S; \alpha_1, \alpha_2) = 1$ if $0 \notin S$, while our bound says that $\rho(S; \alpha_1, \alpha_2) \geq 1 - 4/(|S| - 1)$.

3. If $\mathbb{F}$ is a field of size greater than m, then there is a subset S of $\mathbb{F}$ of size greater than m and the corresponding bound is positive, thus proving the existence of primitive elements. When $|S| \leq m$, our bound says nothing about the density; in particular, when $\mathbb{F}$ is a finite field with less than m elements, our bound says nothing about the existence of primitive elements. In fact, for certain generators $\alpha_1, \ldots, \alpha_n$, the density $\rho(\mathbb{F}; \alpha_1, \ldots, \alpha_n)$ may be zero, so there is no primitive element among the linear combinations of the generators. For example, let $\mathbb{F} = \mathbb{F}_2$, and let $\alpha_1 = \gamma_1 + \gamma_2$ and $\alpha_2 = \gamma_2 + \gamma_3$ where $\gamma_1, \gamma_2, \gamma_3 \in \overline{\mathbb{F}}_2$ have degrees $3, 5, 7$, respectively, over $\mathbb{F}_2$. Note that $\mathbb{F}_{2^{3\cdot5\cdot7}} = \mathbb{F}_2(\alpha_1, \alpha_2)$, however, none of $b_1\alpha_1 + b_2\alpha_2$, where $b_1, b_2 \in \mathbb{F}_2$, generates $\mathbb{F}_{2^{3\cdot5\cdot7}}$!

Testing Primitivity. We now make some comments on how to test whether a given element is primitive. From the proof of the main theorem, it might seem that in order to do this one would need to know the Galois group of the extension $\mathbb{K} = \mathbb{F}(\alpha_1, \ldots, \alpha_n)$ over $\mathbb{F}$, which is in general hard to compute. However, testing primitivity can be done in a much simpler way: To check whether $\beta \in \mathbb{K}$ is primitive, one just needs to find the degree of the minimal polynomial of β over $\mathbb{F}$ via linear algebra, that is, the smallest integer $d > 0$ such that β^d is a linear combination of $1, \beta, \ldots, \beta^{d-1}$ over $\mathbb{F}$. If this degree d of the minimal polynomial is equal to the degree of $\mathbb{K}$ over $\mathbb{F}$ then β is primitive; otherwise it is not. Here the important assumption is that $\mathbb{K} = \mathbb{F}(\alpha_1, \ldots, \alpha_n)$ is given as a *field* and the α's are represented appropriately. Grobner basis techniques provide convenient tools in dealing with representation and in finding the minimal polynomials. We demonstrate below by a concrete example related to geometry. For background on Grobner bases and their applications, we refer the reader to the excellent books [1, 2]. The method indicated below works in general and all steps can be computed efficiently.

Consider $\mathbb{F} = \mathbb{Q}(t)$, the field of rational functions in t where t is transcendental over $\mathbb{Q}$. Let $f(x) = x^4 - (4t - 2)x^2 + 1$ and $g(x) = x^4 - (4t + 2)x^2 + 1$, both irreducible in $\mathbb{F}[x]$. We want to construct a field extension $\mathbb{K}$ of smallest degree that contains one root for each of f and g, called a *composite field* of f and g. Let α_1 be a root of f. Then

$$\mathbb{F}[\alpha_1] \cong \mathbb{F}[x]/(f(x)),$$

and it is a field extension of degree 4 over $\mathbb{F}$. Factoring $g(y)$ in $\mathbb{F}[\alpha_1][y]$ gives

$$g(y) = \left(y^2 + (\alpha_1^3 + (1-4t)\alpha_1)y - 1\right) \cdot \left(y^2 - (\alpha_1^3 + (1-4\,t)\alpha_1)y - 1\right).$$

We remark that the problem of factoring g in $\mathbb{F}[\alpha_1][y]$ can be reduced to that of factoring a polynomial in $\mathbb{Q}[t,z]$ using a Grobner basis for the ideal $\langle f(x), g(y), z - x - y\rangle \subset \mathbb{F}[x,y,z]$ under the lex order $x > y > z$, and we refer the reader to [4] for an efficient algorithm for factoring multivariate polynomials.

We consider the two irreducible factors separatively, and we shall see that the field extensions behave differently. For the first factor of $g(y)$, let

$$g_1(x,y) = y^2 + (x^3 + (1-4t)x)y - 1,$$

and let α_2 be a root of $g_1(\alpha_1, y)$, so a root $g(y)$. Then $\mathbb{K}_1 = \mathbb{F}(\alpha_1, \alpha_2)$ is a field extension of degree 8 over $\mathbb{F}$ and it is isomorphic to the quotient ring $\mathbb{F}[x,y]/\langle f(x), g_1(x,y)\rangle$, which provides an explicit representation for $\mathbb{K}_1$. To see whether $\beta = \alpha_1 + \alpha_2$ is primitive for $\mathbb{K}_1$ over $\mathbb{F}$, define an ideal

$$I_1 = \langle f(x), g_1(x,y), z - x - y\rangle \subset \mathbb{F}[x,y,z].$$

Certainly, $\mathbb{K}_1$ is isomorphic to $\mathbb{F}[x,y,z]/I_1$. Note that the polynomials $f(x), g_1(x,y), z - x - y$ form a Grobner basis for I_1 under the lex order with $z > y > x$. We convert this basis into a reduced Grobner basis under a new lex order with $y > x > z$:

$$\begin{aligned}
&\left(576\,t^2 - 64\right) y - 3\,z^7 + 78\,tz^5 - \left(552\,t^2 + 28\right) z^3 + \left(1152\,t^3 - 288\,t^2 - 344\,t + 32\right) z,\\
&\left(576\,t^2 - 64\right) x + 3\,z^7 - 78\,tz^5 + \left(552\,t^2 + 28\right) z^3 - \left(1152\,t^3 + 288\,t^2 - 344\,t - 32\right) z,\\
&z^8 - 24\,tz^6 + (144\,t^2 + 8)z^4 - (256\,t^3 - 160\,t)z^2 + 16.
\end{aligned}$$

The last polynomial, denoted by h_1, belongs to $\mathbb{F}[z]$. Since h_1 has degree 8 equal to the degree of $\mathbb{K}_1$ over $\mathbb{F}$, we conclude that $\beta = \alpha_1 + \alpha_2$ is a primitive element of $\mathbb{K}_1$ over $\mathbb{F}$. Note that the first two polynomials correspond to explicit expressions of α_1 and α_2 as elements of $\mathbb{F}[\beta]$.

For the second factor of $g(y)$, let

$$g_2(x,y) = y^2 - (x^3 + (1-4\,t)x)y - 1,$$

and let α_3 be a root of $g_2(\alpha_1, y)$, so a root of $g(y)$. Then $\mathbb{K}_2 = \mathbb{F}(\alpha_1, \alpha_3)$ has degree 8 over $\mathbb{F}$ and is isomorphic to the quotient ring $\mathbb{F}[x,y]/\langle f(x), g_2(x,y)\rangle$. Similar to the above case, computing a reduced Grobner basis for the ideal $\langle f(x), g_1(x,y), z - x - y\rangle$ under the lex order with $y > x > z$ yields

$$\begin{aligned}
&y + x - z,\\
&2\,x^2 + \left(-z^3 + (-2 + 4\,t)\,z\right) x - 2,\\
&z^4 - 4\,tz^2 + 4.
\end{aligned}$$

The last polynomial belongs to $\mathbb{F}[z]$, but has degree 4 less than 8, the degree of $\mathbb{K}_2$ over $\mathbb{F}$. Hence $\alpha_1 + \alpha_3$ is not primitive for $\mathbb{K}_1$ over $\mathbb{F}$. To find a primitive element,

consider $\beta = \alpha_1 - \alpha_3$ and the ideal $I_2 = \langle f(x), g_1(x,y), z - x + y\rangle$. Computing a reduced Grobner basis for I_2 gives

$$\left(576\,t^2 - 64\right) y + 3\,z^7 - 78\,tz^5 + \left(552\,t^2 + 28\right) z^3 - \left(1152\,t^3 - 288\,t^2 - 344\,t + 32\right) z,$$
$$\left(576\,t^2 - 64\right) x + 3\,z^7 - 78\,tz^5 + \left(552\,t^2 + 28\right) z^3 - \left(1152\,t^3 + 288\,t^2 - 344\,t - 32\right) z,$$
$$z^8 - 24\,z^6 t + \left(144\,t^2 + 8\right) z^4 - \left(256\,t^3 - 160\,t\right) z^2 + 16.$$

Now the last polynomial, denoted by h_2, belongs to $\mathbb{F}[z]$ and has degree 8. Hence $\beta = \alpha_1 - \alpha_3$ is primitive for $\mathbb{K}_2$ over $\mathbb{F}$. Note that $h_2 = h_1$, hence $\mathbb{K}_2 = \mathbb{K}_1$ (but represented differently).

A reader with algebraic geometry background may see that the above polynomials f_1, f_2 and h_1 (similarly for h_2) have interesting geometric meaning. More precisely, they each define an irreducible curve over $\mathbb{Q}$, and these curves can be viewed as finite covers of the affine line. The curve h_1 defines a finite cover of smallest degree of the affine line so that it can be factored (as map) through each of the covers defined by f_1 and f_2, separately.

References

[1] D. Cox, J. Little and D. O'Shea, *Ideals, varieties, and algorithms*, 2nd ed., Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1997.

[2] David Cox, John Little and Donal O'Shea, *Using algebraic geometry*, Graduate Texts in Mathematics, 185, Springer-Verlag, New York, 1998.

[3] D. S. Dummit and R. M. Foote, *Abstract Algebra*, 2nd edition, 1999.

[4] S. Gao, Factoring multivariate polynomials via partial differential equations, *Mathematics of Computation* **72** (2003), 801–822.

[5] W. V. Vasconcelos, *Computational methods in Commutative Algebra and Algebraic Geometry*, Springer-Verlag, Berlin/Heidelberg/New York, 1998.

[6] B. L. van der Waerden, *Modern Algebra*, Vol. I, Ungar, New York, 1953.

[7] O. Zariski and P. Samuel, *Commutative Algebra*, Vol. I, Van Nostrand, Princeton, 1960.

Department of Mathematical Sciences, Clemson University, Clemson, SC 29634-0975 USA *E-mail address*: brawley@clemson.edu, sgao@math.clemson.edu