Splitting Fields for Characteristic Polynomials of Matrices with Entries in a Finite Field.

Eric Schmutz

Mathematics Department, Drexel University, Philadelphia, Pennsylvania, 19104.

Abstract

Let $M_n(q)$ be the set of all $n \times n$ matrices with entries in the finite field $F_q$. With asymptotic probability one, the characteristic polynomial of a random $A \in M_n(q)$ does not have all its roots in $F_q$. Let $X_n(A)$ be the degree of the splitting field of the characteristic polynomial of $A$, and let $\mu_n$ be the average degree:

$$\mu_n = \frac{1}{|M_n(q)|} \sum_A X_n(A).$$

Using a theorem of Reiner, we prove that

$$\mu_n = e^{c_0\sqrt{n/\log n}+o(1)},$$

where $c_0$ is an explicit constant.

Key words: Finite field, splitting field, random matrix, characteristic polynomial

1 Introduction and notation

If $f \in F_q[x]$, let $X(f)$ be the degree of the splitting field of $f$, i.e. the smallest $d$ such that $f$ factors as a product of linear factors $f = \prod_i (x - r_i)$, with all the roots $r_i$ in $F_{q^d}$. Mignotte and Nicolas [11],[14], and Dixon and Panario [2] asked how large $X(f)$ is for a typical polynomial $f$. More precisely, let $P_{n,q}$ be the set of all monic degree $n$ polynomials with coefficients in the finite field $F_q$, and let $P_{n,q}$ be the uniform probability measure: $P_{n,q}(\{f\}) = q^{-n}$ for all $f$. They studied the asymptotic distribution of the random variable $\log X$, and noted the strong analogies between this problem and the “Statistical

Email address: Eric.Jonathan.Schmutz@drexel.edu (Eric Schmutz).

Preprint submitted to Elsevier 6 July 2006
Group Theory” of Erdős and Turán[3]. Dixon and Panario[2] also estimated the expected value \( E(X) = \text{the average degree} \). Hansen and Schmutz[7] compared random polynomials with the characteristic polynomials of random invertible matrices. Based on this work, it is reasonable to conjecture that a matrix-analogue of the Dixon-Panario theorem should hold.

The number of matrices having a given characteristic polynomial depends, in a complicated way, on the degrees of the irreducible factors that the polynomial has (Reiner[15]). Select a matrix \( A \) uniform randomly from among all \( q^{n^2} \) matrices having entries in the finite field \( \mathbb{F}_q \), and let \( f \) be the characteristic polynomial of \( A \). Hence the characteristic polynomial \( f \) is being selected randomly, but not uniform randomly, from among all monic degree \( n \) polynomials in \( \mathbb{F}_q[x] \). Let \( \mu_n = \text{the average, over all } q^{n^2} \text{ matrices } A, \text{ of the degree of the splitting field of the characteristic polynomial of } A. \) We prove here that

\[
\mu_n = e^{c_0 \sqrt{n / \log n (1+o(1))}}.
\]

where \( c_0 \) is an explicit constant.

The remainder of this section specifies the paper’s symbols and notation. Definitions are listed here in quasi-alphabetical order, and then used later without comment.

- \( c_0 = 2 \sqrt{\int \frac{\log(1+t)}{e^t - 1} \, dt} = 2.990 \ldots \)
- \( c_\infty = \prod_{j=1}^{\infty} (1 - \frac{1}{2^j}) = 0.288 \ldots \)
- \( | \cdot | = \text{degree: if } f \text{ is a polynomial in } \mathbb{F}_q[x], \text{ then } |f| \text{ is its degree.} \)
- \( F(u, r) := \prod_{i=1}^{r} (1 - \frac{1}{u^i}) \text{ for positive integers } u, r, \text{ and } F(u, 0) := 1 \)
- \( g_f = \text{divisor (in } \mathbb{F}_q[x]) \text{ of } f \text{ that is minimal among divisors } g \text{ of } f \text{ for which } X(f) = X(g). \)
- \( G_{n,q} = \{ g_f : f \in \mathcal{P}_{n,q} \}. \)
- \( \mathcal{I}_{n,q} = \text{monic polynomials of degree } n \text{ in } \mathbb{F}_q[x] \)
- \( \mathcal{I} = \bigcup_{n=1}^{\infty} \bigcup_{n,q} \text{monic polynomials in } \mathbb{F}_q[x] \text{ that are irreducible over } \mathbb{F}_q \)
- \( \Lambda_m = \text{partitions of } m \text{ having distinct parts.} \)
- \( \tilde{\Lambda}_m = \text{partitions of } m \text{ (non necessarily distinct parts.)} \)
- \( \mathcal{M}_{n,q} = \text{set of all } n \times n \text{ matrices with entries in the finite field } \mathbb{F}_q. \)
- \( M_{n,q} = \text{probability measure on } \mathcal{P}_{n,q} \text{ defined by } M_{n,q}(\{f\}) = \text{the proportion of matrices in } \mathcal{M}_{n,q} \text{ whose characteristic polynomial is } f. \)
- \( m_\phi(f) = \text{the multiplicity of } \phi \text{ in } f: \text{ for } \phi \in \mathcal{I} \text{ and } f \in \mathbb{F}[x], \phi^{m_\phi(f)} \text{ divides } f \text{ but } \phi^{m_\phi(f)+1} \text{ does not divide } f. \)
- \( \mu_n = \sum_{A \in \mathcal{M}_{n,q}} X(A). \)
- \( \mathcal{P}_{n,q} = \text{set of all } q^n \text{ monic polynomials of degree } n \text{ in } \mathbb{F}_q[x]. \)
- \( P_{n,q} = \text{uniform probability measure on } \mathcal{P}_{n,q}: P_{n,q}(\{f\}) = q^{-n} \)
• $S_{n,q}$ polynomials in $\mathcal{P}_{n,q}$ that factor completely, i.e. have all their roots in $\mathbb{F}_q$.
• $X(f) = \text{degree of the splitting field of } f$, if $f \in \mathbb{F}_q[x]$.
• $X(A) = X(f)$, if $A$ is a matrix with characteristic polynomial $f$.
• $X(\lambda) = \text{least common multiple of the parts of } \lambda$, if $\lambda$ is an integer partition.

The last three definitions overload the symbol $X$. However this is natural and consistent: the degrees of the irreducible factors of a polynomial $f \in \mathbb{F}_q[x]$ form a partition of $|f|$, and it is well known that the degree of the splitting field of $f$ is the least common multiple of the degrees of its irreducible factors.

2 Comparison of the probability measures

There is an explicit formula for the number of matrices with a given characteristic polynomial:

**Theorem 1** (Reiner[15]) If $f = \prod \phi^{m_\phi(f)}$ is a polynomial in $\mathcal{P}_{n,q}$, then

$$M_{n,q}(\{f\}) = \frac{q^nF(q,n)}{\prod_{\phi \in I} F(q^{\phi|}, m_\phi(f))}$$

(See also Crabb[1], Fine-Herstein[5], and Gerstenhaber[9]). In order to apply Theorem 1, we need a simple lemma:

**Lemma 1** For all non-negative integers $a, b$, and all prime powers $q$,

$$F(q, a + b) \geq F(q, a)F(q, b)$$

**Proof.** Since $q^{a+j} \geq q^j$ for all $j$, we have

$$F(q, b) = \prod_{j=1}^{b} (1 - \frac{1}{q^j}) \leq \prod_{j=1}^{b} (1 - \frac{1}{q^{a+j}}).$$

But then

$$F(q, a + b) = F(q, a) \prod_{j=1}^{b} (1 - \frac{1}{q^{a+j}}) \geq F(q, a)F(q, b).$$

In one direction, there is a simple relationship between the probability measures $P_n$ and $M_{n,q}$:
Proposition 1 For all $A \subseteq \mathcal{P}_n$, 

$$M_{n,q}(A) \geq c_\infty P_{n,q}(A).$$

Proof. It is obvious from the definition of $F$ that, for all $u > 1$ and all non-negative integers $r$,

$$0 < F(u,r) \leq 1. \quad (1)$$

If $f \in A$, then by Theorem 1 and (1),

$$M_n(\{f\}) \geq F(q, n) q^{-n} \geq c_\infty q^{-n}.$$ 

Summing over $f \in A$ we get Proposition 1. 

It is interesting to note that the inequality in Proposition 1 has no analogue in the other direction:

Proposition 2 $\limsup_{n \to \infty} \max_{f \in \mathcal{P}_{n,q}} \frac{M_{n,q}(f)}{P_{n,q}(f)} = \infty$

Proof. Consider $f =$ the product of all irreducible polynomials of degree less than or equal to $m$. In this case $n = n_m = \sum_{d=1}^m d |I_{d,q}|$, and

$$M_{n,q}(\{f\}) = q^{-n} \frac{F(q, n)}{\prod_{k=1}^m (1 - \frac{1}{q^k})^{|I_k|}} \quad (2)$$

Since $F(q, n) \geq c_\infty$, it suffices to prove that $\prod_{k=1}^m (1 - \frac{1}{q^k})^{I_k} = o(1)$ as $m \to \infty$. The following bounds appear on page 238 of Mignotte[10]:

$$q^k \geq I_k = \frac{1}{k} \sum_{d|k} \mu(d)q^{k/d} \geq q^k \left( \frac{1}{k} - \frac{2}{k} q^{k/2} \right) \quad (3)$$

Using first the inequality $\log(1 - x) \leq -x$, and then the inequality on the right side of (3), we get

$$\prod_{k=1}^m (1 - \frac{1}{q^k})^{I_k} \leq \exp \left( -\sum_{k=1}^m \frac{1}{k} + O(1) \right) = O\left( \frac{1}{m} \right)$$

•
3 Non-existence of Jordan forms.

Neumann and Praeger[12] estimated the probability that the characteristic polynomial of a random matrix has none of its roots in \( \mathbb{F}_q \). In this section we estimate the probability that the characteristic polynomial of a random matrix has all of its roots in \( \mathbb{F}_q \).

**Theorem 2** For all prime powers \( q \) and all positive integers \( n \),

\[
  c_{\infty} q^{-n} \left( \frac{n + q - 1}{q - 1} \right) \leq M_{n,q} (\mathcal{S}_{n,q}) \leq q^{-n} \left( \frac{n + q - 1}{q - 1} \right)
\]

**Proof.** Suppose \( f \in \mathbb{F}_q[x] \). Then \( f \in \mathcal{S}_{n,q} \) iff two conditions are satisfied:

1. The multiplicities of the linear factors form composition of \( n \) into non-negative integer parts: \( \sum_{\alpha \in \mathbb{F}_q} m_{x-\alpha}(f) = n \), and
2. \( m_\phi(f) = 0 \) for all \( \phi \in \bigcup_{d \geq 2} I_d \); no irreducible factor has degree larger than one.

It is well known that there are exactly \( \binom{n+q-1}{q-1} \) compositions of \( n \) into \( q \) non-negative parts. It therefore suffices to prove that, for any \( f \in \mathcal{S}_{n,q} \), \( c_{\infty} q^{-n} \leq M_{n,q} (\{f\}) \leq q^{-n} \).

Suppose \( f = \prod_{\alpha \in \mathbb{F}_q} (x - \alpha)^{m_{x-\alpha}(f)} \) and \( \sum_{\alpha \in \mathbb{F}_q} m_{x-\alpha}(f) = n \). By Theorem 1, \( M_{n,q}(\{f\}) = \prod_{\alpha \in \mathbb{F}_q} F(q,m_{x-\alpha}(f)). \) Lemma 1 implies that \( \prod_{\alpha \in \mathbb{F}_q} F(q,m_{x-\alpha}(f)) \geq F(q,n) \). Therefore, \( M_{n,q}(\{f\}) \leq q^{-n} \).

For the other direction, apply Proposition 1 with \( A = \{f\} \).

*Corollary 1* For almost every matrix \( A \in M_n(q) \), there is no matrix \( B \in M_n(q) \) such that \( B \) is in Jordan canonical form and is similar to \( A \).

**Comment:** In Corollary 1, “almost every matrix” means “for all but \( o_q(q^n^2) \) matrices”, where the subscript in the little-o indicates \( q \)-dependence. The bounds in Theorem 2 do hold for all \( n \) and \( q \). But for fixed \( n \), we have \( \binom{n+q-1}{q-1} q^{-n} \rightarrow \frac{1}{n} \) as \( q \rightarrow \infty \). On other hand, for fixed \( q \), \( \binom{n+q-1}{q-1} q^{-n} \) approaches zero exponentially fast as \( n \rightarrow \infty \).
4 Average degree

An easy consequence of Proposition 1 is a lower bound for the average degree:

Lemma 2 $\mu_n \geq c_0 \sqrt{n/\log n} (1 + O(\frac{\log \log n}{\sqrt{\log n}}))$

Proof. By Theorem 1,

$$\mu_n = \sum_{f \in \mathcal{P}_{n,q}} q^{-n} \frac{F(q, n)}{\prod_{\phi} F(q^{|\phi|}, m_\phi(f))} X(f)$$

Again using the inequality (1), we get $\prod_{\phi} F(q^{|\phi|}, m_\phi(f)) \leq 1$ and $F(q, n) \geq c_\infty$.
Therefore $\mu_n \geq c_\infty \sum_{f \in \mathcal{P}_{n,q}} q^{-n} X(f)$. The lower bound then follows directly from the results of Dixon and Panario [2].

The upper bound for $\mu_n$ is harder because, as Proposition 2 suggests, we don’t have convenient upper bounds the $M_{n,q}$-probabilities of events. Two lemmas are needed for the proof.

Let $\mathcal{D}(f) = \{ g : g \text{ divides } f \text{ in } \mathbb{F}_q[x] \text{ and } X(g) = X(f) \}$. Then $\mathcal{D}(f)$ is a non-empty finite set that is partially ordered by divisibility. For each $f$, we can choose a minimal element $g_f \in \mathcal{D}(f)$.

Lemma 3 The irreducible factors of $g_f$ appear with multiplicity one and have different degrees.

Proof. Suppose that, on the contrary, $\phi_1$ and $\phi_2$ are irreducible polynomials of degree $d$ and that $\phi_1 \phi_2$ divides $g_f$. Let $g = \frac{g_f}{\phi_1}$. Then $X(g) = X(f)$ and $g$ divides $g_f$. This contradicts the minimality of $g_f$.

Lemma 4 If $|g_f| = d$, then $M_{n,q}(\{f\}) \leq 4M_d(\{g_f\})M_{n-d}(\{h_f\})$.

Proof. Since $f = g_fh_f$, we have $m_\phi(f) = m_\phi(g_f) + m_\phi(h_f)$. It therefore follows from Lemma 1 that

$$F(q^{|\phi|}, m_\phi(f)) \geq F(q^{|\phi|}, m_\phi(g_f))F(q^{|\phi|}, m_\phi(h_f)).$$ (4)
Combining (4) with Theorem 1, we get

\[ M_{n,q}(\{f\}) = \frac{F(q,n)}{q^m \prod_{\phi} F(q^{\phi}, m_{\phi}(f))} \leq \frac{F(q,n)}{q^m \prod_{\phi} F(q^{\phi}, m_{\phi}(g_f))} \]

\[ = \frac{F(q,n)}{F(q,d) F(q,n-d)} \]

\[ M_d(\{g_f\}) M_{n-d}(\{h_f\}). \]

Finally, \( \frac{F(q,n)}{F(q,d) F(q,n-d)} \leq \frac{1}{F(q,d)} \leq \frac{1}{c_{\infty}} \leq 4. \)

\[ \star \]

**Theorem 3** \( \mu_n = \exp \left( c_0 \sqrt{\frac{n}{\log n}} \left( 1 + O\left( \frac{\log \log n}{\log n} \right) \right) \right). \)

**Proof.**

\[ \mu_n = E(X) = \sum_{f \in \mathcal{P}_{n,q}} M_{n,q}(f) X(f) \]

\[ = \sum_{g \in \mathcal{G}_{n,q}} X(g) \sum_{h} M_{n,q}(\{gh\}), \]

where the inner sum is over all \( h \) for which \( g_{gh} = g \). By Lemma 4, this is less than

\[ \sum_{g \in \mathcal{G}_{n,q}} X(g) 4M_{|g|}(\{g\}) \sum_{h} M_{n-|g|,q}(\{h\}) \]

Since the inner sum is less that one, we have

\[ \mu_n \leq 4 \sum_{g \in \mathcal{G}_{n,q}} X(g) M_{n,q}(\{g\}). \]

(5)

If \( g \in \mathcal{G}_{n,q} \), then the degrees of the irreducible factors of \( g \) form a partition of \( |g| \) into distinct parts. Grouping together polynomials that have the same partition, we see that the right side of (5) is less than or equal to

\[ 4 \sum_{m=1}^{n} \sum_{\lambda \in \Lambda_m} LCM(\lambda_1, \lambda_2, \ldots) q^{-m} \prod_{i} \frac{|I_{\lambda_i}|}{1 - \frac{1}{q^{\lambda_i}}}. \]

(6)

If \( \lambda \) has distinct parts \( \lambda_1, \lambda_2, \ldots \), then \( \prod_{i} (1 - \frac{1}{q^{\lambda_i}}) \geq \prod_{i=1}^{m} (1 - \frac{1}{q}) \geq c_{\infty} \). It is well known that \( |I_{\lambda_i}| \leq \frac{2^\lambda_i}{\lambda_i} \). Putting these two estimates back into the right side of (6), we get

\[ \mu_n \leq 4 \frac{\sum_{m=1}^{n} \sum_{\lambda \in \Lambda_m} LCM(\lambda_1, \lambda_2, \ldots)}{\lambda_1 \lambda_2 \ldots} \leq 4 \frac{\sum_{m=1}^{n} \sum_{\lambda \in \Lambda_m} LCM(\lambda_1, \lambda_2, \ldots)}{\lambda_1 \lambda_2 \ldots} \]

(7)
This last quantity has appeared previously in the study of random permutations ([4],[8],[16]) where it was approximated by coefficient of $x^n$ in the generating function
\[ \frac{1}{1-x} \prod_{\text{primes } p}^{\infty} (1 + x^p + \frac{x^{2p}}{2} + \frac{x^{3p}}{3} + \cdots). \]

The conclusion was that the right side of (7) is
\[ \exp \left( c_0 \sqrt{\frac{n}{\log n}} \left( 1 + O(\frac{\log \log n}{\sqrt{\log n}}) \right) \right). \]

References


