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

Eric Schmutz

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

Abstract

Let $\mathcal{M}_n$ be the set of all $n \times n$ matrices with entries in the finite field $\mathbb{F}_q$. Let $\mathbf{X}(A)$ be the degree of the splitting field of the characteristic polynomial of A, and let μ_n be the *average* degree:

$$\mu_n = \frac{1}{|\mathcal{M}_n|} \sum_{A \in \mathcal{M}_n} \mathbf{X}(A).$$

A theorem of Reiner is used to prove that,as $n \to \infty$,

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

where B is an explicit constant.

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

1 Introduction

If $f \in \mathbb{F}_q[x]$,let $\mathbf{X}(f)$ 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 $\mathbb{F}_{q^d}$. Mignotte and Nicolas [12],[16], and Dixon and Panario [2] asked how large $\mathbf{X}(f)$ is for a typical polynomial f. More precisely, let $\mathcal{P}_n$ be the set of all monic degree n polynomials with coefficients in the finite field $\mathbb{F}_q$, and let P_n be the uniform probability measure: $P_n(\{f\}) = q^{-n}$ for all $f \in \mathcal{P}_n$. They studied the asymptotic distribution of the random varaible $\log \mathbf{X}$, and noted the strong analogies between this problem and the "Statistical Group Theory"of Erdős and Turán[3],[4]. Dixon and Panario[2] also estimated the the average degree $q^{-n} \sum_{f \in \mathcal{P}_n} \mathbf{X}(f)$, i.e. the expected value of $\mathbf{X}$. Hansen and

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

Schmutz compared random polynomials with the characteristic polynomials of random invertible matrices. Based on the results in [8], it was 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[17]). 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 μ_n =the average, over all q^{n^2} matrices A, of the degree of the splitting field of the characteristic polynomial of A. We prove here that, as $n \to \infty$,

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

where $B = 2\sqrt{2\int_0^\infty \frac{\log(1+t)}{e^{-t}-1}dt} = 2.990\ldots$ The constant B has appeared previously in the study of random permutations[5] and random polynomials[2].

The remainder of this section specifies the paper's symbols and notations. Definitions are listed here in quasi-alphabetical order, and may be used later without comment.

- $B = 2\sqrt{2\int_0^\infty \frac{\log(1+t)}{e^{t}-1}dt} = 2.990\ldots$
- $c_\infty = \prod_{j=1}^{\infty}(1-\frac{1}{2^j}) = .288\ldots$
- $|\cdot|$ = degree: if f is a polynomial in $\mathbb{F}_q[x]$, then $|f|$ is its degree.
- $F(u,r) := \prod_{i=1}^{r}(1-\frac{1}{u^i})$ for positive integers u, r, and $F(u,0) := 1$
- g_f = divisor (in $\mathbb{F}_q[x]$) of f that is minimal among those monic divisors g of f for which $\mathbf{X}(f) = \mathbf{X}(g)$.
- $\mathcal{G}_n = \{g_f : f \in \mathcal{P}_n\}$.
- $h_f = \frac{f}{g_f}$
- $\mathcal{I}_k$ =the set of monic irreducible polynomials of degree k in $\mathbb{F}_q[x]$
- $H_n = \sum_{k=1}^{n}\frac{1}{k}$, the n'th Harmonic number.
- $I_k = |\mathcal{I}_{k,q}|$, the cardinality of $\mathcal{I}_k$ (The font distinguishes the set from its cardinality. To save space, q is implicit.)
- $\mathcal{I} = \bigcup_{k=1}^{\infty}\mathcal{I}_k$=monic polynomials in $\mathbb{F}_q[x]$ that are irreducible over $\mathbb{F}_q$
- Λ_m= set of partitions of m having distinct parts.
- $\tilde{\Lambda}_m$= partitions of m (not necessarily distinct parts.)
- $\lambda \vdash m$ The conventional notation for $\lambda \in \tilde{\Lambda}_m$.
- $\mathcal{M}_n$=set of all all $n \times n$ matrices with entries in the finite field $\mathbb{F}_q$.
- M_n = probability measure on $\mathcal{P}_n$ defined by $M_n(\{f\})$ = the proportion of matrices in $\mathcal{M}_n$ whose characteristic polynomial is f. (To save space, q is

implicit.)

- $m_\phi(f)$ = the multiplicity of ϕ in f: for $\phi \in \mathcal{I}$ and $f \in \mathbb{F}[x]$, $\phi^{m_\phi(f)}$ divides f but $\phi^{m_\phi(f)+1}$ does not divide f.
- $\mu_n = q^{-n^2} \sum_{A \in \mathcal{M}_n} \mathbf{X}(A)$.
- $\mathcal{P}_n$ = set of all q^n monic polynomials of degree n in $\mathbb{F}_q[x]$.
- P_n = uniform probability measure on $\mathcal{P}_n$: $P_n(\{f\}) = q^{-n}$.
- $\mathcal{S}$ = set of polynomials in $\mathcal{P}_n$ that factor completely, i.e. have all their roots in $\mathbb{F}_q$.
- $\mathbf{X}(f)$ = degree of the splitting field of f, if $f \in \mathbb{F}_q[x]$.
- $\mathbf{X}(A) = \mathbf{X}(f)$, if A is a matrix with characteristic polynomial f.
- $\mathbf{X}(\lambda)$ = least common multiple of the parts of λ, if λ is an integer partition.

The last three definitions overload the symbol $\mathbf{X}$. 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[17])If $f = \prod_\phi \phi^{m_\phi(f)}$ is a polynomial in $\mathcal{P}_n$, then*

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

(See also Crabb[1], Fine-Herstein[6]), and Gerstenhaber[10]). 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 :

Proposition 1 *For all* $\mathcal{A} \subseteq \mathcal{P}_n$,

$$M_n(\mathcal{A}) \geq c_\infty P_n(\mathcal{A}).$$

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

$$0 < F(u,r) \leq 1. \tag{1}$$

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

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

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

•

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

Proposition 2 $\limsup_{n\to\infty} \max_{f\in\mathcal{P}_n} \frac{M_n(\{f\})}{P_n(\{f\})} = \infty.$

Proof. Consider f =the product of all irreducible polynomials of degree less than or equal to m. In this case $n = n_m = \sum_{k=1}^{m} kI_k$, where I_k is the number of monic irreducible polynomials of degree k , and

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

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[11]:

$$q^k \geq I_k = \frac{1}{k}\sum_{d|k}\mu(d)q^{k/d} \geq q^k(\frac{1}{k} - \frac{2}{kq^{k/2}}). \tag{3}$$

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

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

•

3 Non-existence of Jordan forms.

Neumann and Praeger[13] estimated the probability that the characteristic polynomial of a random matrix has *none* of its roots in $\mathcal{F}_q$. 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,*

$$c_\infty q^{-n}\binom{n+q-1}{q-1} \leq M_n(\mathcal{S}) \leq q^{-n}\binom{n+q-1}{q-1}$$

Proof. Suppose $f \in \mathbb{F}_q[x]$. Then $f \in \mathcal{S}$ 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 $\phi \in \bigcup_{d\geq 2}\mathcal{I}_d$; 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 $f \in \mathcal{S}$, $c_\infty q^{-n} \leq M_n(\{f\}) \leq q^{-n}$.

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, $M_n(\{f\}) = \frac{q^{-n}F(q,n)}{\prod_{\alpha\in\mathbb{F}_q}F(q,m_{x-\alpha}(f))}$. Lemma 1 implies that $\prod_{\alpha\in\mathbb{F}_q}F(q,m_{x-\alpha}(f)) \geq F(q,n)$.

Therefore, $M_n(\{f\}) \leq q^{-n}$.

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

•

Note that, for fixed q, $\binom{n+q-1}{q-1}q^{-n}$ approaches zero exponentially fast as $n \to \infty$. Hence we have

Corollary 1 *For almost every matrix $A \in \mathcal{M}_n$, there is no matrix $B \in \mathcal{M}_n$ such that B is in Jordan canonical form and is similar to A.*

Comment: The corollary is stated rather glibly. A clumsier, but more precise statement is that, for every $\epsilon > 0$ and every prime power q, there is an integer $N_{\epsilon,q}$ such that, for all $n > N_{\epsilon,q}$, $\frac{|\mathcal{S}|}{q^{n^2}} < \epsilon$. The subscripts in the number $N_{\epsilon,q}$ are included to emphasize the fact that $N_{\epsilon,q}$ depends on both ϵ and q. It is interesting to compare the fixed-q-large-n limit, namely zero, with the fixed-n-large-q limit: $\lim_{q\to\infty}\binom{n+q-1}{q-1}q^{-n} = \frac{1}{n!} > 0$.

4 Average degree

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

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

Proof. By Theorem 1,

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

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 $\mu_n \geq c_\infty \sum_{f\in\mathcal{P}_n} q^{-n}\mathbf{X}(f)$. The lower bound then follows directly from the results of Dixon and Panario [2].

•

The upper bound for μ_n is harder because, as Proposition 2 suggests, we don't have convenient upper bounds the M_n-probabilities of events. Two lemmas are needed for the proof.

Let $\mathcal{D}(f) = \{g : g \text{ divides } f \text{ in } \mathbb{F}_q[x] \text{ and } \mathbf{X}(g) = \mathbf{X}(f)\}$. 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, ϕ_1 and ϕ_2 are irreducible polynomials of degree d and that $\phi_1\phi_2$ divides g_f. Let $g = \frac{g_f}{\phi_1}$. Then $\mathbf{X}(g) = \mathbf{X}(f)$ and g divides g_f. This contradicts the minimality of g_f.

•

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

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

$$F(q^{|\phi|}, m_\phi(f)) \geq F(q^{|\phi|}, m_\phi(g_f))F(q^{|\phi|}, m_\phi(h_f)). \tag{4}$$

Combining (4) with Theorem 1, we get

$$M_n(\{f\}) = \frac{F(q,n)}{q^n \prod_\phi F(q^{|\phi|}, m_\phi(f))} \leq \frac{F(q,n)}{q^n \prod_\phi F(q^{|\phi|}, m_\phi(g_f))F(q^{|\phi|}, m_\phi(h_f))}$$

$$= \frac{F(q,n)}{F(q,d)F(q,n-d)} M_d(\{g_f\})M_{n-d}(\{h_f\}).$$

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

•

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

Proof.

$$\mu_n = E(\mathbf{X}) = \sum_{f \in \mathcal{P}_n} M_n(f)\mathbf{X}(f)$$

$$= \sum_{g \in \mathcal{G}_n} \mathbf{X}(g) \sum_h M_n(\{gh\}),$$

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

$$\sum_{g \in \mathcal{G}_n} \mathbf{X}(g) 4M_{|g|}(\{g\}) \sum_h M_{n-|g|}(\{h\})$$

Since the inner sum is less that one, we have

$$\mu_n \leq 4 \sum_{g \in \mathcal{G}_n} \mathbf{X}(g) M_n(\{g\}). \tag{5}$$

If $g \in \mathcal{G}_n$, 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

$$4 \sum_{m=1}^{n} \sum_{\lambda \in \Lambda_m} LCM(\lambda_1, \lambda_2, \ldots) q^{-m} \prod_i \frac{I_{\lambda_i}}{(1 - \frac{1}{q^{\lambda_i}})}. \tag{6}$$

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

$$\mu_n \leq \frac{4}{c_\infty} \sum_{m=1}^{n} \sum_{\lambda \in \Lambda_m} \frac{LCM(\lambda_1, \lambda_2, \ldots)}{\lambda_1 \lambda_2 \cdots} \leq \frac{4}{c_\infty} \sum_{m=1}^{n} \sum_{\lambda \in \tilde{\Lambda}_m} \frac{LCM(\lambda_1, \lambda_2, \ldots)}{\lambda_1 \lambda_2 \cdots} \tag{7}$$

This last quantity has appeared previously in the study of random permutations [9],[18] where it was approximated by coefficient of x^n in the generating function

$$\frac{1}{(1-x)^2} \prod_{\text{primes } p}^{\infty} (1 + x^p + \frac{x^{2p}}{2} + \frac{x^{3p}}{3} + \cdots).$$

The conclusion (see appendix) was that the right side of (7) is

$$\exp\left(B \sqrt{\frac{n}{\log n}} (1 + O(\frac{\log \log n}{\sqrt{\log n}})) \right). \tag{8}$$

•

References

[1] M.C.Crabb,Counting nilpotent endomorphisms, *Finite Fields and Their Applications*, **12** (2006) 151–154.

[2] John Dixon and Daniel Panario, The degree of the splitting field of a random polynomial over a finite field, *Electron. J. Combin.* **11** (2004), no. 1, Research Paper 70, 10 pp. (electronic).

[3] Paul Erdős and Paul Turán, On some problems of a statistical group theory III, *Acta Math. Acad. Sci. Hungar.* **18** (1967) 309–320.

[4] Paul Erdős and Paul Turán, On some problems of a statistical group theory IV, *Acta Math. Acad. Sci. Hungar.* **19** (1968) 413–435.

[5] Steven R. Finch, Mathematical constants. *Encyclopedia of Mathematics and its Applications,* **94**. Cambridge University Press, Cambridge, 2003 ISBN 0-521-81805-2, page 287.

[6] N.J. Fine and I.N. Herstein, The probability that a matrix be nilpotent, *Illinois J.Math***2**(1958) 499-504.

[7] Jason Fulman, Random matrix theory over finite fields, *Bull. Amer. Math. Soc.* (N.S.) **39** (2002), no. 1, 51–85

[8] Jennie C. Hansen and Eric Schmutz, How random is the characteristic polynomial of a random matrix? *Math. Proc. Cambridge Philos. Soc.* **114** (1993), no. 3, 507–515.

[9] William M.Y. Goh and Eric Schmutz, The expected order of a random permutation.*Bull. London Math. Soc.***23** (1991), no. 1, 34–42.

[10] Murray Gerstenhaber, On the number of nilpotent matrices with coefficients in a finite field. *Illinois J. Math.* **5** 1961 330–333.

[11] Maurice Mignotte, Mathematics for Computer Algebra, Springer Verlag 1992 ISBN 3-540-97675-2.

[12] M.Mignotte and J.L.Nicolas, Statistiques sur $F_q[X]$. *Ann. Inst. H. Poincar Sect. B (N.S.)***19** (1983), no. 2, 113–121.

[13] Peter M. Neumann and Cheryl E. Praeger, Derangements and eigenvalue-free elements in finite classical groups.

[14] Peter M. Neumann and Cheryl E. Praeger, Cyclic Matrices over Finite Fields *J. London Math. Soc.* (2) **58** (1998), no. 3, 564–586.

[15] Daniel Panario, What Do Random Polynomials over Finite Fields Look Like? `Finite fields and applications,` 89–108, *Lecture Notes in Comput. Sci.*, **2948**, Springer, Berlin, 2004.

[16] J.L. Nicolas A Gaussian law on $F_Q[X]$, *Topics in classical number theory, Vol. I, II* (Budapest, 1981), 1127–1162, Colloq. Math. Soc. Jnos Bolyai, **34**, North-Holland, Amsterdam, 1984.

[17] Irving Reiner, On the number of matrices with given characteristic polynomial. *Illinois J. Math.* **5** 1961 324–329.

[18] Richard Stong, The average order of a permutation. *Electron. J. Combin.***5** (1998), Research Paper 41, 6 pp. (electronic).

5 Appendix

At the suggestion of a referee, further details on the derivation of (8) from (7) are appended. For positive integers m, let $U_m = \sum_{\lambda \vdash m} \frac{LCM(\lambda_1, \lambda_2, \ldots)}{\lambda_1 \lambda_2 \cdots}$, so that the right side of (7) is $\frac{4}{c_\infty} \sum_{m=1}^{n} U_m$. Let $z = \sqrt{n}/\log^2 n$, and let B_n be the coefficient of x^n in $G(x) = \frac{1}{1-x} \prod_{\text{primes } p}^{\infty} (1 + x^p + \frac{x^{2p}}{2} + \frac{x^{3p}}{3} + \cdots)$.

In the middle of page 39 of [9] begins the proof that

$$U_n = O(n) T_1 T_2 T_3, \tag{9}$$

where

$$T_1 \leq n^z H_n^z = \exp\left(O(\frac{\sqrt{n}}{\log n}) \right), \tag{10}$$

$$T_3 \leq H_n^{\log^4 n} = \exp\Big(O(\log^4 n \log\log n) \Big), \tag{11}$$

and

$$T_2 \leq B_n. \tag{12}$$

Thus

$$U_n \leq B_n \exp\left(O(\frac{\sqrt{n}}{\log n}) \right). \tag{13}$$

It is precisely the numbers B_n that were estimated, using a Tauberian theorem[3], in §4 of [18]. (In Stong's notation, $B_n = \sum_{k=1}^{n} a_k$, and $h(t) = (1 - e^{-t}) G(e^{-t})$). His estimate was

$$B_n = \exp\left(B\sqrt{n/\log n} + O(\frac{\sqrt{n}\log\log n}{\log n}) \right). \tag{14}$$

Due to the rapid growth of the numbers B_m, summation does not change the error term in our estimate: combining (14) with (13), we get

$$\sum_{m=1}^{n} U_m \leq n \max_{m \leq n} U_m = \exp(B\sqrt{n/\log n} + O(\frac{\sqrt{n}\log\log n}{\log n})).$$