A note on the van der Waerden conjecture on random polynomials with symmetric Galois group for function fields

Erich L. Kaltofen

Dept. of Mathematics, North Carolina State University
Dept. of Computer Science, Duke University
kaltofen@ncsu.edu; https://users.cs.duke.edu/~elk27/

Abstract

Let $f(x) = x^n + (a_{n-1}t + b_{n-1})x^{n-1} + \cdots + (a_0 t + b_0)$ be of constant degree $n \geq 2$ in x and degree ≤ 1 in t, where all a_i, b_i are randomly and uniformly selected from a finite field $\mathbb{F}_q$ of q elements. Then the probability that the Galois group of f over $\mathbb{F}_q(t)$ is the symmetric group S_n on n elements is $1 - O(1/q)$. Furthermore, the probability that the Galois group of $f(x)$ over $\mathbb{F}_q(t)$ is not S_n is $\geq 1/q$ for $n \geq 3$ and $> 1/q - 1/(2q^2)$ for $n = 2$.

1. Introduction

Let n be an integer constant ≥ 2. The van der Waerden conjecture states that

$$\text{Prob}\Big(\text{Galois group over } \mathbb{Q} \text{ of } x^n + \sum_{i=0}^{n-1} a_i x^i \text{ is not the symmetric group } S_n \;\Big|\; a_i \in \mathbb{Z} \text{ and } |a_i| \leq H\Big) = O\big(\frac{1}{H}\big). \quad (1)$$

In [Bhargava 2021] there is the first proof and an extensive bibliography; see also [Anderson, Gafni, Oliver, Lowry-Duda, Shakan, and Zhang 2021]. Note that the probability (1) is asymptotically sharp: for $a_0 = 0$ all polynomials are reducible and have a smaller Galois group. From (1) one can derive (see Section 3) the following function field analog:

$$\text{Prob}\Big(\text{Galois group over } \mathbb{Q}(t) \text{ of } x^n + \sum_{i=0}^{n-1} (a_i t + b_i) x^i \text{ is not the symmetric group } S_n \;\Big|\; a_i, b_i \in \mathbb{Z} \text{ and } |a_i| \leq H, |b_i| \leq H\Big) = O\big(\frac{1}{H^2}\big). \quad (2)$$

Again, the probability (2) is asymptotically sharp. Here we consider the coefficient field $\mathbb{F}_q(t)$ where $\mathbb{F}_q$ is a finite field with $q = p^\ell$ elements, for $\ell \geq 1$ and the prime characteristic $p \geq 2$. We prove that

$$\text{Prob}\Big(\text{Galois group over } \mathbb{F}_q(t) \text{ of } x^n + \sum_{i=0}^{n-1}(a_i t + b_i)x^i \text{ is not the symmetric group } S_n \\ \;\Big|\; a_i, b_i \in \mathbb{F}_q\Big) = O\big(\frac{1}{q}\big). \quad (3)$$

Again the probability (3) is asymptotically sharp: the probability that $\mathrm{GCD}(A(x), B(x)) \neq 1$ for $A(x) = \sum_{i=0}^{n-1} a_i x^i$ and $B(x) = x^n + \sum_{i=0}^{n-1} b_i x^i$ is exactly $1/q$ (see, for instance, [Benjamin and Bennett 2007]), so at least q^{2n-1} polynomials $x^n + \sum_{i=0}^{n-1}(a_i t + b_i)x^i$ have smaller Galois group. Note that the Galois group of a polynomial of degree $n \geq 3$ in $\mathbb{F}_q[x]$ is not the symmetric group S_n. For $n = 2$, one subtracts the $(q^2 - q)/2$ irreducible $x^2 + b_1 x + b_0$ at $a_1 = a_0 = 0$ from the count.

2. Proof of Probability Estimate (3)

Let K be a field and $f(x) \in \mathsf{K}[x]$ be a polynomial, not necessarily irreducible, over K of degree n with leading coefficient $= 1$. A splitting field $\mathsf{N} = \mathrm{SF}_{\mathsf{K}}(f)$ of f over K is constructed by a tower of fields

$$\mathsf{L}_0 = \mathsf{K} \subset \mathsf{L}_1 \subset \mathsf{L}_2 \subset \cdots \subset \mathsf{L}_\ell = \mathsf{N}, \quad \mathsf{L}_i = \mathsf{L}_{i-1}[y_i]/(g_i(y_i)), \quad i = 1, 2, \ldots, \ell, \quad (4)$$

where $y_1, \ldots, y_\ell$ are fresh variables and where $g_i(x) \in \mathsf{L}_{i-1}[x]$ is an irreducible factor in $\mathsf{L}_{i-1}[x]$ with $\deg_x(g_i) \geq 2$ of

$$f(x) \text{ if } i = 1 \quad \text{and} \quad \frac{f(x)}{(x - y_1)\ldots(x - y_{i-1})} \in \mathsf{L}_{i-1}[x] \text{ if } i \geq 2. \quad (5)$$

At index ℓ we have $f(x) = (x - \alpha_1)\cdots(x - \alpha_n)$ for $\alpha_i \in \mathsf{N}$. All fields N constructed in the manner (4) are isomporhic with an isomorphism that is the identity function on K. Note that for all $1 \leq i \leq \ell$ the fields L_i are the quotient rings $\mathsf{K}[y_1, \ldots, y_i]$ modulo the triangular set $g_1(y_1), g_2(y_1, y_2), \ldots, g_i(y_1, \ldots, y_i)$ over K. Arithmetic in L_i is done recursively as univariate polynomial residue arithmetic in $\mathsf{L}_{i-1}[y_i]/(g_i(y_i))$. By (4), $\mathsf{N} = \mathrm{SF}_{\mathsf{L}_i}(f)$ for all $0 \leq i \leq \ell$. The Galois group $\Gamma_{\mathsf{N}/\mathsf{K}}$ of $f(x)$ over K is the group of all field automorphisms $\psi\colon \mathsf{N} \longrightarrow \mathsf{N}$ with $\psi(a) = a$ for all $a \in \mathsf{K}$. Each automorphism ψ uniquely permutes the distinct roots of f: $\psi(\alpha_i) = \alpha_{\tau(i)}$, and if $f(x)$ is separable, which means all roots are distinct: $\alpha_i \neq \alpha_j$ for all $1 \leq i < j \leq n$, then $\tau \in S_n$, where S_n is the symmetric group of permutations on $1, \ldots, n$, and all permutations τ form a subgroup.

In [van der Waerden 1940, Section 61] the permutations τ in the Galois group of a separable polynomial f over a field K are characterized as follows.

Theorem 2.1. *Let $f(x) = \prod_{i=1}^{n}(x - \alpha_i) \in \mathsf{K}[x]$ where $\alpha_i \in \mathsf{N} = \mathrm{SF}_{\mathsf{K}}(f)$ with $\alpha_i \neq \alpha_j$ for all $1 \leq i < j \leq n$, and let*

$$F(z, u_1, \ldots, u_n) = \prod_{\sigma \in S_n} \Big(z - \big(\sum_{i=1}^{n} \alpha_{\sigma(i)} u_i\big)\Big) \in \mathsf{K}[z, u_1, \ldots, u_n]. \quad (6)$$

Furthermore, let F_1 be an irreducible factor of F in $\mathsf{K}[z, u_1, \ldots, u_n]$ such that $z - (\sum_{i=1}^{n} \alpha_i u_i)$ is a factor of F_1 over N. Then the permutations τ in the Galois group of f over K are exactly those permutations such that $z - (\sum_{i=1}^{n} \alpha_{\tau(i)} u_i)$ is a factor of F_1.

Note that the assumption that the roots α_i of f are distinct is a necessary condition. Let $\mathsf{K} = \mathbb{F}_2(t)$ and $f(x) = x^2 + t = (x + \sqrt{t})^2$. Then $F(z) = (z + u_1\sqrt{t} + u_2\sqrt{t})^2 = z^2 + u_1^2 t + u_2^2 t$, which is irreducible over $\mathbb{F}_2(t)[z, u_1, u_2]$, but the Galois group of $f(x)$ over $\mathbb{F}_2(t)$ has a single element. Because $x^2 + t$ is irreducible in $\mathbb{F}_q(t)[x]$, it is squarefree in $\mathbb{F}_q[x, t]$ but not squarefree (inseparable) over the algebraic closure of $\mathbb{F}_q(t)$.

For generic polynomials the Galois group is the full symmetric group for all fields.

Theorem 2.2. *For the generic polynomial $f^{[v]} = x^n + \sum_{i=0}^{n-1} v_i\, x^i$ over $\mathsf{K}^{[v]} = \mathsf{K}(v_0, \ldots, v_{n-1})$ the polynomial $F^{[v]}$ corresponding to (6) is a separable polynomial in z, hence $\partial F^{[v]}/\partial z \neq 0$, and an irreducible polynomial in $\mathsf{K}[z, u_1, \ldots, u_n, v_0, \ldots, v_{n-1}]$, for all fields K.*

Classically, one uses the Hilbert Irreducibility Theorem to count for which evaluations of the v_i at values in K one preserves irreducibility of F [Kobloch 1956]. For $\mathsf{K} = \mathbb{F}_q(t)$ we can use our effective Hilbert Irreducibility Theorems [Kaltofen 1985, 1995]. We have the following theorem.

Theorem 2.3. *Let $F(X_1, \ldots, X_m) \in \mathsf{K}[X_1, \ldots, X_m]$, K a field, have total degree δ and be irreducible. Assume that $\partial F/\partial X_m \neq 0$. Let $S \subseteq \mathsf{K}$ be a finite set, and let $a_2, \ldots, a_{m-1}$, $b_1, \ldots, b_{m-1}$ be randomly and uniformly sampled elements in S. Then the probability*

$$\operatorname{Prob}\Big(F(b_1, b_2, \ldots, b_{m-1}, z) \in \mathsf{K}[z] \text{ is of degree } \deg_{X_m}(F) \text{ and has discriminant} \neq 0$$
$$\text{and } F(t + b_1, a_2 t + b_2, \ldots, a_{m-1} t + b_{m-1}, z) \text{ is irreducible in } \mathsf{K}[t, z] \Big) \geq 1 - \frac{4\delta\ 2^{\delta}}{|S|}, \qquad (7)$$

where $|S|$ is the number of elements in the set S [Kaltofen 1985, Theorem 2 and its proof].

We apply Theorem 2.3 to

$$F^{[v]}(z, u_1, \ldots, u_n, v_0, \ldots, v_{n-1}) \in \mathsf{K}(u_1, \ldots, u_n)[z, v_0, \ldots, v_{n-1}], \qquad (8)$$

which is defined above for the generic $f^{[v]}(x)$. The leading coefficient of $F^{[v]}$ in z is $= 1$ and $F^{[v]}$ is irreducible over $\mathsf{K}(u_1, \ldots, u_n)$. We have for randomly and uniformly sampled $a_1, \ldots, a_{n-1}$, $b_0, \ldots, b_{n-1} \in S \subseteq \mathsf{K} \subset \mathsf{K}(u_1, \ldots, u_n)$ and

$$\overline{F^{[v]}}(z, u_1, \ldots, u_n, t) \stackrel{\text{def}}{=} F^{[v]}(z, u_1, \ldots, u_n, t + b_0, a_1 t + b_1, \ldots, a_{n-1} t + b_{n-1}) \qquad (9)$$

the probability estimate

$$\operatorname{Prob}\Big(\text{the discriminant of } \overline{F^{[v]}}(z, u_1, \ldots, u_n, t) \text{ in the variable } z \text{ is} \neq 0 \text{ and}$$
$$\overline{F^{[v]}}(z, u_1, \ldots, u_n, t) \text{ is irreducible in } \mathsf{K}[z, u_1, \ldots, u_n, t]\Big) \geq 1 - \frac{4\delta^{[v]}\ 2^{\delta^{[v]}}}{|S|}, \qquad (10)$$

where $\delta^{[v]}$ is the total degree of $F^{[v]}$ in $z, v_0, \ldots, v_{n-1}$. All polynomials $\bar{f}(x) = x^n + (a_{n-1}t + b_{n-1})x^{n-1} + \cdots + (a_0 t + b_0)$ for which $\overline{F^{[v]}}(z, u_1, \ldots, u_n, t)$ is irreducible and separable, the latter of which implies that $\bar{f}(x)$ is separable, have Galois group S_n over $\mathsf{K}(t)$. Because n is a constant, $\delta^{[v]}$ is a constant. For the actual probability estimate (3) we can set $\mathsf{K} = S = \mathbb{F}_q$ and $t = t/a_0$ and multiply the probability (10) by $(1 - 1/|S|)$ for $a_0 \neq 0$. The more specific evaluation $X_1 = t + b_1$ in Theorem 2.3 strengthens our effective Hilbert Irreducibility Theorem.

3. Remarks

Better estimates than (7) in terms of the degree for the effective Hilbert Irreduciblity Theorems for function fields are possible. An estimate $1 - O(\deg(F)^4/|S|)$ is in [Kaltofen 1995] for perfect fields K, which includes all $\mathbb{F}_q$.

The estimate (2) follows from (1) by counting the irreducible $F^{[v]}(z, u_1, \ldots, u_n)$ for $v_i = a_i + tb_i$ with integers bounded by $|a_i| \leq H$ and $|b_i| \leq H$ and the variable evaluation $t = 2H+1$ which implies $|v_i| \leq 2H^2 + 2H$, with $(2H+1)^2$ values for each v_i. The count implies that $\mathrm{GCD}(x^n + \sum_{i=0}^{n-1} a_i x^i, \sum_{i=0}^{n-1} b_i x^i) \neq 1$ occurs with probability $O(1/H^2)$ for fixed n.

Acknowledgment: I thank Theresa C. Anderson for her correspondence about the topic of the paper. This research was supported by the National Science Foundation under Grant CCF-1717100.

References

Anderson, Theresa C., Gafni, Ayla, Oliver, Robert J. Lemke, Lowry-Duda, David, Shakan, George, and Zhang, Ruixiang. Quantitative Hilbert irreducibility and almost prime values of polynomial discriminants, 2021. URL: https://arxiv.org/abs/2107.02914.

Benjamin, Arthur T. and Bennett, Curtis D. The probability of relatively prime polynomials. *Mathematics Magazine*, 80(3):196–202, 2007. URL: https://doi.org/10.1080/0025570X.2007.11953481.

Bhargava, Manjul. Galois groups of random integer polynomials and van der Waerden's Conjecture, 2021. URL: https://arxiv.org/abs/2111.06507.

Kaltofen, E. Effective Hilbert irreducibility. *Information and Control*, 66:123–137, 1985. URL: http://users.cs.duke.edu/~elk27/bibliography/85/Ka85_infcontr.pdf.

Kaltofen, E. Effective Noether irreducibility forms and applications. *J. Comput. System Sci.*, 50(2):274–295, 1995. URL: http://users.cs.duke.edu/~elk27/bibliography/95/Ka95_jcss.pdf.

Kobloch, Hans-Wilhelm. Die Seltenheit der reduziblen Polynome. *Jahresbericht d. DMV*, 59:12–20, 1956. URL: http://resolver.sub.uni-goettingen.de/purl?PPN37721857X_0059.

van der Waerden, B. L. *Moderne Algebra.* Springer Verlag, Berlin, 1940. English transl. publ. under the title "Modern algebra" by F. Ungar Publ. Co., New York, 1953.

4. Appendix

The norm $\mathrm{norm}_{\mathsf{N}/\mathsf{K}}(\beta(y_1,\ldots,y_\ell))$ of an element $\beta(y_1,\ldots,y_\ell) \in \mathsf{N}$ over K, where N is the splitting field (4) of a possible inseparable polynomial f, is defined recursively:

$$\mathrm{norm}_{\mathsf{L}_{\ell-1}/\mathsf{K}}\underbrace{\left(\prod_{j=1}^{k}\beta(y_1,\ldots,y_{\ell-1},\gamma_j)\right)}_{\mathrm{norm}_{\mathsf{N}/\mathsf{L}_{\ell-1}}(\beta(y_1,\ldots,y_\ell))\in\mathsf{L}_{\ell-1}{}^{1}} \in \mathsf{K}, g_\ell(x) = (x-\gamma_1)\cdots(x-\gamma_k), \gamma_i \in \mathsf{N}, \gamma_1 = y_\ell. \quad (11)$$

The definition (11) extends to the rational function fields $\mathsf{N}(X_1,\ldots,X_m)$ over $\mathsf{K}(X_1,\ldots,X_m)$, where we have the following theorem.

Theorem 4.1. *Let $G \in \mathsf{N}[X_1,\ldots,X_m]$ be an irreducible polynomial over N, where N is the splitting field (4) of a not necessarily separable polynomial. Then $\mathrm{norm}_{\mathsf{N}/\mathsf{K}}(G) = H^k$ where $H \in \mathsf{K}[X_1,\ldots,X_m]$ is irreducible over K and $k \geq 1$.*

Proof. Suppose $\mathrm{norm}_{\mathsf{N}/\mathsf{K}}(G) = H_1H_2$ with $H_1, H_2 \in \mathsf{K}[X_1,\ldots,X_m]$ and $\mathrm{GCD}(H_1,H_2) = 1$. Note that relatively primeness as an arithmetic property over K remains valid over N. Now suppose that $G(X_1,\ldots,X_m,y_1,\ldots,y_\ell)$ is an irreducible factor of H_1 over N. By definition (11) there exist roots $\gamma_i \in \mathsf{N}$ of $g_i(x)$ such that $G(X_1,\ldots,X_m,\gamma_1,\ldots,\gamma_\ell)$ divides H_2 over N. The field N is isomorphic to $\mathsf{K}(\gamma_1,\ldots,\gamma_\ell)$ by $\psi\colon y_i \mapsto \gamma_i$ and $\psi(a) = a$ for all $a \in \mathsf{K}$, so $G(X_1,\ldots,X_m,\gamma_1,\ldots,\gamma_\ell)$ divides $\psi(H_1) = H_1$ over N, which contradicts that H_1, H_2 are relatively prime. ☐

Note that for $\beta \in \mathsf{N}$ we have $\mathrm{norm}_{\mathsf{N}/\mathsf{K}}(x-\beta) = h(x)^k$ where $h(x) \in \mathsf{K}[x]$ is the irreducible minimum polynomial with $h(\beta) = 0$, which means that $\mathrm{norm}_{\mathsf{N}/\mathsf{K}}(\beta)$ is k-th power of the product of all conjugates of β over K, which are the roots of h with multiplicities. For a separable polynomial $f(x)$ and $\beta \in \mathsf{N} = \mathrm{SF}_{\mathsf{K}}(f)$, we have $\mathrm{norm}_{\mathsf{N}/\mathsf{K}}(\beta) = \prod_{\psi\in\Gamma_{\mathsf{N}/\mathsf{K}}}\psi(\beta)$, where $\Gamma_{\mathsf{N}/\mathsf{K}}$ is the Galois group as a group of field automorphisms.

Proof of Theorem 2.1. Let $F_1 = \big(z - (\sum_{i=1}^n \alpha_i u_i)\big)G_1$ with $G_1 \in \mathsf{N}[z,u_1,\ldots,u_n]$. Then $F_1 = \psi(F_1) = \big(z - (\sum_{i=1}^n \psi(\alpha_i)u_i)\big)\ \psi(G_1)$ for all $\psi \in \Gamma_{\mathsf{N}/\mathsf{K}}$. Because f is separable all $\sum_{i=1}^n \psi(\alpha_i)u_i$ are distinct, and therefore all $z - (\sum_{i=1}^n \psi(\alpha_i)u_i)$ divide F_1 over N, whose product is the norm in the splitting field $\mathsf{N}(z,u_1,\ldots,u_n)$ of $f(x)$ over $\mathsf{K}(z,u_1,\ldots,u_n)$, and therefore $\in \mathsf{K}[z,u_1,\ldots,u_n]$. ☐

Second proof of Theorem 2.1. By Theorem 4.1 the norm of $z - \sum_{i=1}^n \alpha_i u_i$ is $H(z,u_1,\ldots,u_n)^k$ with H irreducible in $\mathsf{K}[z,u_1,\ldots,u_n]$. The norm's discriminant in z is $\neq 0$ because the roots are distinct, which implies $k = 1$. ☐

[1] Note that $\mathrm{norm}_{\mathsf{N}/\mathsf{L}_{\ell-1}}(\beta(y_1,\ldots,y_\ell))$ is the Sylvester resultant of $g_\ell(x)$ and $\beta(y_1,\ldots,y_{\ell-1},x)$ with respect to the variable x.

Proof of Theorem 2.2. First, $f^{[v]}(x)$ is separable in x because it is irreducible over $\mathsf{K}^{[v]}$ and its derivative with respect to z is $\neq 0$. The univariate polynomial discriminant is a non-zero polynomial in the coefficients over fields of all characteristics, which is $\neq 0$ for exactly the separable polynomials. Therefore, $F^{[v]}$ is also separable in z implying that $\partial F^{[v]}/\partial z \neq 0$.

Let $\prod_{i=1}^{n}(x - w_i) = x^n + e_{n-1}(w_1, \ldots, w_n)x^{n-1} + \cdots + e_0(w_1, \ldots, w_n) \in \mathsf{K}[z, w_1, \ldots, w_n]$, where e_i are plus/minus the $(n - i)$'th elementary symmetric functions in fresh variables $w_1, \ldots, w_n$, and let $\bar{F}^{[v]}$ be $F^{[v]}$ evaluated at $v_i = e_i(w_1, \ldots, w_n)$. We have $\bar{F}^{[v]} = \prod_{\sigma \in S_n} \big(z - \sum_{i=1}^{n} w_{\sigma(i)} u_i\big)$. Now let $F_1^{[v]}$ be an irreducible factor of $F^{[v]}$ in $\mathsf{K}[z, u_1, \ldots, u_n, v_0, \ldots, v_{n-1}]$ and let $\bar{F}_1$ be $F_1^{[v]}$ evaluated at $v_i = e_i(w_1, \ldots, w_n)$. Then by definition of $\bar{F}^{[v]}$, there is a permutation $\tau \in S_n$ such that $z - (w_{\tau(1)}u_1 + \cdots + w_{\tau(n)}u_n)$ divides $\bar{F}_1$ with co-factor $\bar{G}_1 \in \mathsf{K}[z, u_1, \ldots, u_n, w_1, \ldots, w_n]$. Permuting the w_i's in that factorization of $\bar{F}_1$ does not change $\bar{F}_1$ and shows that $z - (w_{\sigma(1)}u_1 + \cdots + w_{\sigma(n)}u_n)$ divides $\bar{F}_1$ for all permutations $\sigma \in S_n$. Therefore F_1 has degree $n!$ in z. $\square$