Factoring polynomials in finite fields: an application of Lang-Weil to a problem in graph theory.

Katz, Nicholas M.

in: Mathematische Annalen | Mathematische Annalen | Periodical Issue | Article

625 - 638

Terms and Conditions

Contact:

Niedersächsische Staats- und Universitätsbibliothek
Digitalisierungszentrum
37070 Goettingen
Germany
Email: gdz@www.sub.uni-goettingen.de

Purchase a CD-ROM

Math. Ann. 286, 625–637 (1990)

Mathematische Annalen

Factoring polynomials in finite fields: an application of Lang–Weil to a problem in graph theory

Nicholas M. Katz

Department of Mathematics, Princeton University, Fine Hall, Princeton, NJ 08544, USA

Introduction

Recently Chung [Ch] has constructed some interesting directed graphs using finite fields, and proven non-trivial upper bounds for their diameters. Her proof combines a general graph theoretic upper bound for the diameter in terms of the eigenvalues of the "adjacency matrix", the observation that for her graphs the eigenvalues are certain character sums over finite fields, and an estimate [Ka] for these character sums that comes from the Riemann Hypothesis for varieties over finite fields.

In this paper, we study Chung's graphs by a direct algebraic geometric method, interpreting a given bound for diameter as the statement that each of a certain collection of algebraic varieties has a rational point, and then using Lang–Weil to prove the existence of the required rational points. In order to apply Lang–Weil in the situation at hand, we are naturally led to showing that certain explicit two-parameter families of polynomials in one indeterminate have Galois group the full symmetric group; so far as we are aware, these constitute new examples of such polynomials.

The basic set-up

Chung's construction is this: fix an integer $n \geqq 2$, a finite field $k, K/k$ the field extension of degree n, and $\xi \in K$ an element such that $K = k(\xi)$. Out of this data Chung builds the directed graph whose nodes are the elements α of $K^\times$, and in which there is a directed edge $\alpha \to \beta$ if and only if $\beta/\alpha = \xi + a$ for some $a \in k$.

From the point of view of graph theory it is natural to ask the following questions:

1) Is this directed graph connected?
2) If so, what is its diameter?
3) What is its girth (shortest length of a cycle)?

For Chung's graph, these questions amount to:

1) Do the elements $\{\xi + a\}_{a \in k}$ generate $K^\times$?

2) If so, what is the least integer d such that every element of $K^\times$ is the product of at most d factors each in $\{\xi + a\}_{a \in k}$?
3) What is the least integer $g \geqq 1$ such that 1 is the product of precisely g factors each in $\{\xi + a\}_{a \in k}$?

For the question 1) of connectedness, the estimate of [Ka] shows that Chung's graph is connected if $\mathrm{Card}(k) > (n-1)^2$, but connectedness does not seem to be known in general. Along this line, one might ask whether at least one among the elements $\{\xi + a\}_{a \in k}$ is a *generator* of $K^\times$?

For question 2) of diameter, Chung shows that for each fixed n there is an *explicit* constant $A(n)$ such that $d \leqq 2n+1$ provided that $\mathrm{Card}(k) \geqq A(n)$. She also points out that, by an elementary counting argument, we must have $d \geqq n+1$ provided $\mathrm{Card}(k) \geqq 6 - n$. In this paper we will show the existence of an *inexplicit* constant $B(n)$ such that $d \leqq n+2$ provided $\mathrm{Card}(k) \geqq B(n)$. In fact, we will show that every element α of $K^\times$ can be written as the product of *precisely* $n+2$ *distinct* factors each in $\{\xi + a\}_{a \in k}$, and we will give an estimate for the number of such representations. We will also show that for $n = 2$ our result $d = 4$ is best possible, but we cannot rule out the possibility that for $n \gg 0$ the actual truth is $d = n+1$.

For question 3) of girth, it is easy to see that $g \geqq n$ (otherwise the element ξ would satisfy an equation over k of degree $< n$), and that $g \leqq n+2$ (by our result above for $\alpha = 1$); one has $g = n$ if and only if there exist n not necessarily distinct elements $a_1, \ldots, a_n$ in k such that $\Pi(X - a_i) - 1$ is the irreducible polynomial for ξ over k.

First variation

Let us return to the basic data (n, k, K, ξ) which go into Chung's construction. Let us denote by $f(X) \in k[X]$ the monic irreducible polynomial of degree n satisfied by ξ over k; then we can recover K and ξ from the data (n, k, f) by

$$K = k[X]/(f), \quad \xi = \text{the image of } X \text{ in } K.$$

This observation suggests the followng variant of Chung's construction: given an integer $n \geqq 2$, a finite field k, and a polynomial $f \in k[X]$ of degree n, consider the n-dimensional commutative k-algebra $K := k[X]/(f)$, the element $\xi :=$ the image of X in K; denoting by $K^\times$ the multiplicative group of invertible elements of K, form the directed graph $\mathcal{G}(n, k, f)$ whose nodes are the elements α of $K^\times$, and in which there is a directed edge $\alpha \to \beta$ if and only if $\beta/\alpha = \xi + a$ for some $a \in k$.

It is plausible that in this generality there exists for each $n \geqq 2$ a constant $c(n)$ such that as soon as $\mathrm{Card}(k) \geqq c(n)$ this directed graph $\mathcal{G}(n, k, f)$ is connected and has diameter $\leqq n+2$, but we are unable to prove this. Jusqu'a nouvel ordre we need to make the additional assumption that the k-algebra K is *etale*, or, what is the same, that the polynomial f has n distinct roots in "the" algebraic closure of k. For brevity, we will say that such an f is 'etale of degree n". Of course over a finite (indeed perfect) field, any irreducible polynomial is etale in this sense, so we still have the original Chung construction as a particular case of ours.

Theorem 1. *For every integer $n \geqq 2$ there exists a constant $B(n)$ such that for any finite field k with* Card $(k) \geqq B(n)$, *and any polynomial $f \in k[X]$ which is etale of degree n, the directed graph $\mathcal{G}(n, k, f)$ is connected and has diameter $\leqq n+2$.*

To prove this, we will prove the stronger result that as soon as Card $(k) \geqq B(n)$, any element α of $K^\times$ can be written as the product of $n+2$ terms $\xi + a_i$, with all the $n+2$ a_i's *distinct* in k. The strategy for doing this is simple: suppose, given α, that we seek $n+2$ elements a_i in k such that

$$\alpha = \Pi(\xi + a_i) \text{ in } K^\times. \tag{*}$$

Now α, being in $K^\times$, may be viewed as the class modulo f of a polynomial $g \in k[X]$ with $(f, g) = 1$, and we may choose g to be monic of degree $n+2$ (first choose the "standard" polynomial representative of α of degree $< n$, and then add to it $b^{-1}X^2 f(X)$, where b is the leading coefficient of f). Then we can rewrite the above equation (*) as the congruence

$$g(X) \equiv \Pi(X + a_i) \text{ modulo } f(X). \tag{**}$$

Subtracting the two sides and dividing by f, we see that (**) holds if and only if there exist two elements $u, v \in k$ such that we have an equality of monic polynomials of degree $n+2$ in $k[X]$

$$g(X) + (uX + v)f(X) = \Pi(X + a_i). \tag{***}$$

(Notice that if u and v exist they are uniquely determined by the a_i.) What about the condition that the a_i be all distinct? This is none other than the condition that the polynomial $g(X) + (uX + v)f(X)$ be etale of degree $n+2$.

To summarize, then, in order to show that any element α of $K^\times$ can be written as the product of $n+2$ terms $\xi + a_i$ (resp., and with all the $n+2$ a_i's *distinct* in k), it is equivalent to show that for any monic $g \in k[X]$ of degree $n+2$ with $(f, g) = 1$, there exist $n+4$ elements $(u, v, a_1, \ldots, a_{n+2})$ in k such that (***) holds (resp., and such that $g(X) + (uX + v)f(X)$ is etale of degree $n+2$).

The key observation is that if we view $(u, v, a_1, \ldots, a_{n+2})$ as "unknowns", then equating like powers of X in (***) leads to $n+2$ equations in these $n+4$ variables, thus defining inside $\mathbb{A}^{n+4}$ a certain *algebraic variety* $V(n, k, f, g)$ over k; moreover, the condition that $g(X) + (uX + v)f(X)$ be etale of degree $n+2$ is precisely that its discriminant, a k-polynomial in u and v, be invertible, and thus this condition defines an *open subvariety* $U(n, k, f, g)$ of $V(n, k, f, g)$.

Thus a representation of $\alpha := g \bmod f$ as an $n+2$ fold product $\Pi(\xi + a_i)$ in $K^\times$, where the order of the a_i's counts, is none other than a k-rational point of the variety $V(n, k, f, g)$, and such a representation with all the a_i distinct is none other than a point of the open subvariety $U(n, k, f, g)$.

Theorem 2. *For each integer $n \geqq 2$ these exists a constant $C(n)$ with the following property: Given (n, k, f, g) with $f \in k[X]$ etale of degree n and $g \in k[X]$ monic of degree $n+2$ with $(f, g) = 1$, the variety $U(n, k, f, g)$ is a smooth, finite type, geometrically connected affine k-scheme of dimension 2, and we have the estimate*

$$|\text{Card}\,(U(n, k, f, g)(k)) - \text{Card}\,(k)^2| \leqq C(n)\,\text{Card}\,(k)^{3/2}.$$

This trivially gives the desired sharpening of Theorem 1, with $B(n) := C(n)^2$:

Corollary. *Given (n, k, f) with $f \in k[X]$ etale of degree n, and with $\mathrm{Card}(k) > C(n)^2$, any α in $K^\times$ has (counting order) at least*

$$\mathrm{Card}(k)^2 - C(n)\,\mathrm{Card}(k)^{3/2} > 0$$

representations as an $n+2$ fold product of distinct $\xi + a_i$'s.

Interlude: the Lang–Weil method

In this section, we recall the general Lang–Weil method (compare [Ka-2, Proof of 12.4.3]) of obtaining uniform estimates for the number of rational points on a variety over a finite field as that variety varies in a suitable family.

Theorem (Lang–Weil). *Let S be a scheme of finite type over $\mathbb{Z}$, $f: X \to S$ a morphism of finite type, and $d \geqq 1$ an integer. Suppose that for every finite field k, and every k-valued point $s_k \in S(k)$, the k-scheme fibre $X(s_k) := f^{-1}(s_k)$ is geometrically irreducible of dimension d. Then there exists a constant $A(X/S)$, with the property that for every finite field k, and every k-valued point $s_k \in S(k)$, we have the estimate*

$$|\mathrm{Card}(X(s_k)(k)) - \mathrm{Card}(k)^d| \leqq A(X/S)\,\mathrm{Card}(k)^{d-1/2}.$$

Proof. Writing S as the union of the two open sets $S[1/2]$ and $S[1/3]$, working separately on each and then taking for A the maximum of the constants obtained on each of these open sets, we are reduced to the case when there exists a prime number ℓ which is invertible on S. In this case, consider the constructible sheaves $R^i f_! \mathbb{Q}_\ell$ on S: these vanish for $i > 2d$ and $R^{2d} f_! \mathbb{Q}_\ell \approx \mathbb{Q}_\ell(-d)$. In view of Grothendieck's Lefschetz Trace Formula and Deligne's Weil II result that $R^i f_! \mathbb{Q}_\ell$ is mixed of weight $\leqq i$, our assertion is immediate if we take the constant $A(X/S)$ to be the sup, over all geometric points s of S, of the $\mathbb{Q}_\ell$-dimension of the stalk at s of the constructible sheaf

$$\bigoplus_{0 \leqq i \leqq 2d-1} R^i f_! \mathbb{Q}_\ell. \quad \text{QED}$$

Applying Lang–Weil: reduction to a problem in Galois theory

In this section we will explain how, for each fixed $n \geqq 2$, to apply Lang–Weil in the above form to the varieties $U(n, k, f, g)$.

Lemma. *There exists an affine scheme S_n of finite type over $\mathbb{Z}$ and a smooth affine morphism $h: U_n \to S_n$ of relative dimension 2 such that for any finite field k, the k-valued points of S_n are precisely the pairs (f, g) consisting of an $f \in k[X]$ which is etale of degree n and of a $g \in k[X]$ of degree $n+2$ which satisfies $(f, g) = 1$, and such that the fibre of $h: U_n \to S_n$ over such a k-valued point is the k-scheme $U(n, k, f, g)$.*

Proof. Over any ring B, a polynomial $f(X) = \Sigma b_i X^i$ *in* $R[X]$ will be said to be everywhere of degree n if $b_i = 0$ for $i > n$ and if $b_n \in R^\times$. Fix such an f: the B-algebra $B[X]/(f)$ is free of rank n. Given any element g in $B[X]$, we have $(f, g) = 1$ in $B[X]$ if and only if g is invertible in the B-algebra $B[X]/(f)$, or equivalently if and only if the element $\mathrm{Norm}(g)$, the norm taken from $B[X]/(f)$ to B, is

invertible in B. The universal f which is everywhere of degree n lives over the ring $\mathbb{Z}[r_0,\dots,r_n][1/r_n]$, where the r_i are independent indeterminates. Similarly, the universal monic polynomial $g(X)=\Sigma s_i X^i$ of degree $n+2$ lives over the polynomial ring $\mathbb{Z}[s_0,\dots,s_{n+1}]$. Over the ring

$$A_n := \mathbb{Z}[r_0,\dots,r_n,s_0,\dots,s_{n+1}][1/r_n]$$

we have the universal pair (f,g) consisting of an f which is everywhere of degree n and of a g which is monic of degree $n+2$.

Let us say that an f in $B[X]$, B an arbitrary ring, which is everywhere of degree n is everywhere etale of degree n if $(f, df/dX)=1$ in $B[X]$; then the universal pair (f,g) consisting of a polynomial f which is everywhere of degree n and everywhere etale of degree n and of a monic polynomial g of degree $n+2$ which satisfies $(f,g)=1$ lives over the ring $B_n := A_n[1/\mathrm{Norm}(g\cdot df/dX)]$, the norm from $A_n[X]/(f)$ to A_n. The desired scheme S_n is none other than $\mathrm{Spec}(B_n)$.

Now consider two independent indeterminates U, V, and the monic polynomial $h(X) := g(X)+(UX+V)f(X)$ of degree $n+2$ over the ring $B_n[U,V]$; the open set where $h(X)$ is everywhere etale of degree $n+2$ is the spec of the ring $C_n := B_n[U,V][1/\mathrm{Norm}(dh/dX)]$, the norm from $B_n[U,V][X]/(h)$ to $B_n[U,V]$.

Over any ring C, if we are given any monic polynomial h of degree $n+2$ which is everywhere etale of degree $n+2$, then the scheme of complete factorizations of h is represented by the spec of the C-algebra $D = C[A_1,\dots,A_{n+2}]/I$, where the A_i are independent indeterminates over C and I is the ideal of relations obtained by equating like powers of X in

$$h(X) = \Pi(X - A_i).$$

Because h is everywhere etale of degree $n+2$, D is finite etale Galois over C with structural group the symmetric group $\mathfrak{S}_{n+2}$.

Applying this to the ring C_n and h, we obtain as above a finite etale Galois C_n-algebra $D_n := C_n[A_1,\dots,A_{n+2}]/I$, where the A_i are independent indeterminates over C_n and I is the ideal of relations obtained by equating like powers of X in

$$h(X) = \Pi(X - A_i).$$

It is now tautological that if we define $U_n := \mathrm{Spec}(D_n)$, then the natural morphism $U_n \to S_n$ "view D_n as an B_n-algebra" is the required morphism. If we factor this morphism as

$$U_n := \mathrm{Spec}(D_n) \xrightarrow{(1)} \mathrm{Spec}(C_n) \xrightarrow{(2)} \mathrm{Spec}(B_n[U,V]) \xrightarrow{(3)} \mathrm{Spec}(B_n) := S_n,$$

we see that it is smooth and affine of relative dimension 2, being the composition of the finite etale morphism $\mathrm{Spec}(D_n)\to\mathrm{Spec}(C_n)$, the affine open immersion $\mathrm{Spec}(C_n)\to\mathrm{Spec}(B_n[U,V])$, and of the projection $\mathrm{Spec}(B_n[U,V])\to\mathrm{Spec}(B_n)$.

QED

In view of this lemma, Theorem 2 would be an immediate consequence of Lang–Weil if only we knew, for each $n \geqq 2$, the

Proposition. *All of the geometric fibres of $U_n \to S_n$ are irreducible.*

We now explain how to reduce this Proposition to a statement in Galois theory. Consider a geometric fibre, say over an algebraically closed field F. This means that we begin with a polynomial $f(X) \in F[X]$ of degree n which has n distinct roots in F (remember that F is algebraically closed), and with a monic $g(X) \in F[X]$ of degree $n+2$ which satisfies $(f,g)=1$. We then pass to the open set $\mathscr{U}$ of $\mathbb{A}^2 := \mathrm{Spec}(F[U,V])$ where $h(X) := g(X) + (UX+V)f(X)$ has invertible discriminant, and then to the finite etale $\mathfrak{S}_{n+2}$ covering $U(n,F,f,g)$ of $\mathscr{U}$ defined as the scheme of complete factorizations of h.

Let us admit temporarily the following Galois theoretic statement, and show how it implies the required irreducibility of $U(n,F,f,g)$.

Theorem 3. *Let F be an algebraically closed field, $n \geqq 1$ an integer, $f(X) \in F[X]$ a polynomial of degree n which has n distinct roots in F, $g(X) \in F[X]$ a polynomial of degree $n+2$ which satisfies $(f,g)=1$, and U, V independent indeterminates. Consider the polynomial*

$$h(X) := g(X) + (UX+V)f(X)$$

of degree $n+2$ over the field $F(U,V)$. The splitting field of $h(X)$ over $F(U,V)$ is galois with group the full symmetric group $\mathfrak{S}_{n+2}$.

If we admit Theorem 3, then in particular $h(X)$ has $n+2$ distinct roots in "the" algebraic closure of $F(U,V)$, and consequently its discriminant is not identically zero in $F(U,V)$. This nonvanishing means precisely that $\mathscr{U}$ is a *nonvoid* open set in $\mathbb{A}^2$. Now consider the finite etale $\mathfrak{S}_{n+2}$-covering $U(n,F,f,g) \to \mathscr{U}$; because $\mathscr{U}$ is normal, the total space $U(n,F,f,g)$ is normal, so equal to the disjoint union of its connected ($=$ irreducible, by normality) components $\mathscr{V}_i$, *each* of which is finite etale over $\mathscr{U}$. Passing to the generic point η of $\mathscr{U}$, we see that the generic fibre $U(n,F,f,g)_\eta$ over η is the disjoint union of the generic fibres $(\mathscr{V}_i)_\eta$, each of which is the spec of a field. Thus $U(n,F,f,g)$ is irreducible if and only if its generic fibre $U(n,F,f,g)_\eta$ is the spec of a field. This last condition means precisely that the $F(U,V)$-algebra $F(U,V)[A_1,\ldots,A_{n+2}]/I$, where I is the ideal of relations obtained by equating coefficients in

$$h(X) = (\text{leading coef. of } g(X))\Pi(X - A_i),$$

is a field, and this is in turn equivalent to saying that $h(X)$ is a separable irreducible polynomial over $F(U,V)$ with Galois group $\mathfrak{S}_{n+2}$.

Thus Theorem 3 does indeed imply the Proposition. QED

Galois theory variations

We begin by reducing Theorem 3 to

Theorem 3bis. *Let F be an algebraically closed field, $n \geqq 0$ an integer, $f(X) \in F[X]$ a polynomial of degree $n+1$ which has $n+1$ distinct roots in F, $g(X) \in F[X]$ a polynomial of degree $n+2$ which satisfies $(f,g)=1$, and U, V independent indeterminates. Consider the polynomial*

$$h(X) := g(X) + (UX+V)f(X)$$

of degree $n+2$ over the field $F(U,V)$. The splitting field of $h(X)$ over $F(U,V)$ is galois with group the full symmetric group $\mathfrak{S}_{n+2}$.

To see that Theorem 3bis implies Theorem 3, we argue as follows. By means of a preliminary change of variable $X \mapsto X+\alpha$ with $\alpha \in F$, we may assume in Theorem 3 that the polynomials f and g satisfy the additional hypothesis $f(0)g(0) \neq 0$. The splitting field of $h(X)$ is then the same as that of $X^{n+2}h(1/X)$; the identity

$$X^{n+2}h(1/X) = X^{n+2}g(1/X) + (U + VX)\cdot(X^{n+1}f(1/X)))$$

now puts us in the situation of Theorem 3bis for the same n, with U and V interchanged, and for the polynomials $\tilde{f} := X^{n+1}f(1/X)$ and $\tilde{g} := X^{n+2}g(1/X)$.

We now introduce new indeterminates

$$S := 1/U \quad \text{and} \quad T := -V/U$$

in Theorem 3bis, and divide through f, g, and h by the leading coefficient of f; Theorem 3bis becomes trivially equivalent to

Theorem 3ter. *Let F be an algebraically closed field, $n \geqq 0$ an integer, $f(X) \in F[X]$ a monic polynomial of degree $n+1$ which has $n+1$ distinct roots $a_1, \ldots, a_{n+1}$ in F, $g(X) \in F[X]$ a polynomial of degree $n+2$ which satisfies $(f,g)=1$, and S, T independent indeterminates. Consider the polynomial*

$$h(X) := Sg(X) + (X - T)f(X)$$

of degree $n+2$ over the field $F(S,T)$. The splitting field of $h(X)$ over $F(S,T)$ is galois with group the full symmetric group $\mathfrak{S}_{n+2}$.

Notice that in Theorem 3ter it is obvious that the polynomial $h(X)$ has nonvanishing discriminant, since this is true at $S=0$. Thus we can speak of the Galois group of its splitting field over $F(S,T)$.

Theorem 3ter results from the stronger

Theorem 4. *Let F be an algebraically closed field, $n \geqq 0$ an integer, $f(X) \in F[X]$ a monic polynomial of degree $n+1$ which has $n+1$ distinct roots $a_1, \ldots, a_{n+1}$ in F, $g(X) \in F[X]$ a polynomial of degree $n+2$ which satisfies $(f,g)=1$, and S, T independent indeterminates. Consider the polynomial*

$$h(X) := Sg(X) + (X - T)f(X)$$

of degree $n+2$ over the integral domain $F[T][[S]]$. The splitting field of $h(X)$ over the fraction field of $F[T][[S]]$ is galois with group the full symmetric group $\mathfrak{S}_{n+2}$. More generally, for any polynomial $p(X) \in F[X]$ which satisfies $(f,p)=1$, the splitting field of $h(X)$ over the fraction field of $F[T, 1/p(T)][[S]]$ is galois with group the full symmetric group $\mathfrak{S}_{n+2}$.

That Theorem 4 implies 3ter is obvious, since after the extension of ground field from $F(S,T)$ to the fraction field of $F[T][[S]]$ the Galois group can only decrease.

Before begininning the proof of Theorem 4, we explain a general criterion that a given polynomial $h(X)$ of degree $n+2$ with nonzero discriminant over a given

field K have Galois group $\mathfrak{S}_{n+2}$ over K. Suppose that in *some* overfield L of K, our polynomial h factors completely, with roots $\alpha_0, \alpha_1, \ldots, \alpha_{n+1}$. Because $\mathfrak{S}_{n+2}$, viewed as the group of permutations on $\{0, \ldots, n+1\}$, is generated by the $n+1$ transpositions $(0,1),(0,2),\ldots,(0,n+1)$, it suffices to show that for any integer i in $\{1,\ldots,n+1\}$, the field $K(\alpha_1,\ldots,\hat{\alpha}_i,\ldots,\alpha_{n+1})$ obtained by adjoining to K all the roots *except* for α_0 and α_i does *not* contain α_0.

We will prove Theorem 4 by using a standard form of Newton's Lemma to "write down" all the roots, and then verifying "by inspection" the above criterion.

Newton's Lemma. *Let R be a ring, and $\mathfrak{J}$ an ideal of R. Suppose that R is complete and separated for the $\mathfrak{J}$-adic topology. Suppose we are given a polynomial $h(X) \in R[X]$. Denote by $\bar{h}(X) \in (R/\mathfrak{J})[X]$ the reduction* mod $\mathfrak{J}$ *of $h(X)$, and by $\bar{h}'(X)$ its derivative. Suppose we are given an element $\bar{\alpha}$ in $R/\mathfrak{J}$ which satisfies*

$$\bar{h}(\bar{\alpha}) = 0 \text{ in } R/\mathfrak{J}, \quad \bar{h}'(\bar{\alpha}) \text{ is a unit in } R/\mathfrak{J}.$$

Then there exists a unique element α in R which satisfies

$$h(\alpha) = 0 \text{ in } R, \quad \alpha \equiv \bar{\alpha} \bmod \mathfrak{J}.$$

Proof. The construction-proof proceeds modulo successive powers of $\mathfrak{J}$ by induction. If $\alpha_n \in R$ reduces mod $\mathfrak{J}$ to $\bar{\alpha}$ and satisfies $h(\alpha_n) \equiv 0 \bmod \mathfrak{J}^n$, for some $n \geqq 1$, then $\alpha_{n+1} := \alpha_n - h(\alpha_n)/h'(\alpha_n)$ is congruent to α_n mod $\mathfrak{J}^n$, satisfies $h(\alpha_{n+1}) \equiv 0 \bmod \mathfrak{J}^{n+1}$, and is unique mod $\mathfrak{J}^{n+1}$ with these properties. QED

We now turn to the proof of Theorem 4. We will first apply Newton's Lemma to our $h(X)$ over the ring $R := F[T, 1/p(T)f(T)][[S]]$, the ideal $\mathfrak{J} := (S)$. The factorization of $\bar{h}(X)$ into $n+2$ linear factors

$$\bar{h}(X) = (X - T)(X - a_1)\cdots(X - a_{n+1})$$

gives us $n+2$ roots $T, a_1, \ldots, a_{n+1}$ of $\bar{h}$, at each of which the value of $\bar{h}'$ is invertible in $F[T, 1/p(T)f(T)]$; Newton's Lemma then refines $T, a_1, \ldots, a_{n+1}$ into zeroes $\tilde{T}, \alpha_1, \ldots, \alpha_{n+1}$ of $h(X)$ in $F[T, 1/p(T)f(T)][[S]]$ which reduce mod S to $T, a_1, \ldots, a_{n+1}$ respectively. Thus we may take for L the fraction field of $F[T, 1/p(T)f(T)][[S]]$.

In order to verify the criterion, we must show, denoting by K the fraction field of $F[T, 1/p(T)][[S]]$, that for any $1 \leqq i \leqq n+1$, the field $K(\alpha_1,\ldots,\hat{\alpha}_i,\ldots,\alpha_{n+1})$ obtained by adjoining to K all the roots *except* for $\tilde{T}$ and α_i does *not* contain $\tilde{T}$. Renumbering the roots of f, we see that it suffices to treat the case $i = 1$.

To see this, we argue as follows. Let us denote by $f_1(X) \in F[X]$ the polynomial

$$f_1(X) := f(X)/(X - a_1) = (X - a_2)\cdots(X - a_{n+1}).$$

Then over the ring $F[T, 1/p(T)f_1(T)][[S]]$ with the S-adic topology we can apply Newton's Lemma to $h(X)$, but only to the roots $a_2, \ldots, a_{n+1}$ of $\bar{h}$; at each of these roots the value of $\bar{h}'$ is invertible in $F[T, 1/p(T)f_1(T)]$. Newton's Lemma then refines $a_2, \ldots, a_{n+1}$ into zeroes $\tilde{\alpha}_2, \ldots, \tilde{\alpha}_{n+1}$ of $h(X)$ in $F[T, 1/p(T)f_1(T)][[S]]$ which reduce mod S to $a_2, \ldots, a_{n+1}$ respectively.

By the *uniqueness* in Newton's Lemma, and the fact that $F[T, 1/p(T)f_1(T)][[S]]$

is a *subring* of $F[T, 1/p(T)f(T)][[S]]$, we have

$$\alpha_i = \tilde{\alpha}_i \quad \text{for} \quad i = 2, \ldots, n+1.$$

It remains to see that $\tilde{T}$ does not lie in $K(\alpha_2, \ldots, \alpha_{n+1})$; as this field is contained in the fraction field of $F[T, 1/p(T)f_1(T)][[S]]$ by the analysis above, it suffices to show that $\tilde{T}$ does not lie in the fraction field of $F[T, 1/p(T)f_1(T)][[S]]$. Because the ring $F[T, 1/p(T)f_1(T)][[S]]$ is *normal*, and $h(X)$ has leading coefficient a unit (namely $1+S$), we see that if $\tilde{T}$ were to lie in this fraction field, then $\tilde{T}$ would necessarily lie in $F[T, 1/p(T)f_1(T)][[S]]$ itself.

To show that this is not the case, we will use the first step of Newton's method to write $\tilde{T}$ in the larger ring $F[T, 1/p(T)f(T)][[S]]$, and show that, when we develop $\tilde{T}$ as a series in S, already the coefficient of S does not lie in $F[T, 1/p(T)f_1(T)]$. Indeed, Newton's method tells us that the root $\tilde{T}$ near the approximate root T of $h(X) := Sg(X) + (X - T)f(X)$ is given modulo $S^2F[T, 1/p(T)f(T)][[S]]$ by

$$\begin{aligned}\tilde{T} &\equiv T - h(T)/h'(T) \\ &\equiv T - Sg(T)/(Sg'(T) + f(T)) \\ &\equiv T - (g(T)/f(T))S.\end{aligned}$$

That the coefficient of S, $g(T)/f(T)$, does not lie in $F[T, 1/p(T)f_1(T)]$ is now obvious: the function $g(T)/f(T)$ has a simple pole at $T = a_1$, because $(f, g) = 1$, while every element of $F[T, 1/p(T)f_1(T)]$ is holomorphic at $T = a_1$, because both f_1 and p are invertible at $T = a_1$, f_1 because f has all distinct roots and p because it is prime to f. This concludes the proof of Theorem 4. QED

Here is an "arithmetic" variant on Theorem 4.

Theorem 4bis. *Let R be a complete discrete valuation ring with uniformizing parameter π and algebraically closed residue field F, $n \geqq 0$ an integer, $f(X) \in R[X]$ a monic polynomial of degree $n+1$ whose reduction* $\bmod \pi$ *$\bar{f}(X) \in F[X]$ has $n+1$ distinct roots $a_1, \ldots, a_{n+1}$ in F, $g(X) \in R[X]$ a polynomial everywhere of degree $n+2$ which satisfies $(\bar{f}, \bar{g}) = 1$, and T an indeterminate. Consider the polynomial*

$$h(X) := \pi g(X) + (X - T)f(X)$$

of degree $n+2$ over

$$R\langle\langle T \rangle\rangle := \text{the } \pi\text{-adic completion of } R[T].$$

The splitting field of $h(X)$ over the fraction field of $R\langle\langle T \rangle\rangle$ is galois with group the full symmetric group $\mathfrak{S}_{n+2}$. More generally, for any polynomial $p(X) \in R[X]$ which satisfies $(\bar{f}, \bar{p}) = 1$, the splitting field of $h(X)$ over the fraction field of the π-adic completion of $R[T, 1/p(T)]$ is galois with group the full symmetric group $\mathfrak{S}_{n+2}$.

Proof. The proof via Newton's Lemma is essentially identical to that given above of Theorem 4 (itself the special case $R = F[[S]]$, $\pi = S$, of Theorem 4bis). QED

Corollary. *Let $n \geqq 0$ be an integer, $f(X) \in \mathbb{Z}[X]$ a monic polynomial of degree $n+1$ whose discriminant $\delta \in \mathbb{Z}$ is nonzero, and whose constant term $f(0)$ is nonzero. Let T be an indeterminate. Then for any prime number p which does not devide $\delta \cdot f(0)$, the*

polynomial

$$h(X) := pX^{n+2} + (X - T)f(X)$$

over $\mathbb{Q}(T)$ *has Galois group the full symmetric group* $\mathfrak{S}_{n+2}$.

Proof. Indeed if we denote by W the Witt vectors of an algebraically closed field of characteristic p, then by Theorem 4bis the polynomial $h(X)$ still has Galois group $\mathfrak{S}_{n+2}$ over the fraction field of $W\langle\langle T\rangle\rangle$. QED

Second variation

In this section we will study a second variant of Chung's graph: given an integer $n \geqq 2$, a finite field k, and a polynomial $f \in k[X]$ of degree n, consider as before the n-dimensional commutative k-algebra $K := k[X]/(f)$, and the element $\xi :=$ the image of X in K; denoting by $K^\times$ the multiplicative group of invertible elements of K, form the directed graph $\mathcal{2G}(n, k, f)$ whose nodes are the elements α of $K^\times$, and in which there is a directed edge $\alpha \to \beta$ if and only if $\beta/\alpha = b(\xi + a)$ for some pair $a \in k$, $b \in k^\times$.

Lemma. *The directed graph* $\mathcal{2G}(n, k, f)$ *has girth* $g \geqq n$; *it has diameter* $d \geqq n$ *unless* $n = 2$ *and* k *is* $\mathbb{Z}/2\mathbb{Z}$.

Proof. If the girth $g < n$, then ξ satisfies an equation over k of too low degree. If $k^\times$ contains an element $b \neq 1$, then the diameter $d \geqq n$ because b cannot be written as a product of fewer than n terms for the same reason. If $n \geqq 3$, then $d \geqq n$ because otherwise every nonconstant polynomial $g(X)$ in $k[X]$ of degree $< n$ with $(f, g) = 1$ would factor completely over k; but taking for $g(X)$ a k-irreducible polynomial of degree $n - 1 \geqq 2$ shows that this is not the case. QED

Theorem 6. *For every integer* $n \geqq 2$ *there exists a constant* $D(n)$ *such that for any finite field* k *with* Card $(k) \geqq D(n)$, *and any polynomial* $f \in k[X]$ *which is etale of degree* n, *the directed graph* $\mathcal{2G}(n, k, f)$ *is connected and has diameter* $d \leqq n + 1$.

To prove this, we will prove the stronger result that as soon as Card $(k) \geqq D(n)$, any element α of $K^\times$ can be written as the product of an element b in $k^\times$ and of $n + 1$ terms $\xi + a_i$, with all the $n + 1$ a_i's *distinct* in k. The strategy for doing this is simple: suppose given α, that we seek $n + 1$ elements a_i in k and an element b in $k^\times$ such that

$$\alpha = b\Pi(\xi + a_i) \text{ in } K^\times. \qquad (*)$$

Now α, being in $K^\times$, may be viewed as the class modulo f of a polynomial $g \in k[X]$ with $(f, g) = 1$, and we may choose g to be monic of degree $n + 1$ (first choose the "standard" polynomial representative of α of degree $< n$, and then add to it $c^{-1}Xf(X)$, where c is the leading coefficient of f). Then we can rewrite the above equation (*) as the congruence

$$g(X) \equiv b\Pi(X + a_i) \text{ modulo } f(X). \qquad (**)$$

Subtracting the two sides and dividing by f, we see that (**) holds if and only if

there exist two elments $u, v \in k$ such that we have an equality of polynomials of degree $n+1$ in $k[X]$

$$g(X) + (uX + v) f(X) = b\Pi(X + a_i). \qquad (***)$$

(Notice that if u and v exist they are uniquely determined by b and the a_i.) What about the condition that the a_i be all distinct? This is none other than the condition that the polynomial $g(X) + (uX + v) f(X)$ be etale of degree $n+1$.

The key observation is that if we view $(u, v, b, a_1, \ldots, a_{n+1})$ as "unknowns", then equating like powers of X in $(***)$ leads to $n+2$ equations in these $n+4$ variables, thus defining inside $\mathbb{A}^{n+4}$ a certain *algebraic variety* $\mathcal{2}V(n, k, f, g)$ over k; moreover, the condition that $g(X) + (uX + v) f(X)$ be etale of degree $n+1$ is precisely that its discriminant, a k-polynomial in u and v, be invertible, and thus this condition defines an *open subvariety* $\mathcal{2}U(n, k, f, g)$ of $\mathcal{2}V(n, k, f, g)$.

Thus a representation of $\alpha := g \bmod f$ of type $(*)$, where the order of the a_i's counts, is none other than a k-rational point of the variety $\mathcal{2}V(n, k, f, g)$, and such a representation with all the a_i distinct is none other than a point of the open subvariety $\mathcal{2}U(n, k, f, g)$. Thus Theorem 6 (with $D(n) := e(n)^2$) results from

Theorem 7. *For each integer $n \geqq 2$ there exists a constant $e(n)$ with the following property: Given (n, k, f, g) with $f \in k[X]$ etale of degree n and $g \in k[X]$ monic of degree $n+1$ with $(f, g) = 1$, the variety $\mathcal{2}U(n, k, f, g)$ is a smooth, finite type, geometrically connected affine k-scheme of dimension 2, and we have the estimate*

$$|\mathrm{Card}(\mathcal{2}U(n, k, f, g)(k)) - \mathrm{Card}(k)^2| \leqq e(n)\,\mathrm{Card}(k)^{3/2}$$

and its

Corollary. *Given (n, k, f) with $f \in k[X]$ etale of degree n, and with $\mathrm{Card}(k) > e(n)^2$, any α in $K^\times$ has (counting order of the $\xi + a_i$) at least*

$$\mathrm{Card}(k)^2 - e(n)\,\mathrm{Card}(k)^{3/2} > 0$$

representations as the product of an element of $k^\times$ and of an $n+1$ fold product of distinct $\xi + a_i$'s.

To prove Theorem 7 we have only to recopy the steps in the proof of Theorem 2 by means of Lang–Weil, applying Theorem 3bis (with $n := n-1$) to our (f, g) to know that the variety $\mathcal{2}U(n, k, f, g)$ is smooth and geometrically irreducible of dimension two. QED

Examples

We first give a family of examples showing that for $n = 2$, the Theorem 1 cannot be improved. Let k be a finite field, and take for $f(X)$ the quadratic polynomial $X^2 + 1$. The image of X in $K := k[X]/(f)$ will be denoted i rather than ξ.

Lemma. *Notations as above, suppose that none of the numbers*

$$-1, 15, 27$$

is a square in k (a condition which by quadratic reciprocity is fulfilled if k is $\mathbb{Z}/p\mathbb{Z}$

for any $p \equiv 19$ or $31 \bmod 60$). Then the element $8i$ of $K^\times$ cannot be represented as the product of fewer than 4 elements of the form $i + a$ with $a \in k$.

Proof. To see this, we first observe that $8i$ is not itself of the form $i + a$ with $a \in k$. We next consider the possibility that for some a, b in k we have

$$8i = (i + a)(i + b) \text{ in } K.$$

This amounts to the pair of equations

$$8 = a + b, \quad ab = 1,$$

which is to say

$$8 = a + (1/a) \text{ with } a \text{ in } k^\times,$$

and this equation has a solution in k if and only if 15 is a square in k. Finally we consider the possibility that for some a, b, c in k we have

$$8i = (i - a)(i - b)(i - c) \text{ in } K.$$

Writing $8i = i^3 + 9i$, this becomes

$$X^3 + 9X \equiv (X - a)(X - b)(X - c) \bmod (X^2 + 1),$$

i.e., there exist a, b, c, d in k such

$$X^3 + 9X + d(X^2 + 1) = (X - a)(X - b)(X - c).$$

Consider the discriminant δ of the left hand polynomial: the alleged factorization would give

$$\delta = (a - b)^2(a - c)^2(b - c)^2,$$

but a tedious direct calculation shows that

$$\delta = -(2(d^2 - 27))^2.$$

Since -1 is not a square in k, these two equations for δ show that δ must vanish; but if δ vanishes then d must be a square root of 27 in k, and k contains no such element. QED

We next give a family of examples showing that for $n = 3$ the Theorem 6 cannot be improved. (Unfortunately K is not a field in these examples.)

Let k be a finite field, and take for f the cubic polynomial $X^3 + 9X$. The image of X in $K := k[X]/(f)$ is denoted ξ.

Lemma. *Notations as above, suppose that none of the numbers*

$$-1, 15, 27$$

is a square in k (a condition which by quadratic reciprocity is fulfilled if k is $\mathbb{Z}/p\mathbb{Z}$ for any $p \equiv 19$ or $31 \bmod 60$). Then no $k^\times$-multiple of the element $\xi^2 + 1$ of $K^\times$ can be represented as the product of fewer than 4 elements of the form $\xi + a$ with $a \in k$.

Proof. If $b(\xi^2 + 1) = (\xi + a_1)$ or $(\xi + a_1)(\xi + a_2)$ or $(\xi + a_1)(\xi + a_2)(\xi + a_3)$ in K,

with $b \in k^\times$ and the $a_i \in k$, then modulo $(X^3 + 9X)k[X]$ we have a congruence

$$b(X^2+1) \equiv (X+a_1) \text{ or } (X+a_1)(X+a_2) \text{ or } (X+a_1)(X+a_2)(X+a_3).$$

In the first two cases this forces, for degree reasons, an equality

$$b(X^2+1) = (X+a_1) \text{ or } (X+a_1)(X+a_2) \text{ in } k[X],$$

which is visibly impossible, since X^2+1 is of degree 2 and k-irreducible. In the third case it means that

$$X^3 + 9X + b(X^2+1) = (X-a_1)(X-a_2)(X-a_3),$$

and this is impossible by the first Lemma above. QED

Questions

(1). What happens if we drop the hypothesis that f be etale, and insist only that $(f,g)=1$? The "worst case" $f(X) := X^n$, when $K^\times = (k[X]/(X^n))^\times$ is the product of $k^\times$ with a group of truncated Witt vectors, already seems quite interesting.
(2). How far is Theorem 6 from the truth? We have already pointed out that over a large field k, we must have $d \geqq n$. In the case $n=2$, however, it is trivially true that the diameter is 2 rather than 3. On the other hand, the example above shows that for $n=3$ Theorem 6 is best possible, at least in the context of etale f's. It still leaves open the possibility that for $n \gg 0$ and $\mathrm{Card}(k) \geqq D(n)$, one has $d=n$.
(3). What are the smallest values of the constants $B(n)$ and $D(n)$ for which Theorems 1 and 6 respectively are true?

References

[Ch] Chung, F.R.K.: Diameters and eigenvalues. (to appear)
[De] Deligne, P.: La conjecture de Weil II. Publ. Math. IHES **52**, 313–428 (1981)
[Ka] Katz, N.M.: An estimate for character sums. J. Am. Math. Soc. **2**, 197–200 (1989)
[Ka-2] Katz, N.M.: Gauss sums, Kloosterman sums and monodromy groups. Annals of Math. Study **116**. Princeton: Princeton University Press 1988
[L–W] Lang, S., Weil, A.: Numbers of points of varieties in finite fields. Am. J. Math. **76**, 819–827 (1954)

Received September 2, 1988