INFORMATION AND COMPUTATION 90, 139–155 (1991)

Toward a Theory of Pollard's Rho Method

ERIC BACH*

Computer Sciences Department, University of Wisconsin, Madison, Wisconsin 53706

Pollard's "rho" method for integer factorization iterates a simple polynomial map and produces a nontrivial divisor of n when two such iterates agree modulo this divisor. Experience and heuristic arguments suggest that a prime divisor p should be detected in $O(\sqrt{p})$ steps, but this has never been proved. Indeed, nothing seems to be have been rigorously proved about the probability of success that would improve the obvious lower bound of $1/p$. This paper shows that for fixed k, this probability is at least $\binom{k}{2}/p + O(p^{-3/2})$ as $p \to \infty$. This leads to an $\Omega(\log^2 p)/p$ estimate of the success probability.

1. INTRODUCTION

In 1975, J. M. Pollard published his famous "rho" method for integer factorization. The algorithm is simple, elegant, and often used in practice when a brute-force search for divisors fails.

However, very little is known in a rigorous sense about why it works. Experience and probabilistic intuition indicate that it will remove a prime factor p from n after about $\sqrt{p}$ steps; arguments to this effect have been given by Brent (1980), Brent and Pollard (1981), Gold and Sattler (1983), Guy (1976), Montgomery (1987), Pollard (1975), and Riesel (1982). However, this running time bound has never been proved. Indeed, nothing seems to have been shown that improves the obvious bound that the probability of success is at least $1/p$. For this reason alone, it is of interest to see what can be proved rigorously about the algorithm.

Pollard's method is the followng. Choose integers x and y and compute the sequence

$$f_0 = x, \qquad f_1 = x^2 + y, \qquad f_2 = (x^2 + y)^2 + y, \ldots$$

modulo n. A divisor of n modulo which two such iterates collide can be found by examining $\gcd(f_{2i+1} - f_i, n)$, $i = 0, 1, 2, \ldots$. Proceeding in this manner allows one to avoid storing all the iterates, since the pair (f_{2i+1}, f_i) can be easily computed from (f_{2i-1}, f_{i-1}) by the recurrence relation.

* This research was supported by National Science Foundation Grants DCR-8504485 and DCR-8552596.

0890-5401/91 $3.00

Pollard originally suggested using $x=2$, $y=1$. He also noted that the values $y=0$ and $y=-2$ should be avoided, because of the simple closed forms for the iterates f_i in these cases.

I show below that if x and y are chosen at random subject to $0 \leqslant x, y < n$, then the probability that a prime factor p is discovered before the kth step of this process is, for k fixed and $p \to \infty$, at least

$$\binom{k}{2} \Big/ p + O(p^{-3/2}).$$

Analyzing the dependence of this bound on k allows the probability of a successful factorization to be raised to $\Omega(\log^2 p)/p$. I believe this to be the first result on Pollard's algorithm that does not rely on some heuristic assumptions.

The main new idea in the proof is to associate with each pair $i<j$ a polynomial $\rho_{i,j}$ whose "generic" roots are pairs (x, y) for whom the first collision occurs when $f_i = f_j$ (see Section 3). These polynomials have integer coefficients (see Section 4); this fact allows one to use results of Weil to estimate the success probability (see Section 5). The polynomials $\rho_{i,j}$ obey an interesting "exclusion principle," which states roughly that two such polynomials corresponding to relatively prime cycle lengths will not have common zeros (see Section 6).

The results of this paper show that the first few iterates behave roughly as independently chosen random numbers. It is customary to use this unrealistic assumption when dealing with a large number of trials, but one is always interested in reducing the amount of disbelief that must be suspended in such analyses. To this end, I show that the observed running time of the rho method follows from two less stringent assumptions: first, that the average number of points per curve $\rho_{i,j}$ is not too much smaller than its "expected" value, and that the total number of points lying on pairs of such curves approximates a law of "average pairwise independence" (see Section 7). These assumptions are concrete and capable of being proved or disproved, unlike that of stochastic independence.

2. Notation and Background

In this section I list the algebraic facts (and their intuitive geometric interpetations) that will be needed later. Mostly the results are standard; appropriate references include the books of van der Waerden (1970), Zariski and Samuel (1958) (for algebra), Fulton (1969), Brieskorn and Knörrer (1986) (for curves), Hartshorne (1977), and Fried and Jarden (1986) (for varieties over finite fields).

$\mathbb{Z}$ denotes the integers, $\mathbb{Q}$ the rational numbers, $\mathbb{C}$ the complex numbers,

and $\mathbb{F}_q$ a finite field with q elements. $A[x_1, ..., x_n]$ denotes the polynomial ring in n indeterminates with coefficients in the ring A. If A is a unique factorization domain (UFD), so is $A[x_1, ..., x_n]$; extending the coefficient ring to include fractions does not affect the irreducibility of a polynomial. Similarly $A[[x_1, ..., x_n]]$ denotes the ring of formal power series; this is also a UFD if A is a field.

Let k be a field, and $f \in k[x, y]$ be a polynomial of degree d. The *homogenization* of f is $\bar{f} = z^d f(x/z, y/z)$. Similarly, a homogeneous polynomial $h \in k[x, y, z]$ has a *dehomogenization* with respect to z given by $h_z = h(x, y, 1)$. These operations are partial inverses; $(\bar{f})_z = f$, and if $z \nmid h$ then $(\bar{h}_z) = h$. From this follows a 1–1 correspondence between irreducible factors of f and (necessarily homogeneous) irreducible factors of $\bar{f}$.

Geometrically, homogenization amounts to embedding the solutions of $f(x, y) = 0$ into the projective plane. For this reason, a nonzero triple $(x : y : z)$ with $\bar{f}(x, y, z) = 0$ is called a *projective zero* of f. (Such triples "name" the same point if they differ by a constant factor.)

Let k be a field, and let $f, g \in k[x, y]$ be polynomials without a common factor. Associated with each point P is a nonnegative integer $\mu(P)$ (the *intersection multiplicity*) that measures the extent to which the curves defined by f and g touch at P. It is positive if and only if P is a common zero of f and g, and greater than 1 if and only if f and g are tangent there, or one of the curves has a singularity. *Bézout's theorem* states that $\sum_P \mu(P) = \deg(f) \cdot \deg(g)$, where the sum is taken over all projective points with coordinates in the algebraic closure of k. This gives an upper bound for the number of intersection points of two curves, as well as a lower bound of 1.

Let A be an integral domain, and K the algebraic closure of its quotient field. A polynomial $f(x, y) \in A[x, y]$ is called *absolutely irreducible* if it is irreducible in $K[x, y]$. If $f \in \mathbb{Z}[x, y]$ is absolutely irreducible, then its reduction mod p is absolutely irreducible over $\mathbb{F}_p$, with finitely many exceptions.

Let $f \in \mathbb{F}_p[x, y]$ be absolutely irreducible, of degree d. *Weil's theorem* states that N_P, the number of projective zeros in $\mathbb{F}_p$ of $\bar{f}$, satisfies $|N_p - (p+1)| \leqslant 2\binom{d-1}{2}\sqrt{p}$. This bound holds for all but finitely many p provided that $f \in \mathbb{Z}[x, y]$ is absolutely irreducible.

Let $f \in k[x, y]$ vanish at the origin. A *branch* of f is an irreducible factor of f in the UFD $k[[x, y]]$, with constant term 0. Geometrically, a branch defines an indecomposable local piece of a curve.

Heine's theorem (1858) states that if a rational-coefficient power series defines an algebraic function, then that function is specified by a polynomial with rational coefficients. Formally, let $k \subset K$ be fields, and let the power series $y(x) \in k[[x]]$. If $y(x)$ satisfies a polynomial of degree d in $K[x, y]$, then it satisfies one of the same degree in $k[x, y]$.

Since the proof may not be readily available I give it here. The hypothesis of the theorem asserts that there is a nontrivial K-linear dependence between finitely many of the power series $x^i y^j$, say $v_1, ..., v_k$. If $v_1, ..., v_k$ were k-linearly independent, then for some n, the polynomials $v_1 \bmod x^n, ..., v_k \bmod x^n$ would be linearly independent, implying that the system $\sum t_i v_i \equiv 0 \pmod{x^n}$ has no nonzero solution t_i in K. This is a contradiction.

Finally, *Taylor's formula*: if $f \in A[t_1, ..., t_m]$,

$$f(t_1, ..., t_m) = \sum_{\alpha} \frac{1}{|\alpha|!} \frac{\partial^\alpha f}{\partial t^\alpha}(0, ..., 0)\, t^\alpha,$$

where the summation is taken over multi-indices $\alpha = (\alpha_1, ..., \alpha_m)$ with $\alpha_i \geqslant 0$ and $|\alpha| = \sum \alpha_i$. In characteristic p this formula holds provided that it is given the following interpretation: any denominator occurring in the right-hand side is to be cancelled with a numerator resulting from differentiation.

3. A Formalization of the Rho Method

Define polynomials $f_i \in \mathbb{Z}[x, y]$, by

$$f_0 = x, \qquad f_{i+1} = f_i^2 + y, \qquad i = 1, 2, ...$$

Then f_i has degree 2^i, and

$$f_i \equiv x^{2^i} \pmod{y},$$
$$f_i \equiv y^{2^{i-1}} + \cdots + y \pmod{x}$$

(the elided terms "..." are not asserted to have any special form). As a consequence of the definition,

$$f_{i+j}(x, y) = f_i(f_j(x, y), y).$$

It can also be shown that f_i is absolutely irreducible, by Eisenstein's criterion.

Let $1 \leqslant i < j$. To study the points for which f_i and f_j take the same value, one must consider the factorization of $f_i - f_j$. Evidently,

$$\begin{aligned} f_j - f_i &= f_{j-1}^2 - f_{i-1}^2 = (f_{j-1} + f_{i-1})(f_{j-1} - f_{i-1}) \\ &= (f_{j-1} + f_{i-1})(f_{j-2} + f_{i-2}) \cdots (f_{j-i} + x)(f_{j-i} - x). \end{aligned}$$

However, this is not a complete factorization; indeed, for all $k \geqslant 1$,

$$f_{l+n} - f_l \mid f_{l+kn} - f_l \tag{1}$$

and

$$f_{l+n}+f_l \mid f_{l+kn}+f_l, \tag{2}$$

as can be seen by induction on k.

I now associate a unique polynomial $\rho_{i,j} \in \mathbb{Z}[x, y]$ with each pair (i, j), $i<j$. It is the unique polynomial satisfying the following two properties:

(a) $\rho_{i,j}$ is a monic (in y) irreducible divisor of $f_j - f_i$.

(b) Let $\omega_{i,j}$ denote a primitive (2^j-2^i)th root of unity. Then $\rho_{i,j}(\omega_{i,j}, 0)=0$.

To show that this is a good definition, I claim that property (b) captures $\rho_{i,j}$ uniquely. Modulo y, $f_j - f_i \equiv x^{2^i}(x^{2^j-2^i}-1)$, so there is some divisor of $f_j - f_i$ that vanishes at $(\omega_{i,j}, 0)$. Moreover, there is at most one, as can be seen from the squarefree factorization

$$x^{2^j-2^i}-1 = \prod_{\mu \mid 2^j-2^i} \Phi_\mu(x).$$

(Φ_μ is the μth cyclotomic polynomial.) Finally, $\omega_{i,j}$ and hence 2^j-2^i is uniquely associated with the pair (i, j) (consider binary notation: 2^j-2^i consists of $(j-i)$ ones followed by i zeros).

Since $\rho_{0,j} \mid f_j - f_0$ and $\rho_{i,j} \mid f_{j-1}+f_{i-1}$ if $i \geqslant 1$, (1) implies that

$$\rho_{0,j} \left| \frac{f_j - f_0}{\prod_{d \mid j,\, d \neq j} \rho_{0,d}} \right.$$

and (2) implies for $i \geqslant 1$,

$$\rho_{i,j} \left| \frac{f_{j-1}+f_{i-1}}{\prod_{d \mid j-i,\, d \neq j-i} \rho_{i,i+d}} \right..$$

Possibly the above divisibility relations are really equalities, but I have not been able to prove this. They do, however, give a good way to compute the ρ_{ij}'s: perform the above division and test the result for irreducibility. For example,

$$\rho_{0,1} = y + x^2 - x$$

$$\rho_{0,2} = y + x^2 + x + 1$$

$$\rho_{1,2} = y + x^2 + x$$

$$\rho_{0,3} = y^3 + (3x^2 + x + 2)\, y^2 + (3x^4 + 2x^3 + 3x^2 + 2x + 1)\, y + (x^6 + x^5 + x^4 + x^3 + x^2 + x + 1)$$

$$\rho_{1,3} = y + x^2 - x + 1$$

$$\rho_{2,3} = y^2 + 2(x^2+1)\, y + (x^2+1)x^2.$$

4. Squarefree Irreducible Factorization

This section contains the main technical result of this paper: to factor $f_k - f_l$ in $\mathbb{C}[x, y]$, it suffices to factor it in $\mathbb{Z}[x, y]$. As a bonus, the proof shows that the factorization is squarefree.

Briefly, the proof goes as follows. First, any factor h of $f_k - f_l$ must have the projective zero $(0:1:0)$. Using x and z as local coordinates near this point, a computation shows that all the branches of $f_k - f_l$ are expressible in the form $z - t(x)$, where t is a power series with rational coefficients. One of these branches must represent h, from which it follows that h itself must have rational coefficients.

I now develop this argument formally. Let $k > l$; because $f_k - f_0 \mid f_{k+1} - f_1$, it may be assumed without loss of generality that $l \geqslant 1$. The homogenizations of the f_i's are

$$\bar{f}_1 = x^2 + yz; \qquad \bar{f}_i = \bar{f}_{i-1}^2 + yz^{2^{i-1}}, \qquad i \geqslant 2, \tag{3}$$

so the homogenization of $f_k - f_l$ is

$$\bar{f}_{l,k} = \bar{f}_k - z^{2^k - 2^l} \bar{f}_l. \tag{4}$$

By induction on i, each $\bar{f}_i$ vanishes at $(0:1:0)$, and so the same must be true of $\bar{f}_{l,k}$.

Now let $g_i = \bar{f}_i(x, 1, z)$ and $g_{l,k} = \bar{f}_{l,k}(x, 1, z)$. Then (3) and (4) imply

$$\begin{gathered} g_1 = x^2 + z, \qquad g_i = g_{i-1}^2 + z^{2^{i-1}}, \qquad i = 2, \ldots, k-1, \\ g_{l,k} = g_{k-1}^2 + z^{2^{k-1}} - z^{2^k - 2^l} g_l. \end{gathered} \tag{5}$$

As is customary when dealing with formal power series, I use $O(x^m)$ to indicate terms divisible by x^m.

Lemma 1. *Let $1 \leqslant l < k$, and define $g_{l,k}$ as above. Then the equation $g_{l,k}(x, z) = 0$ has 2^{k-1} distinct power series solutions $z \in \mathbb{Q}[[x]]$. They are all of the form $z(x) = -x^2 + O(x^3)$.*

Proof. Consider the k polynomial equations relating the $k+1$ variables $x, z, z_1, \ldots, z_{k-1}$:

$$\begin{gathered} F_1 = x^2 + z - z_1 = 0 \\ F_i = z_{i-1}^2 + z^{2^{i-1}} - z_i = 0, \qquad i = 2, \ldots, k-1 \\ F_k = z_{k-1}^2 + z^{2^{k-1}} - z^{2^k - 2^l} z_l = 0. \end{gathered} \tag{6}$$

(the new indeterminate z_i stands for g_i).

By recursion on n, I define polynomials $z^{(n)}, z_1^{(n)}, ..., z_{k-1}^{(n)}$ in $\mathbb{Q}[x]$ for which

$$F_i(z^{(n)}, ..., z_{k-1}^{(n)}) \equiv 0 \pmod{x^{2^{i+1}-1+n}}, \qquad i = 1, ..., k. \tag{7}$$

If $n = 0$, let

$$z^{(0)} = -x^2; \qquad z_i^{(0)} = \pm x^{2^{i+1}-1}, \qquad i = 1, ..., k-1.$$

This certainly satisfies (7).

For the recursive step, let $n \geqslant 0$, and assume that $z^{(n)}, z_1^{(n)}, ..., z_{k-1}^{(n)}$ satisfy (7). I seek constants $\varepsilon_0, ..., \varepsilon_{k-1}$ such that if

$$z^{(n+1)} = z^{(n)} + \varepsilon_0 x^{3+n}; \qquad z_i^{(n+1)} = z_i^{(n)} + \varepsilon_i x^{2^{i+1}+n}, \qquad i = 1, ..., k-1$$

then

$$F_i(z^{(n+1)}, ..., z_{k-1}^{(n+1)}) \equiv 0 \pmod{x^{2i+1+n}}, \qquad i = 1, ..., k. \tag{8}$$

By Taylor's fomula, (8) will be satisfied if

$$\begin{aligned}
&F_1(z^{(n)}, ..., z_{k-1}^{(n)}) + \varepsilon_0 x^{3+n} \frac{\partial F_1}{\partial z} + \varepsilon_1 x^{4+n} \frac{\partial F_1}{\partial z_1} \\
&\qquad + \cdots + \varepsilon_{k-1} x^{2^k+n} \frac{\partial F_1}{\partial z_{k-1}} \equiv 0 \pmod{x^{4+n}} \\
&F_2(z^{(n)}, ..., z_{k-1}^{(n)}) + \varepsilon_0 x^{3+n} \frac{\partial F_2}{\partial z} + \varepsilon_1 x^{4+n} \frac{\partial F_2}{\partial z_1} \\
&\qquad + \cdots + \varepsilon_{k-1} x^{2^k+n} \frac{\partial F_2}{\partial z_{k-1}} \equiv 0 \pmod{x^{8+n}} \qquad (8) \\
&\qquad \cdots \\
&F_k(z^{(n)}, ..., z_{k-1}^{(n)}) + \varepsilon_0 x^{3+n} \frac{\partial F_k}{\partial z} + \varepsilon_1 x^{4+n} \frac{\partial F_k}{\partial z_1} \\
&\qquad + \cdots + \varepsilon_{k-1} x^{2^k+n} \frac{\partial F_k}{\partial z_{k-1}} \equiv 0 \pmod{x^{2^{k+1}+n}}
\end{aligned}$$

(the higher-order terms can be ignored). Because of the powers of x above the diagonal, this system is lower-triangular. The partial derivatives along the diagonal, evaluated at $z^{(n)}, ..., z_{k-1}^{(n)}$, are

$$\frac{\partial F_1}{\partial z} = 1 + O(x), \qquad \frac{\partial F_i}{\partial z_{i-1}} = \pm 2x^{2^i-1} + O(x^{2^i}), \qquad i = 2, ..., k.$$

Calculation of the terms below the diagonal shows that the first congruence is divisible by x^{3+n}, the second by x^{7+n}, ..., and the kth by $x^{2^{k+1}-1+n}$. Therefore (9) will hold if

$$F_1(z^{(n)}, ..., z_{k-1}^{(n)})/x^{3+n} + \varepsilon_0 \equiv 0 \pmod{x}$$
$$F_2(z^{(n)}, ..., z_{k-1}^{(n)})/x^{7+n} + \cdots \pm 2\varepsilon_1 \equiv 0 \pmod{x}$$
$$\cdots$$
$$F_k(z^{(n)}, ..., z_{k-1}^{(n)})/x^{2^{k+1}-1+n} + \cdots + \cdots \pm 2\varepsilon_{k-1} \equiv 0 \pmod{x}$$

which allows the constants $\varepsilon_0, ..., \varepsilon_{k-1}$ to be found by back-substitution.

This construction procedures formal power series $z, z_1, ..., z_{k-1}$ in $\mathbb{Q}[[x]]$ that satisfy (6). Furthermore, Eqs. (6) express $z_1, ..., z_{k-1}$ explicitly in terms of x and z. Hence for each possible choice of the 2^{k-1} signs there is a distinct power series for z in terms of x; the result follows. ∎

LEMMA 2. *In* $\mathbb{C}[x, y]$, $f_k - f_l$ *has a squarefree factorization into absolutely irreducible polynomials with integer coefficients.*

Proof. Let $g_{l,k} \in \mathbb{C}[x, z]$ be defined by (5). The intersection of the curve $g_{l,k} = 0$ with the line $z = 0$ (which has to exist) occurs only at $x = z = 0$, which must have multiplicity 2^k by Bézout's theorem. Lemma 1 gives 2^{k-1} branches of $g_{l,k} = 0$, each tangent to $z = 0$ at the origin. Hence the total multiplicity of these branches is 2^k, and there cannot be any more branches.

Now let h be an irreducible polynomial factor of $g_{l,k}$; h also must vanish at $(0, 0)$, and by Lemma 1 and the above, there is a power series $z(x) \in \mathbb{Q}[[x]]$ for which $h(x, z) = 0$. By Heine's theorem, we can take h to have rational coefficients.

Now, there is a 1–1 correspondence between factors of $f_k - f_l$ and factors of $g_{l,k}$ given by $f \mapsto (\tilde{f})_z$. This correspondence does not change the coefficient field, so that each irreducible factor of $f_k - f_l$ can be taken to be in $\mathbb{Q}[x, y]$ as well.

Finally, $f_k - f_l$ is *primitive* (the greatest common divisor of its coefficients is 1). Thus the factorization of $f_k - f_l$ in $\mathbb{Z}[x, y]$ already gives the factorization in $\mathbb{Q}[x, y]$, and hence (by the above) the factorization in $\mathbb{C}[x, y]$. ∎

5. ESTIMATES FOR THE SUCCESS PROBABILITY

This section contains various estimates of the probability that the rho method is successful. First, I show that the probability that two iterates

collide modulo p is asymptotically $\binom{k}{2}/p$ and develop a similar bound for the probability that p is removed at or before the kth step of the process. Finally, I show how this bound can be made explicit in its dependence on k; this leads to an $\Omega(\log^2 p)/p$ bound for the success probability.

THEOREM 1. *Fix $k \geqslant 1$. Choose x and y ar random subject to $0 \leqslant x, y < p$. Then the probability that for some $i, j < k$, $i \neq j$, $f_i(x, y) \equiv f_j(x, y) \pmod p$ is at least $\binom{k}{2}/p + O(1/p^{3/2})$ as $p \to \infty$.*

Proof. The probability in question is at least the probability that for some i, j, $i < j < k$, $\rho_{i,j}(x, y) \equiv 0$. By inclusion–exclusion, Weil's, and Bézout's theorems, this is at least

$$\sum_{i<j<k} \Pr[\rho_{i,j}(x, y) \equiv 0] - \sum_{\substack{i<j<k \\ i'<j'<k \\ (i,j)\neq(i',j')}} \Pr[\rho_{i,j}(x, y) \equiv \rho_{i',j'}(x, y) \equiv 0]$$

$$\geqslant \sum_{i<j<k} (1/p + O(p^{-3/2})) - O(1/p^2) = \binom{k}{2}\Big/ p + O(p^{-3/2}). \quad \blacksquare$$

This result approximates what one would expect if $f_0, \ldots, f_{k-1}$ were numbers chosen at random. In this latter case the probability that there exists a pair $i < j$ with $f_i = f_j$ is $\binom{k}{2}/p + O(p^{-2})$. However, in the theorem given here, the dependence on k in the error term is much greater.

THEOREM 2. *Fix $k \geqslant 1$. Let n have two prime factors p and q with $p < q$. Then $\gcd(f_{2i+1} - f_i, n) \neq 1, n$ for some $i < k$ with probability at least $\binom{k}{2}/p + O(p^{-3/2})$ as $p \to \infty$.*

Proof. If there is a pair (i, j), $0 \leqslant i < j < k$, such that $f_i \equiv f_j \pmod p$, then for some $t \leqslant j$, $f_{2t+1} \equiv f_t \pmod p$. (Consider the cycle formed by $f_i, \ldots, f_j$.) Therefore the probability of success is at least

$$\Pr[\text{for some } i < j < k,\ f_i \equiv f_j \pmod p \text{ and for no } i < k,\ f_i \equiv f_{2i+1} \pmod q].$$

The two conditions are independent, by the Chinese remainder theorem. Since $f_i - f_{2i+1}$ splits into absolutely irreducible factors in $\mathbb{Z}[x, y]$, Weil's theorem gives a bound on the number of zeroes in $\mathbb{F}_q$ of each factor, so

$$\Pr[\text{for some } i < k,\ f_i \equiv f_{2i+1} \pmod q] = O(1/q).$$

Hence the gcd is nontrivial before step k with probability at least

$$\left[\binom{k}{2}\Big/ p + O(1/p^{3/2})\right][1 - O(1/q)] \geqslant \binom{k}{2}\Big/ p + O(1/p^{3/2}),$$

since $q > p$. $\blacksquare$

THEOREM 3. *Let n have two prime divisors p and q with $p<q$. Let $k(p)=\lfloor \frac{1}{4}\log_2 p \rfloor$. Then there is some $i<k(p)$ such that $\gcd(f_{2i+1}-f_i, n)\neq 1$, n with probability at least $\Omega(\log^2 p)/p$.*

Proof. It is first necessary to check that the theory of Sections 3 and 4 can be extended to characteristic $p\neq 2$. If $i<j<\log_2 p$, then the polynomial $x^{2^j-2^i}-1$ has distinct roots mod p; this allows one to choose a primitive (2^j-2^i)th root of unity in the algebraic closure $\bar{\mathbb{F}}_p$, and hence define $\rho_{i,j}$. Since the cyclotomic polynomial may not be irreducible modulo p, $\rho_{i,j}$ may depend on the particular root of unity chosen, but this does not affect the later results. Lemmas 1 and 2 are still true with $\mathbb{Q}$ and $\mathbb{C}$ replaced by $\mathbb{F}_p$ and $\bar{\mathbb{F}}_p$, because the power series coefficients for the branches of f_j-f_i at $(0:1:0)$ lie in $\mathbb{Z}[\frac{1}{2}]$, and $p\neq 2$.

The polynomial $\rho_{i,j}$ has degree at most 2^{k-1} and a projective zero $(0:1:0)$, so by Weil's theorem, the number of solutions in $\mathbb{F}_p$ to $\rho_{i,j}(x,y)=0$ is at least $p-2^{2k-2}\sqrt{p}\geqslant \frac{3}{4}p$. By Bézout's theorem, the number of solutions to $\rho_{i,j}=\rho_{i',j'}=0$ is at most $2^{2k-2}\leqslant \sqrt{p}/4$. Therefore the probability that there exist $i<j$, less than k, with $f_i\equiv f_j \pmod p$ is at least

$$\binom{k}{2}\frac{3p}{4}-\binom{\binom{k}{2}}{2}\frac{1}{4p^{3/2}}=\Omega(\log^2 p)/p.$$

Now consider $f_{2i+1}-f_i \bmod q$. It splits into absolutely irreducible factors of degrees $d_1, \ldots, d_m$, where $\sum d_j = 2^{2i+1}$ and $m\leqslant 2^{2i+1}$, so it vanishes modulo q with probability at most

$$\frac{1}{q^2}\sum_{j=1}^{m}\left[q+2\sqrt{q}\binom{d_j-1}{2}\right]\leqslant \frac{m}{q}+\frac{(\sum d_j)^2}{q^{3/2}}\leqslant \frac{2^{2i+1}}{q}+\frac{2^{4i+2}}{q^{3/2}}.$$

Summing this over $i=0,\ldots,k-1$, the probability that for some $i<k$, $f_{2i+1}\equiv f_i \pmod q$ is at most $6/\sqrt{q}$, since $k\leqslant \frac{1}{4}\log_2 q$.

The result now follows as in the proof of Theorem 2. ∎

6. AN EXCLUSION PRINCIPLE

The curves $\rho_{i,j}=0$ have a striking tendency not to intersect. For instance, there are 210 resultants $R_y(\rho_{i,j},\rho_{k,l})$ corresponding to indices up to 6; the resultant is ± 1—indicating a lack of intersection points—in 165 cases.

This behavior can be explained using the following geometric intuition. The projective point $(0:1:0)$ is a singularity common to all the curves, and there the intersection has high multiplicity. Most of the intersections

that one might expect by Bézout's theorem are "consumed" by this singularity, leaving relatively few to occur in the affine plane.

In this section I prove results that explain this behavior; Theorems 4 and 5 taken together show that the intersection of $\rho_{i,j}=0$ and $\rho_{k,l}=0$ is empty on a set of indices of positive density (rougly 61%, in fact). Throughout this section, $\rho_{i,j}$ is the polynomial in $\mathbb{Z}[x, y]$ defined in Section 3, but the field over which it is evaluated is arbitrary, except for having a characteristic different from 2. It is easiest to consider $j=l=0$ and $j=l=1$ separately, then the general case.

LEMMA 3. *If (ξ, η) is a common zero of $\rho_{0,i}$ and $\rho_{0,j}$ (with $i \neq j$), then it is also a zero of $f_d - f_0$, where $d = \gcd(i, j)$.*

Proof. The hypothesis implies that $f_0(\xi, \eta) = f_i(\xi, \eta) = f_j(\xi, \eta)$. Let d be the least positive integer for which $f_0(\xi, \eta) = f_d(\xi, \eta)$; then $d \mid i$ and $d \mid j$, so $d \mid \gcd(i, j)$. The result follows. ∎

LEMMA 4. *If i adnd j are relatively prime, greater than 1, then $\rho_{0,i}$ and $\rho_{0,j}$ have no zeroes in common.*

Proof. First I need to compute a resultant. Let $\varepsilon = y - x + x^2$, and consider the polynomials f_i modulo ε^2. Then $f_0 = x$, $f_1 = x + \varepsilon$, and in general

$$f_k \equiv x + \varepsilon\phi_k(x) \pmod{\varepsilon^2},$$

where ϕ_k satisfies the recurrence $\phi_0(x) = 0$, $\phi_{k+1} = 1 + 2x\phi_k(x)$. Since $\varepsilon = f_1 - f_0$,

$$\frac{f_k - f_0}{f_1 - f_0} \equiv \phi_k(x) \pmod{\varepsilon}.$$

It follows that ϕ_k is the resultant with respect to y of $(f_k - f_0)/(f_1 - f_0)$ and $f_1 - f_0$.

Now let (ξ, η) be a common zero of $\rho_{0,i}$ and $\rho_{0,j}$; it must also be a common zero of $(f_i - f_0)/(f_1 - f_0)$ and $(f_j - f_0)/(f_1 - f_0)$. By Lemma 3 this must be a zero of $f_1 - f_0$ as well; hence the two resultants ϕ_i and ϕ_j vanish when $x = \xi$. Assume without loss of generality that $i < j$, and let $i' = j - i$. Then

$$\phi_i(\xi) = 0$$

$$\phi_{i+1}(\xi) = 1 = \phi_1(\xi)$$

$$\cdots$$

$$\phi_{i+i'}(\xi) = \phi_{i'}(\xi) = 0.$$

Now apply the same reasoning to i and i', and so on. It eventually will turn out that $\phi_d(\xi) = 0$, where d is the greatest common divisor of i and j. But $d = 1$ and $\phi_d(x) = 1$, so this is impossible. ∎

LEMMA 5. *Let* $1 < i < j$, *and* $\gcd(i, j) = 1$. *Then if* (ξ, η) *is a common zero of* $f_0 + f_i$ *and* $f_0 + f_j$, *it is also a zero of* $f_0 + f_1$.

Proof. The hypothesis implies that $f_1(\xi, \eta) = f_{i+1}(\xi, \eta) = f_{j+1}(\xi, \eta)$. Let d be the least positive number for which $f_1 = f_{d+1}$. Then $d \mid i$ and $d \mid j$, so that $d = 1$ (since i and j are coprime). It follows that (ξ, η) is a zero of one of $f_0 \pm f_1$. Now, $f_0(\xi, \eta) = f_1(\xi, \eta) \neq 0$ is impossible, for then at (ξ, η), $f_0 + f_i = 2f_0 \neq 0$. Hence if $f_0 - f_1 = 0$ at (ξ, η), $f_0 = f_1 = 0$, so $f_0 + f_1 = 0$ there as well. ∎

LEMMA 6. *If* i *and* j *are relatively prime, greater than* 1, *then* $\rho_{1,i+1}$ *and* $\rho_{1,j+1}$ *have no zeroes in common.*

Proof. A computation similar to that in the proof of Lemma 4 shows that if $\psi_k(x)$ denotes the resultant with respect to y of $(f_0 + f_k)/(f_0 + f_1)$ and $f_0 + f_1$, then ψ_k satisfies the recurrence

$$\psi_0 = 0; \qquad \psi_k = 1 - 2x\psi_{k-1}.$$

In particular, $\psi_1 = 1$ and, as in the proof of Lemma 4, if i and j are coprime integers greater than 1, $(f_i + f_0)/(f_1 + f_0)$ and $(f_j + f_0)/(f_1 + f_0)$ have no zeroes in common. The result follows by noting that $\rho_{0,i} \mid (f_i + f_0)/(f_1 + f_0)$ and similarly for j. ∎

LEMMA 7. *If* i *and* j *are coprime, greater than* 1, *then for all* s, $\rho_{s,s+i}$ *and* $\rho_{s,s+j}$ *have no common zeroes.*

Proof. By Lemmas 4 and 6, it remains to consider $s \geqslant 2$. But if (ξ, η) is a common zero of $(f_s + f_{s+i})/(f_s + f_{s+1})$ and $(f_s + f_{s+j})/(f_s + f_{s+1})$ then $(f_s(\xi, \eta), \eta)$ is a common zero of $(f_0 + f_i)/(f_0 + f_1)$ and $(f_0 + f_j)/(f_0 + f_1)$. This is impossible by Lemma 6. ∎

LEMMA 8. *Let* $1 \leqslant i < j$. *Then* $\rho_{i,j}(x, y) \mid \rho_{i-1,j-1}(x^2 + y, y)$.

Proof. The polynomial $\rho_{i-1,j-1}(x^2 + y, y)$ must divide $f_{j-1}(x^2 + y, y) - f_{j-1}(x^2 + y, y) = f_j - f_i$. Let the factorization of $f_j - f_i$ be $g_1 \cdots g_k$, where $\rho_{i,j} = g_1$. Then $\rho_{i-1,j-1}(x^2 + y, y) = \prod_{i \in I} g_i$ for some index set I. Now, only g_1 has the zero $(\omega_{i,j}, 0)$, so it is enough to show that $\rho_{i-1,j-1}(x^2 + y, y)$ vanishes at this point. But in general, $\Phi_{2\mu}(x) \mid \Phi_\mu(x^2)$; from this the result follows. ∎

LEMMA 9. *Let* $t \leqslant i < j$. *Then* $\rho_{i,j}(\xi, \eta) = 0$ *implies* $\rho_{i-t,j-t}(f_t(\xi, \eta), \eta) = 0$.

Proof. By the previous lemma and induction. ∎

LEMMA 10. *Let* $d \geqslant 2$ *and* $j \geqslant 1$. *Then there are no common solutions to* $f_0 = f_1$ *and* $\rho_{j,j+d} = 0$.

Proof. Assume there were a common solution (ξ, η). Then since $f_0(\xi, \eta) = f_1(\xi, \eta) = \cdots$,

$$f_{j-1}(\xi, \eta) + f_{j-1+d}(\xi, \eta) = 0$$
$$f_{j-1}(\xi, \eta) - f_{j-1+d}(\xi, \eta) = 0,$$

so $f_0(\xi, \eta) = f_1(\xi, \eta) = 0$. This implies that $\xi = \eta = 0$. Now let M denote the polynomial ideal (x, y) and consider the sequence $f_0, f_1, \ldots$ modulo M^2. A computation then shows that

$$\frac{f_{j-1+d} + f_{j-1}}{f_j + f_{j-1}} \equiv 1 \pmod{M^2};$$

this implies that $\rho_{j,j+d}(0, 0) \neq 0$. ∎

THEOREM 4. *Let* d *and* e *be relatively prime integers, each greater than* 1. *Then there are no solutions to*

$$\rho_{i,i+d} = \rho_{j,j+e} = 0.$$

Proof. In light of Lemma 7 it can be assumed that $i < j$. If there were a solution (ξ, η), then by Lemma 9 $(f_i(\xi, \eta), \eta)$ would be a common zero of $\rho_{0,d}$ and $\rho_{j-i,j-i+e}$. But an argument based on period lengths (using the hypothesis on d and e) shows that $(f_i(\xi, \eta), \eta)$ would be a common zero of $f_0 - f_1$ and $\rho_{j-i,j-i+e}$, which contradicts Lemma 10. ∎

THEOREM 5. *The number of distinct pairs* (i, j) *and* (i', j') *with* $0 \leqslant i < j < k$, $0 \leqslant i' < j' < k$ *and* $\gcd(i - j, i' - j') = 1$ *is asymptotic to* $(6/\pi^2)\binom{k}{2}$ *as* $k \to \infty$.

Proof. I will count the ordered pairs of such pairs and show that they are asymptotically $(6/\pi^2)\binom{k}{2}^2$ in number. There are $k - 1$ tuples whose components are identical; subtracting this and dividing by 2 will give the result.

Let m be a positive integer and let $n_k(m) = \#\{(i, j): 0 \leqslant i < j < k \;\&\; i \equiv j \pmod{m}\}$. By considering pairs with $j < k - 1$ and $j = k - 1$ separately,

$$n_k(m) = n_{k-1}(m) + \left\lfloor \frac{k-1}{m} \right\rfloor.$$

Since $n_1(m) = 0$, this gives

$$n_k(m) = \sum_{i=0}^{k-1} \left\lfloor \frac{i}{m} \right\rfloor = \frac{\binom{k}{2}}{m} + O(k),$$

where the constant implied by the "O" symbol is absolute.

Now let pairs (i, j) and (i', j') be drawn independently subject to $0 \leqslant i < j < k$. Let E_m denote the event that $i - j \equiv i' - j' \equiv 0 \pmod{m}$. By the estimate for n_k above,

$$\Pr[E_m] = \left(\frac{1}{m} + O\left(\frac{1}{k}\right)\right)^2 = \frac{1}{m^2} + O\left(\frac{1}{mk}\right) + O\left(\frac{1}{k^2}\right).$$

E_m holds for some $m > 1$ precisely when E_p holds for some prime p. By inclusion–exclusion,

$$\begin{aligned}\Pr[E_p \text{ holds for some } p] &= \sum_{p \leqslant k} \Pr[E_p] - \sum_{pq \leqslant k} \Pr[E_p \,\&\, E_q] + \cdots \\ &= \sum_{p \leqslant k} \Pr[E_p] - \sum_{pq \leqslant k} \Pr[E_{pq}] + \cdots.\end{aligned}$$

After a rearrangement,

$$\Pr[E_p \text{ holds for some } p] = \sum_{p \leqslant k} \frac{1}{p^2} - \sum_{pq \leqslant k} \frac{1}{(pq)^2} + \cdots + O(\varepsilon_1) + O(\varepsilon_2),$$

where

$$\varepsilon_1 = \sum_{p \leqslant k} \frac{1}{kp} + \sum_{pq \leqslant k} \frac{1}{kpq} + \sum_{pqr \leqslant k} \frac{1}{kpqr} + \cdots \leqslant \frac{1}{k} \prod_{p \leqslant k} \left(1 + \frac{1}{p}\right) = O\left(\frac{\log k}{k}\right)$$

(since $\prod_{p \leqslant k}(1 - 1/p) \sim e^{-\gamma}/\log k$) and

$$\varepsilon_2 = \sum_{p \leqslant k} \frac{1}{k^2} + \sum_{pq \leqslant k} \frac{1}{k^2} + \cdots = O\left(\frac{1}{k}\right).$$

As $k \to \infty$, both ε_1 and ε_2 tend to zero, so the probability that some E_p holds tends to

$$\sum_p \frac{1}{p^2} - \sum_{pq} \frac{1}{(pq)^2} + \cdots = 1 - \zeta(2)^{-1}$$

(ζ being the Riemann zeta function) and so the probability that *no* E_p holds tends to $\zeta(2)^{-1} = 6/\pi^2$. ∎

7. HEURISTICS

In this section I briefly review some of the heuristic ideas about the rho method and offer one of my own. The new model leads to two interesting open questions.

The crudest assumption to make is that the successive iterates behave like randomly chosen numbers. It is well known in probability theory that one must sample $O(\sqrt{p})$ times from a set of size p to expect a duplicate (see Feller, 1968), so one might guess that the iterates should collide after this many steps.

However, this does not use the functional nature of the iteration. Pollard (1975) assumed this to be a "random mapping" of the residues modulo p; using this assumption he found an expected value close to $\sqrt{p}$ for the least i for which $f_{2i+1} = f_i$. Similar results were conjectured by Brent (1980), Brent and Pollard (1981), and Gold and Sattler (1983) for a more sophisticated version of the algorithm; for this method, results on random mappings by Arney and Bender (1975) and Broder (1981) are relevant.

Guy (1976) has conjectured that the maximum number of iterations needed to detect a prime less than x is $O(x \log x)^{1/2}$. Riesel (1982) argues that this should be $O(\sqrt{x/(\log x \log\log x)})$, based on the assumption that each algebraic factor of $f_{2i+1} - f_i$ splits like a randomly chosen integer of the same size. However, unless I misunderstand his argument it contains an error: the expected number of random samples need to collect N different "coupons" is $O(N \log N)$, not $O(N)$ as seems to be claimed.

What might one conjecture using the ideas of this paper? There is in any case a lower bound for the probability of collision given by

$$\sum \Pr[\rho_{i,j} = 0] - \sum \Pr[\rho_{i,j} = \rho_{i',j'} = 0] + \cdots.$$

For a small number of iterations one can use the worst-case bounds given by Weil and Bézout to estimate the first two sums. As the number of iterations grows large, however, one would expect deviations to cancel and the sums to tend to some "average" value.

In effect, the first sum measures the average number of roots in $\mathbb{F}_p$ of the $\rho_{i,j}$'s and the second measures the average number of $\mathbb{F}_p$-intersection points of the curves. One can get intuition about this situation by considering a probabilistic model where each point (x, y) decides independently and with probability $1/p$ whether or not to be a zero of each $\rho_{i,j}$. Since "most" polynomials in $\mathbb{F}_p[x, y]$ are absolutely irreducible (see the paper of Fredman, 1972), this amounts to assuming the $\rho_{i,j}$'s to be random plane curves.

Using this model I find that the expected time until a collision occurs modulo p is roughly $\sqrt{\pi/2 \cdot p}$, which agrees with Pollard's calculation and empirical results. Probabilistically, the expected number of intersection

points per curve pair is 1; the sample of 210 resultants mentioned in Section 6 had an average of 0.814 intersection points.

Apropos of this model, it would be interesting to prove or disprove that for $k = \sqrt{p}$ and $p \to \infty$,

$$\sum_{i<j<k} \Pr[\rho_{i,j}(x, y) \equiv 0 \pmod{p}] \geqslant C_1 + o(1) \tag{10}$$

and

$$\sum_{\substack{i<j<k \\ i'<j'<k \\ (i,j)\neq(i',j')}} \Pr[\rho_{i,j}(x, y) \equiv \rho_{i',j'}(x, y) \equiv 0 \pmod{p}] \leqslant C_2 + o(1). \tag{11}$$

If these estimates were true for positive constants $C_1 > C_2$, then an $O(\sqrt{p})$ bound on the running time of the rho method would follow.

In this context it should be noted that if one could prove something like (10) even for the set of curves $\rho_{i,j}$ for which $j - i$ is prime, then the results of Section 6 give an analog to (11) that could be used to improve the results of this paper substantially.

It would also be of interest to find other problems for which the above "random curve" model is useful.

Acknowledgments

The support of the National Science Foundation is gratefully acknowledged. I also thank the IBM Almaden Research Center and the Computer Science Division of the University of California at Berkeley for sponsoring visits during which I did much of the research leading to this paper. Richard Fateman at Berkeley deserves special thanks for providing access to the computer algebra system VAXIMA, which computed the statistics mentioned in Section 6.

Received March 9, 1988; revised July 28, 1989

References

Arney, J., and Bender, E. A. (1982) Random mappings with constraints on coalescence and number of origins, *Pacific J. Math.* **103**, 269–294.

Brent, R. P. (1980), An improved Monte Carlo factorization algorithm, *BIT* **20**, 176–184.

Brent, R. P., and Polland, J. M. (1981), Factorization of the eighth Fermat number, *Math. Comp.* **36**, 627–630.

Brieskorn, E., and Knörrer, H. (1986), "Plane Algebraic Curves," Birkhäuser, Boston.

Broder, A. (1985), "Weighted random mappings; Properties and application," Stanford University Computer Science Department Report STAN-CS-85-1054.

Feller, W. (1968), "An Introduction to Probability Theory and its Applications," Wiley, New York.

FREDMAN, M. L. (1972), The distribution of absolutely irreducible polynomials in several indeterminates, *Proc. Amer. Math. Soc.* **31**, 387–390.

FRIED, M., AND JARDEN, M. (1986), "Field Arithmetic," Springer, Berlin.

FULTON, W. (1969), "Algebraic Curves," Benjamin, New York.

GOLD, R., AND SATTLER, J. (1983), Modifikationen des Pollard-algorithmus, *Computing* **30**, 77–79.

GUY, R. (1976), How to factor a number, *Congre. Numer.* **16**, 48–89. (Proceedings, Fifth Manitoba Conference on Numerical Matematics.)

HARTSHORNE, R. (1977), "Algebraic Geometry," Springer, New York.

HEINE, E. (1854), Ueber die Entwickelung von Wurzeln algebraischer Gleichungen in Potenzreihen, *J. Reine Angew. Math.* **48**, 267–275.

MONTGOMERY, P. (1987), Speeding the Pollard and elliptic curve methods of factorization, *Math. Comp.* **48**, 243–264.

POLLARD, J. M. (1975), A Monte-Carlo method for factorization, *BIT* **15**, 331–334.

RIESEL, H. (1982), "Prime Numbers and Computer Methods for Factorization," Birkhäuser, Boston.

VAN DER WAERDEN, B. L. (1970), "Algebra," Ungar, New York.

ZARISKI, O., AND SAMUEL, P. (1958), "Commutative Algebra," 2 Vols., Van Nostrand, Princeton, NJ.