Contemporary Mathematics
Volume 143, 1993

Relatively Prime Values Of Polynomials

James Lee Hafner
IBM Research Division
Almaden Research Center, K53/802
650 Harry Road
San Jose, CA 95120-6099

Peter Sarnak
IBM Research Division
Almaden Research Center, K53/802
650 Harry Road
San Jose, CA 95120-6099

Kevin McCurley *
Sandia National Laboratories
Albuquerque, NM 87185

Abstract

We consider the problem of determining the probability (in a precisely defined sense) that randomly chosen values of a multivariate polynomial are relatively prime. Our result applies to any polynomial with integer coefficients. The resulting probability is given by a product over primes and as such is determined precisely from local considerations.

Introduction

It is well known that the probability that two integers are relatively prime is $6/\pi^2$. We show below that the probability that two large integer matrices have

1991 *Mathematics Subject Classification.* Primary 11N32; Secondary 11N64.
This paper is in final form and no version of it will be submitted for publication elsewhere.
*This work was performed under U. S. Department of Energy contract number DE-AC04-76DP00789

relatively prime determinants is about 1/3. Besides the intrinsic interest of this result, there is also the possible application to the running time analysis of certain probabilistic algorithms [3] and [1].

Another interesting example of our result shows that the probability that two random integers are relatively prime can be made very near to 1 by restricting those integers to be generated by a very simple polynomial.

We actually prove a very general result. Let $P(x_1, \dots, x_l)$ be a non-zero polynomial in l variables and with integer coefficients. For $k \geq 1$ and squarefree, set

$$\rho(k) = \rho(k, P) = \#\{x \in (\mathbb{Z}/k\mathbb{Z})^l : P(x) \equiv 0 \pmod{k}\}.$$

It is easy to see that $\rho(k)$ is multiplicative in k and (see below)

$$\rho(k) = O_\epsilon(k^{l-1+\epsilon}). \tag{1}$$

By the probability that $P(x^{(1)}), P(x^{(2)}), \dots, P(x^{(r)})$, ($r \geq 2$ and fixed) be relatively prime we mean the limit (if it exists)

$$\lim_{t\to\infty} \frac{1}{t^{lr}} \sum_{\substack{|x^{(1)}|\leq t, \dots, |x^{(r)}|\leq t \\ \gcd(P(x^{(1)}), \dots, P(x^{(r)}))=1}} 1,$$

where $|x| = \max\{|x_j|\}$, $x = (x_1, \dots, x_l)$. Our result is the following.

Theorem 1 *The above limit exists for every fixed polynomial and is equal to*

$$C_r(P) = \prod_p \left(1 - \frac{\rho^r(p)}{p^{lr}}\right) = \sum_{k=1}^{\infty} \frac{\mu(k)\rho^r(k)}{k^{lr}},$$

where μ is the Möbius function.

Note that this value is determined precisely from the condition that the numbers $P(x^{(1)}), \dots, P(x^{(r)})$ do not all vanish locally.

For the proof, we need a couple of lemmas.

Lemma 1 *For squarefree k we have*

$$\rho(k) = O(k^{l-1} d^B(k))$$

for some constant $B > 0$ depending only on the polynomial P, and where $d(k) = \sum_{d|k} 1$. The implied constant of course depends on P.

Note: Since $d(k) = O_\epsilon(k^\epsilon)$ for all $\epsilon > 0$, (1) follows from this lemma.

Proof. We have already noted that $\rho(k)$ is multiplicative so it suffices to prove

$$\rho(p) \ll p^{l-1},$$

for all primes p sufficiently large (depending on P). This was first observed by Ore [4] in a slightly stronger form (see also [2, Chapter 6]), but we give a simple proof below for completeness.

Write

$$\begin{aligned} P(x_1,\dots,x_l) &= R_0(x_1,\dots,x_{l-1}) + R_1(x_1,\dots,x_{l-1})x_l \\ &\quad + \cdots + R_j(x_1,\dots,x_{l-1})x_l^j. \end{aligned} \tag{2}$$

For p sufficiently large, P does not vanish identically for every choice of x (mod p). Hence, we can assume some R_ν, say R_0, also has this property. For each selection of $x_1,\dots,x_{l-1}$ mod p for which $R_0(x_1,\dots,x_{l-1}) \not\equiv 0 \pmod p$, we have $O(1)$ values of x_l (mod p) which are solutions of $P(x_1,\dots,x_{l-1},x_l) \equiv 0 \pmod p$. Thus

$$\rho(p) \ll p^{l-1} + p \sum_{\substack{R_0(x_1,\dots,x_{l-1})\equiv 0 \ (\mathrm{mod}\ p) \\ x_j \ (\mathrm{mod}\ p)}} 1.$$

A simple induction argument then completes the proof. □

Lemma 2 *Let P be as above and let A be a large positive constant. For square-free k with $t < k \le t^A$ we have*

$$\sum_{\substack{|x|\le t \\ P(x)\equiv 0 \ (\mathrm{mod}\ k)}} 1 = O_\epsilon(t^{l-1/2+\epsilon}), \qquad \textit{as } t \to \infty.$$

The implied constant depends on P and A but not on k.

Proof. We proceed by induction on the number of variables l. For $l = 1$

$$\sum_{\substack{|x|\le t \\ P(x)\equiv 0 \ (\mathrm{mod}\ k)}} 1 \le \rho(k) = O_\epsilon(k^\epsilon) = O_\epsilon(t^\epsilon),$$

by Lemma 1.

For the general case, write $P(x)$ as in (2) and assume that $R_0(x_1,\dots,x_{l-1})$ is not identically zero. Now again for a selection of $x_1,\dots,x_{l-1}$, the number of solutions in x_l, $|x_l| \le t$, of

$$R_0(x_1,\dots,x_{l-1}) + R_1(x_1,\dots,x_{l-1})x_l + \cdots + R_j(x_1,\dots,x_{l-1})x_l^j \equiv 0 \pmod k$$

is easily seen to be

$$\ll_\epsilon \min\{t, k^\epsilon(R_0(x_1,\dots,x_{l-1}),k)\}.$$

Hence

$$\begin{aligned} \sum_{\substack{|x|\le t \\ P(x)\equiv 0 \ (\mathrm{mod}\ k)}} 1 &\ll_\epsilon \sum_{|x_1|\le t,\dots,|x_{l-1}|\le t} \min\{t, k^\epsilon(R_0(x_1,\dots,x_{l-1}),k)\} \\ &\ll_\epsilon t^\epsilon \sum_{\substack{|x_1|\le t,\dots,|x_{l-1}|\le t \\ (R_0(x_1,\dots,x_{l-1}),k)\le\sqrt{t}}} t^{1/2} \end{aligned}$$

$$+\,t^{1+\epsilon}\sum_{\substack{\nu\mid k\\ \nu>\sqrt{t}}}\ \sum_{\substack{|x_1|\le t,\ldots,|x_{l-1}|\le t\\ (R_0(x_1,\ldots,x_{l-1}),k)=\nu}} 1$$

$$\le\ t^{l-1/2+\epsilon}+t^{1+\epsilon}\sum_{\substack{\nu\mid k\\ \nu>\sqrt{t}}}\ \sum_{\substack{|x_1|\le t,\ldots,|x_{l-1}|\le t\\ R_0(x_1,\ldots,x_{l-1})\equiv 0\ (\mathrm{mod}\ \nu)}} 1. \qquad (3)$$

For $\nu \le t$, the last inner sum is

$$O\left(\left(\frac{t}{\nu}\right)^{l-1}\rho_{R_0}(\nu)\right)$$

and so by Lemma 1 this sum is

$$\ll_\epsilon \left(\frac{t}{\nu}\right)^{l-1}\nu^{l-2+\epsilon}\le\frac{t^{l-1+\epsilon}}{\nu}\le t^{l-3/2+\epsilon}. \qquad (4)$$

since $\nu > \sqrt{t}$.

For $\nu > t$, the last inner sum in (3) may be estimated by the inductive hypothesis and gives

$$O_\epsilon(t^{l-1-1/2+\epsilon}). \qquad (5)$$

Hence, combining the two cases (4) and (5), we can bound the last term in (3) by

$$O_\epsilon\left(\sum_{\nu\mid k}t^{l-1/2+\epsilon}\right)=O_\epsilon(t^{l-1/2+\epsilon}d(k))=O_\epsilon(t^{l-1/2+\epsilon}).$$

This proves Lemma 2. □

We now turn to the proof of the Theorem. We take $r = 2$, as the general case is similar. From the defining property of μ we have

$$\sum_{\substack{|x|\le t,\ |y|\le t\\ (P(x),P(y))=1}} 1 \;=\; \sum_{\substack{|x|\le t,\ |y|\le t\\ P(x)^2+P(y)^2\ne 0}}\left(\sum_{k\mid(P(x),P(y))}\mu(k)\right)$$

$$=\sum_{1\le k\le t^A}\mu(k)\sum_{\substack{P(x)\equiv 0\ (\mathrm{mod}\ k)\\ P(y)\equiv 0\ (\mathrm{mod}\ k)\\ P(x)^2+P(y)^2\ne 0}} 1$$

$$=\sum_{1\le k\le t^A}\mu(k)\left(\sum_{\substack{P(x)\equiv 0\ (\mathrm{mod}\ k)\\ P(y)\equiv 0\ (\mathrm{mod}\ k)\\ |x|\le t,\ |y|\le t}} 1-\sum_{\substack{P(x)=0\\ P(y)=0\\ |x|\le t,\ |y|\le t}} 1\right), \qquad (6)$$

where $A > \deg P$ (and t is large).

We split the k sum in (6) as $S_1 + S_2$ where S_1 corresponds to the range $1 \le k \le t$ and S_2 corresponds to the rest.

Now clearly,

$$\begin{aligned} S_1 &= \sum_{1 \le k \le t} \mu(k) \Bigg(\sum_{\substack{P(x)\equiv 0 \ (\mathrm{mod}\ k) \\ |x| \le t}} 1 \Bigg)^2 + O(t^{2l-1}) \\ &= \sum_{1 \le k \le t} \mu(k) \left[\left\{ \left(\frac{t}{k}\right)^l + O\left(\frac{t}{k}\right)^{l-1} \right\} \rho(k) \right]^2 + O(t^{2l-1}). \end{aligned}$$

Using Lemma 1 in what follows, we get

$$\begin{aligned} S_1 &= t^{2l} \sum_{1 \le k \le t} \frac{\mu(k)\rho^2(k)}{k^{2l}} + O(t^{2l-1+\epsilon}) \\ &= t^{2l} C_2(P) + O(t^{2l-1+\epsilon}). \end{aligned} \tag{7}$$

For S_2 we have

$$\begin{aligned} S_2 &= \sum_{t < k \le t^A} \mu(k) \Bigg(\sum_{\substack{P(x)\equiv 0 \ (\mathrm{mod}\ k) \\ |x| \le t \\ P(x) \ne 0}} 1 \Bigg) \Bigg(\sum_{\substack{P(y)\equiv 0 \ (\mathrm{mod}\ k) \\ |y| \le t}} 1 - \sum_{\substack{P(y)=0 \\ |y| \le t}} 1 \Bigg) \\ &\le \sum_{\substack{t \le k \le t^A \\ k \text{ square-free}}} \Bigg(\sum_{\substack{P(x)\equiv 0 \ (\mathrm{mod}\ k) \\ |x| \le t \\ P(x) \ne 0}} 1 \Bigg) \Bigg(\sum_{\substack{P(y)\equiv 0 \ (\mathrm{mod}\ k) \\ |y| \le t}} 1 \Bigg) \end{aligned}$$

By Lemma 2, we have that the last expression in parentheses is $O_\epsilon(t^{l-1/2+\epsilon})$, hence

$$\begin{aligned} S_2 &\ll_\epsilon t^{l-1/2+\epsilon} \sum_{t < k \le t^A} \sum_{\substack{P(x)\equiv 0 \ (\mathrm{mod}\ k) \\ |x| \le t \\ P(x) \ne 0}} 1 \\ &\le t^{l-1/2+\epsilon} \sum_{\substack{|x| \le t \\ P(x) \ne 0}} \sum_{\substack{m,\ k \\ mk = P(x)}} \\ &= t^{l-1/2+\epsilon} \sum_{\substack{|x| \le t \\ P(x) \ne 0}} d(P(x)) \\ &\ll_\epsilon t^{l-1/2+\epsilon} t^{l+\epsilon} \ll_\epsilon t^{2l-1/2+\epsilon}. \end{aligned} \tag{8}$$

The key to this last estimate is the second step where we reverse the order of summation again, putting the sum on x as the outer sum. In this way we can exploit the fact that though the k's can get rather large, the number of such k is small (a divisor function).

The estimates of (7) and (8) prove the Theorem.

Some Examples

In the case that $P(x)$ is the determinant of an $n \times n$ matrix, one can easily compute $\rho(p)$. We find that

$$C_2(\det_n) = \prod_p \left\{ 1 - \left(1 - \prod_{\nu=1}^{n} (1 - p^{-\nu}) \right)^2 \right\}.$$

Thus

$$C_2(\det_1) = \frac{6}{\pi^2} > C_2(\det_2) > \cdots > C_2(\det_n) \to C_2(\infty),$$

where

$$C_2(\infty) = \prod_p \left\{ 1 - \left(1 - \prod_{\nu=1}^{\infty} (1 - p^{-\nu}) \right)^2 \right\}.$$

Vardi [5] has developed a method to compute such products over primes efficiently and finds that

$$C_2(\infty) = 0.353\ldots.$$

Another interesting example hinted at in the introductory paragraphs is the polynomial

$$P(x) = x^2 + x + 41.$$

Then $C_2(P) \geq 0.986$ which is very large for a polynomial of such small degree and with such small coefficients. The polynomial $Q(x) = x^{12} + 4094$ has an exceptionally small value (without being zero) for this probability, namely, $C_2(Q) \leq 0.00552$. It would be an interesting problem to find other examples with small degree and coefficients, perhaps multivariate, where this probability is extremal.

Note Added in the Proof

After this paper was submitted and accepted, we learned of a new (unpublished) technique due to Michael Rosen that may apply to this problem and yield a very simple proof of the main result. It is based on the observation that the rational integers are uniformly distributed in the topological group $\prod_p \mathbb{Z}_p$, the product of the p-adic integers over all finite primes, and on Weyl's criterion for uniform distribution. Since the details of this method and how it might apply to this problem have not all been worked out, we leave this as a simple note at this time.

References

[1] James Lee Hafner and Kevin S. McCurley, Asymptotically fast triangularization of matrices over rings, *Proc. of SODA*, 1990; also IBM Research Report RJ 6921, 7/12/89.

[2] Rudolf Lidl and Harald Niederreiter, *Finite Fields*, Encylopedia of Mathematics and Its Appl., Vol. 20, Addison-Wesley, Reading, Mass., 1983 (now published by Cambridge University Press).

[3] Kevin S. McCurley, "Cryptographic key distribution and computation in class groups," in *Number Theory and Applications* (Proceedings of the NATO Advanced Study Institute on Number Theory and Applications, Banff, 1988), Richard A. Mollin, Ed., Kluwer, Boston, pages 459–479.

[4] Ö. Ore, Über höhere Kongruenzen, Norsk. Mat. Forenings Skrifter Ser. I (1922), no. 7, 15pp.

[5] Ilan Vardi, personal communication.