THE INSOLUBILITY OF CLASSES OF DIOPHANTINE EQUATIONS.*

By N. C. Ankeny and P. Erdös.

Introduction. Consider the non-trivial rational integer solutions in the variables $X_1, X_2, \cdots, X_n$ of the equation

$$(1) \qquad a_1X_1^m + a_2X_2^m + \cdots + a_nX_n^m = 0,$$

where $m, a_1, a_2, \cdots, a_n$ are non-zero rational integers, and $m > 0$. By a non-trivial solution we mean one in which not all $X_j = 0$, $j = 1, 2, \cdots, n$.

Let U be a large positive real number tending to infinity, and let $D(U, a_1, a_2, \cdots, a_n) = D(U)$ be the number of $m \leqq U$ for which (1) has a non-trivial rational solution. Putting a mild but necessary restriction on the coefficients, something may be said about the order of magnitude of $D(U)$.

THEOREM I. *If, for every selection of $e_j = 0$ or ± 1, $(j = 1, 2, \cdots, n)$ except $(e_1, \cdots, e_n) = (0, 0, \cdots, 0)$, we have $a_1e_1 + \cdots + a_ne_n \neq 0$, then $D(U) = o(U)$ as $U \to +\infty$.*

Theorem I could be interpreted as stating that equation (1) is "almost always" unsolvable; or the density of m, for which (1) has a non-trivial solution, is zero.

One very important case that the hypothesis of Theorem I excludes is when $a_1 = a_2 = a_3 = 1$. However, our methods still yield a result of some interest in this case.

THEOREM II. *The density of integers m, for which the equation $X_1^m + X_2^m + X_3^m = 0$ has a rational solution and for which $(X_1X_2X_3, m) = 1$, is zero.*

The restriction $(X_1X_2X_3, m) = 1$ is sometimes referred to as the first case in Fermat's equation.

The result $M(U) = o(U)$ can be strengthened to $M(U) = O(U(\log U)^{-c})$ for some positive constant c. The proof of this stronger inequality requires a good deal more effort and will not be presented in this paper.

The result of Theorem I can be generalized from the rational number

* Received September 14, 1953; revised December 7, 1953.

field to any algebraic number field F. The restriction on a_j, which are now any non-zero algebraic integers contained in F, is that $a_1\epsilon_1 + \cdots + a_n\epsilon_n \neq 0$ where $\epsilon_j = 0$ or any root of unity contained in F. The proof of this generalization will not be given in complete detail, but will be briefly outlined at the end of this paper.

In Section 1 we shall present some introductory Lemmas and in Section 2, the proof of Theorems I and II will be presented.

1. *Notations.* U denotes a large positive variable. $c_1, c_2, \cdots$ denote absolute constants. p, q are rational primes. ζ_g is a primitive g-th root of unity.

LEMMA 1. *Let $a_1, \cdots, a_n$ satisfy (2), $g > 2$, and $4 \nmid g$.* If $(e_1, \cdots, e_n)$ *is any one of the $3^n - 1$ n-tuples referred to in the statement of Theorem* I *and if $h_1, \cdots, h_n$ are any non-negative integers then*

$$\sum_{k=1}^{n} a_k e_k \zeta_g^{h_k} \neq 0. \tag{3}$$

Proof. Suppose first that $g = p$ or $2p$ where p is an odd prime. Since $\zeta_g^p = \pm 1$, the assumption that (3) is false leads to a relation

$$\sum_{j=0}^{p-1} b_j \zeta_g^j = 0, \text{ where } b_j = \sum_{k \in S_j} a_k e_k, \qquad (j = 0, \cdots, p-1),$$

and S_j is a (possibly void) subset of the set of numbers $\{1, \cdots, n\}$. The sets $S_0, \cdots, S_{p-1}$ are non-overlapping and their union is the set $\{1, \cdots, n\}$. Thus, because of (2), there is an i such that $b_i \neq 0$ and, for every $i' \neq i$, $b_{i'} \neq b_i$. On the other hand, ζ_g is a root of either $x^{p-1} + x^{p-2} + \cdots + x + 1$ or $x^{p-1} - x^{p-2} + - \cdots + (-1)^{p-1}$, both of which are irreducible polynomials over the rational field R. It follows that $b_0 = \pm b_1 = \cdots = \pm b_{p-1}$, a contradiction.

To complete the proof of the lemma, let $g = p_1^{d_1} p_2^{d_2} \cdots p_s^{d_s}$ or $2p_1^{d_1} p_2^{d_2} \cdots p_s^{d_s}$ where the p's are distinct odd primes and the d's are positive integers. Assume by induction on $d_1 + \cdots + d_s$ that the lemma holds for $g' = g/p_1$ (> 2). Since $\zeta_g^{p_1} = \zeta_{g'}$ the assumption that (3) is false leads to a relation

$$\sum_{j=0}^{p_1-1} \beta_j \zeta_g^j = 0, \text{ where } \beta_j = \sum_{k \in S_j} a_k e_k \zeta_{g'}^{f_k}, \qquad (j = 0, \cdots, p_1 - 1),$$

the f's being non-negative integers and the sets $S_0, \cdots, S_{p_1-1}$ having a meaning similar to that in the first part of the proof.

By the inductive hypothesis there is an i such that $\beta_i \neq 0$ and, for every

$i' \neq i$, $\beta_{i'} \neq \beta_i$. On the other hand, the irreducible equation satisfied by ζ_g in the field $R(\zeta_{g'})$ is either $x^{p_1-1} + \cdots + x + 1 = 0$, $(d_1 = 1)$, or $x^{p_1} - \zeta_{g'} = 0$, $(d_1 > 1)$. Thus $\beta_0 = \beta_1 = \cdots = \beta_{p_1-1}$, a contradiction.

LEMMA 2. *If $3 \nmid g$ then, for any non-negative integers h_1, h_2, h_3,*

$$\zeta_g^{h_1} + \zeta_g^{h_2} + \zeta_g^{h_3} \neq 0. \tag{4}$$

Proof. Assume there exist h_1, h_2, h_3 such that $\zeta_g^{h_1} + \zeta_g^{h_2} + \zeta_g^{h_3} = 0$. Divide through by $\zeta_g^{h_3}$, yielding

$$\zeta_g^{k_1} + \zeta_g^{k_2} + 1 = 0, \tag{5}$$

for 2 integers k_1, k_2. Taking the imaginary parts of both sides of (5) yield that $\sin(2\pi k_1/g) + \sin(2\pi k_2/g) = 0$. This implies $k_1 \equiv -k$, or $k_1 + g/2 \pmod g$ where only the former is possible if $2 \nmid g$.

Now taking the real part of (5) yields $\cos(2\pi k_1/g) + \cos(2\pi k_2/g) = -1$ or, on substituting $k_2 \equiv -k_1$ or $k_1 + g/2 \pmod g$, yields that

$$2\cos(2\pi k_1/g) = -1, \text{ or } \cos(2\pi k_1/g) + \cos(2\pi(k_1 + 2g/g)) = -1.$$

This last equation is clearly impossible. The former equation implies that $3 \mid g$, which is contrary to our hypothesis.

THEOREM III. *If $a_1, a_2, \cdots, a_n$ satisfy condition (2), then for a given m there exists no non-trivial rational solutions of (1) provided we can find a rational prime p such that*

$$m \text{ divides } p-1, \qquad mr = p-1, \tag{6}$$

$$4 \nmid r \tag{7}$$

$$\phi(r) < \alpha^{-1}\log p, \tag{8}$$

where $\alpha = \log(|a_1| + |a_2| + \cdots + |a_n|)$, and $\phi(r)$ is a Euler ϕ function.

Proof. (cf. [4], H. S. Vandiver). Assume there exists a p which satisfies (3), (4) and (5), and that (1) has a rational solution such that $X_1X_2\cdots X_n \not\equiv 0 \pmod p$. Without loss of generality, assume $(X_1, X_2, \cdots, X_n) = 1$. Then consider (1) in the field $R(\zeta_r)$.

As $p \equiv 1 \pmod r$, the ideal factorization of p is $(p) = P_1P_2\cdots P_s$ in $R(\zeta_r)$, where $s = \phi(r)$, and $N_{R(\zeta_r),R}(P_1) = p$. Hence, the group of m-th power residues of the multiplicative cyclic group of residues $(\operatorname{mod} P_1)$ has $(p-1)/m = r$ elements. One sees that the elements ζ_r^j, $j = 0, 1, \cdots, r-1$ are incongruent $(\operatorname{mod} P_1)$. So ζ_r^j form a subgroup of r elements in a multiplicative subgroup of residues (mod P_1). Hence, these two subgroups must coincide.

As $a_1X_1^m + \cdots + a_nX_n^m = 0$, à fortiori, $a_1X_1^m + \cdots + a_nX_n^m \equiv 0$ (mod P_1) or, by the coinciding of the two subgroups, $a_1\zeta_r^{t_1} + \cdots + a_n\zeta_r^{t_n} \equiv 0$ (mod P_1) for some n-tuple of integers $(t_1, \cdots, t_n)$. Hence, $p = N_{R(\zeta_r),R}(P_1)$ divides $N_{R(\zeta_r),R}(a_1\zeta_r^{t_1} + \cdots + a_n\zeta_r^{t_n})$. But,

$$| N_{R(\zeta_r),R}(a_1\zeta_r^{t_1} + \cdots + a_n\zeta_r^{t_n}) \leqq (| a_1 | + \cdots + | a_n |)^{\phi(r)}.$$

Thus $p \leqq (| a_1 | + \cdots + | a_n |)^{\phi(r)}$, which is a contradiction to our hypothesis unless $a_1\zeta_r^{t_1} + \cdots + a_n\zeta_r^{t_n} = 0$. This case, however, is impossible by Lemma 1.

Hence, we have shown that $X_1X_2 \cdots X_n \equiv 0$ (mod p). Hence, p divides one of the variables, say X_n. However, proceeding in the same way with the truncated equation $a_1X_1^m + a_2X_2^m + \cdots + a_{n-1}X_{n-1}^m$ we will see that p will divide each X_i, $i = 1, 2, \cdots, n$. This is a contradiction to $(X_1, X_2, \cdots, X_n) = 1$. This proves Theorem 1.

COROLLARY. *If $n = 3$, $a_1 = a_2 = a_3 = 1$, m square free, and a prime p exists that satisfies* (6), (7), (8) *in Theorem* III, $3 \nmid r$, *then* (1) *has no non-trivial solution relatively prime to m.*

Proof. Using the proof of Theorem III and Lemma 2, we immediately infer that there exists no solution of $X_1^m + X_2^m + X_3^m \equiv 0$ (mod p) and $X_1X_2X_3 \not\equiv 0$ (mod p). Hence, if there exists a rational solution $X_1^m + X_2^m + X_3^m = 0$, then $p \mid X_1X_2X_3$.

If q denotes any prime factor of m, and $(X_1X_2X_3, m) = 1$ we have, by using Furtwangler's criterion on Fermat's Equation (cf. Landau [2]), that for any $p \mid X_1X_2X_3$, $p^{q-1} \equiv 1$ (mod q^2). As $p \equiv 1$ (mod m), $p \equiv 1$ (mod q). Therefore, $p \equiv 1$ (mod q^2). As m is square free, $p \equiv 1$ (mod m^2); therefore $p - 1 \geqq m^2$.

By hypothesis, $\phi(r) < \log p/\log 3$. Thus $r < (\log p/\log 3)^2$. Now $m^2 \leqq p - 1 = mr < (\log p/\log 3)^2 m$. Hence, $m < (\log p/\log 3)^2$ or $p - 1 < (\log p/\log 3)^4$.

This last inequality is clearly contradictory and this completes the proof of the corollary.

2. To prove Theorems I and II, we shall derive a set of integers m which satisfy Theorem III and such that almost all integers are divisible by at least one element of our set.

Denote by $\lambda(n)$ the least prime diivsor of n.

LEMMA 3. *If γ denotes Euler's constant,*

$$\sum_{n \in J_1(U)} 1 = e^{-\gamma} U (\log\log\log U)^{-1} + O(\log U),$$

where $J_1(U)$ denotes the rational integers lying between U and $2U$ which have all their prime factors $> \log\log U$.

LEMMA 4. *If $d < U^{\frac{1}{2}}$, then*

$$\sum_{n \in J_2(U)} 1 = e^{-\gamma} \phi(d, \log\log U) U (\log\log\log U)^{-1} + O(\log U)$$

where $\phi(d, V) = d \prod_{\substack{p \mid d \\ p < V}} (1 - 1/p)$, and $J_2(U)$ is the set of integers n between U and $2U$, and $n \equiv 1 \pmod d$.

LEMMA 5. *For any constant c_1, and U sufficiently large,*

$$\sum_{\log U < r < 2\log U} \sum_{J_3(U,r)} 1 > c_2 U (\log\log\log U)^{-1},$$

where $J_3(U, r)$ denotes the set of primes $p < c_1 U \log U$, $p \equiv 1 \pmod r$, and $\lambda((p-1)/r) > \log\log U$. The constant c_2 depends only upon the choice of c_1.

Lemmas 3, 4, and 5 are quite elementary in nature. The proofs of them are very similar. We shall give here only a proof of Lemma 3.

Proof of Lemma 3. Let d be any square free number $< \log\log U$, and let $f(d, U)$ denote the number of integers which lie between U and $2U$ and which are divisible by d. Then $f(d, U) = U/d + O(1)$. If $\mu(d)$ denotes the Moebius function

$$\sum_{n \in J_1(U)} 1 = \sum_{U < n < 2U} \sum_{d \mid (n, h)} \mu(d)$$

where $h = \prod_{p \leqq \log\log U} p$, as this last inner sum is 1 if n has no prime factors $\leqq \log\log U$, and zero otherwise. Hence,

$$\sum_{n \in J_1(U)} 1 = \sum_{d \mid h} \mu(d) \sum_{\substack{U < n < 2U \\ d \mid n}} 1 = \sum_{d \mid h} \mu(d) f(d, U)$$

$$= \sum_{d \mid h} (\mu(d) U/d + O(1)) = U \prod_{p \mid h} \mu(d)/d + O(\sum_{d \mid h} 1)$$

$$= U \prod_{p \mid h} (1 - 1/p) + O(h) = e^{-\gamma} U (\log\log\log U)^{-1} + O(\log U)$$

by using Merton's Theorem on prime numbers. This proves Lemma 3. The proof of Lemma 4 is almost identical to the above. Lemma 5 is again the same as above using the Siegel-Walficz theorem on primes in arithmetic progressions.

Lemma 4 implies that to the primes $< c_3 U \log U$ there corresponds at least $c_2 U(\log\log\log U)^{-1}$ numbers $m < U$, $m = (p-1)/r$, $\lambda(m) > \log\log U$. However, this correspondence is not necessarily unique, and possibly many primes could correspond to the same m. With this in mind we prove

LEMMA 6. *Let $H(U, g)$ denote the number of integers m, $U < m < 2U$, $\lambda(m) > \log\log U$, and such that there are exactly g distinct primes $p_1, p_2, \cdots, p_g$ where $p_j \equiv 1 \pmod m$, and $p_j < c_1 U \log U$ for $j = 1, 2, \cdots, g$. Then $H(U, g) = O(g^{-2} U \log\log\log U)$.*

Proof. The function g^{-2} in the above Lemma could easily be replaced by a function of g which tends to zero far more rapidly as g increases, but this improvement is not needed for this present paper.

To prove Lemma 6 we shall derive an upper bound on the number of times that $(r_1 m + 1)$, $(mr_2 + 1)$, $(mr_3 + 1)$ can simultaneously be primes where $\lambda(m) > \log\log U$, $U \leq m \leq 2U$, $1 \leq r_1, r_2, r_3 \leq \log U$.

Now for the moment regard r_1, r_2, r_3 as fixed and m varying as we described. Then the problem is to derive an upper bound on the number of elements

$$(9) \qquad (mr_1 + 1)(mr_2 + 1)(mr_3 + 1)$$

which have no prime factors $\leq U^{\frac{1}{2}}$. This corresponds to the slight generalization to the twin prime problem where there the number we sieve is $(m)(m+2)$. In our problem we have a polynomial in m composed of 3 linear factors. Utilizing the general method developed by Selberg (cf. [3], especially pp. 291-292), we can easily prove that there are less than

$$c_4 U (\log\log\log U)^{-1} (\log U)^{-3} \psi(r_1) \psi(r_2) \psi(r_3)$$

element of (9) with r_1, r_2, r_3 fixed which have no prime factors $< U^{\frac{1}{2}}$, where $\psi_2(r) = \prod_{p \mid r} (1 - 1/p)$.

Hence, summing over r_1, r_2, r_3, we have that the number of elements of (9) is less than $c_4 U (\log\log\log U)^{-1} (\log U)^{-3} \left(\sum_{r=1}^{\log U} \psi(r) \right)^3$. Now

$$\psi(r) = \prod_{p \mid r} (1 - 1/p)^{-1}$$
$$= \prod_{p \mid r} (1 + 1/p)(1 - 1/p^2)^{-1} < \prod_{p \mid r} (1 + 1/p) \prod_{p} (1 - 1/p^2)^{-1}$$
$$\leqq \pi^2/6 \prod_{p \mid r} (1 + 1/p) = c_5 \sum_{d \mid r} 1/d.$$

Hence

$$\sum_{r=1}^{\log U} \psi(r) < c_5 \sum_{r=1}^{\log U} \sum_{d \mid r} 1/d = c_5 \sum_{d=1}^{\log U} 1/d \sum_{\substack{d \mid r \\ r < \log U}} 1 \leqq c_5 \sum_{d=1}^{\log U} \log U/d^2 < c_6 \log U.$$

Hence, the number of elements of (9) which have all their prime factors $< U^{\frac{1}{2}}$ is $< c_7 U (\log \log \log U)^{-1}$. However, if $c_m{}^n$ is the binomial coefficient, this number is equal to $\sum_{g=3} c_3{}^g H(U, g)$ where $g > 2$. Therefore

(10) $$\sum_{g=3}^{\infty} c_3{}^g H(U, g) < c_6 U (\log \log U)^{-1}.$$

Lemma 6 follows immediately from (10).

Lemmas (4) and (5) showed that to the primes-corresponded various numbers m. Lemma 6 shows conversely that to each m there cannot correspond too many primes.

Therefore, if we define the set $M(U)$ to be all integers $m < U$, $\lambda(m) > \log \log U$, and m satisfies the conditions of Theorem III, then

(11) $$\sum_{m \in M(U)} 1 > c_7 U (\log \log \log U)^{-1}.$$

3. In this section we shall establish that almost all rational integers are divisible by an integer of the set M where $M = \bigcup_{c_6 < U < \infty} M(U)$.

Let $D(M(U)) = D(U)$ denote the density of integers not divisible by any integer of our set $M(U)$. Let U_1 be some large real number.

LEMMA 7. *There exists a constant $0 < c_8 < 1$ such that $D(U_1) < c_8$.*

Proof. As $D(U_1)$ denotes the density of integers not divisible by any $m \in M(U_1)$, $1 - D(U_1)$ denotes the density of integers which are divisible by some $m \in M(U_1)$. Hence

(12) $$1 - D(U_1) \geqq \sum_{m \in M(U_1)} \delta(m),$$

where $\delta(m)$ denotes the density of integers divisible by m but by no other integer of our set $M(U_1)$ except a divisor of m which might be contained in $M(U_1)$. Now, the density of integers t which have no prime factors between

$\log\log U_1$ and U_1, is well known to be $> \frac{1}{2}(\log\log\log U_1)(\log U_1)^{-1}$. Now the set mt is divisible by m and no other element of our set $M(U_1)$ except possibly a divisor of m, as all numbers of $M(U_1)$ have all of their prime divisors lying between $\log\log U_1$ and U_1, and hence to divide mt implies, by our definition of the integers t, that it would divide m. Therefore,

$$\delta(m) \geqq \tfrac{1}{2}(\log\log\log U_1)(\log U_1)^{-1}m^{-1}. \tag{13}$$

By (12), (13) and (11),

$$1 - D(U_1) \geqq \tfrac{1}{2}(\log\log\log U_1)(\log U_1)^{-1} \sum_{m \in M(U_1)} 1/m > c_9 > 0.$$

Letting $c_8 = 1 - c_9$, we have proved Lemma 7.

If U_2 is another large constant, then by Lemma 7, $D(U_2) < c_8$ also. If $U_2 > \exp\{\exp\{2U_1\}\}$, then the elements of $M(U_1)$ are relatively prime to the elements of $M(U_2)$. As all prime factors of $M(U_2)$ are greater than $\log\log U_2 > 2U_1$, and all prime factors of the elements of $M(U_1)$ are less than U_1.

As $D(M(U))$ denotes the density of integers not divisible by any element in $M(U)$,

$$D(M(U_1) \cup D(M(U_2))) = D(M(U_1)) \cdot D(M(U_2)) < c_8{}^2.$$

Similarly, defining $U_3 = \exp\{\exp\{2U_2\}\}$, $U_4 = \exp\{\exp\{2U_3\}\}, \cdots$, gives that

$$D(\bigcup_{j=1}^{n} M(U_j)) = \prod_{j=1}^{n} D(M(U_j)) < c_8{}^n \to 0$$

as $n \to \infty$. Hence,

$$D(M) \leqq \lim_{n\to\infty} D(\bigcup_{j=1}^{n} M(U_j)) = 0, \tag{14}$$

where $D(M)$ denotes the density of integers not divisible by any element of M.

Conversely (14) may be interpreted as saying that almost all integers are divisible by some element of our set M. If an integer n is divisible by an m, $m \in M$, we see that the equation (1) has no non-trivial solution for m, and hence, no non-trivial solution for n. This completes the proof of Theorem I.

To prove Theorem II we would need to add to our conditions on M that the $(p-1)/m$ be relatively prime to 3, and that m be square free. These additional assumptions could easily be incorporated in Section II, and present no real difficulties.

To establish the generalization of Theorem I to an algebraic number field F we merely sum in Lemma 5 over the rational primes which are norms of prime ideals in F. Theorem III can be bodily carried over by changing the definition of α to $(F:R)\log(|a_1|+\cdots+|a_n|)$ where $(F:R)$ denotes the degree of F over the rational number. The remainder of the proof is almost identical.

THE JOHNS HOPKINS UNIVERSITY
AND
AMERICAN UNIVERSITY.

REFERENCES.

[1] N. C. Ankeny, "The insolubility of sets of diophantine equations in the rational numbers," *Proceedings of the National Academy of Sciences*, vol. 38 (1952), pp. 880-884.

[2] E. Landau, *Vorlesungen über Zahlentheorie*, vol. III, 1950, p. 315.

[3] A. Selberg, "Sieve-method in prime number theory," *Proceedings of the international Congress of Mathematicians*, vol. I (1950), pp. 286-292.

[4] H. S. Vandiver, "On classes of diophantine equations of higher degrees which have no solutions," *Proceedings of the National Academy of Sciences*, vol. 32 (1946), pp. 101-106.