COLLOQUIUM MATHEMATICUM

VOL. 96 2003 NO. 2

PRODUCTS OF FACTORIALS MODULO p

BY

FLORIAN LUCA (Morelia) and PANTELIMON STĂNICĂ (Montgomery, AL)

Abstract. We show that if $p \neq 5$ is a prime, then the numbers

$$\left\{ \frac{1}{p}\binom{p}{m_1,\ldots,m_t} \;\middle|\; t \geq 1,\, m_i \geq 0 \text{ for } i = 1,\ldots,t \text{ and } \sum_{i=1}^{t} m_i = p \right\}$$

cover all the nonzero residue classes modulo p.

1. Introduction. Let p be a fixed odd prime and let s and t be fixed positive integers which depend on p. Consider the following subset of $\mathbb{Z}_p^*$:

$$P_{s,t}(p) = \left\{ x_1!\ldots x_t! \pmod p \;\middle|\; x_i \geq 1 \text{ for } i = 1,\ldots,t \text{ and } \sum_{i=1}^{t} x_i = s \right\}. \tag{1}$$

The problem that we investigate in this note is the following: given p, find sufficient conditions on s and t to ensure that $P_{s,t}(p)$ contains the entire $\mathbb{Z}_p^*$.

Let $\varepsilon > 0$ be any small number. Throughout this paper, we denote by $c_1, c_2, \ldots$ computable positive constants which are either absolute or depend on ε. From the way we have formulated the above problem, we see that its answer is easily decidable if either both s and t are very small or very large with respect to p. For example, if $s < c_1(\log p)^2$ with a suitable constant c_1, then it is clear that $P_{s,t}(p)$, or even the union of all $P_{s,t}(p)$ for all allowable values of t, cannot possibly contain the entire $\mathbb{Z}_p^*$ when p is large. Indeed, the reason is that the cardinality of that union is at most $p(s) = O(\exp(c_2\sqrt{s}))$, and this is much smaller than p when p is large if c_1 is chosen such that $c_1 > c_2^2$. Here, we have denoted by $p(s)$ the number of unrestricted partitions of s, and the constant c_2 can be chosen to be $\pi\sqrt{2/3}$. It is also obvious that $P_{s,t}(p)$ does not generate the entire $\mathbb{Z}_p^*$ (for any s) when $t = 2$. Moreover, there exist infinitely many prime numbers p for which the smallest nonquadratic residue modulo p is at least $c_3 \log p$, and so if one wants to generate the entire $\mathbb{Z}_p^*$ from $P_{s,t}(p)$, then one should allow in (1) partitions of s where $\max(x_i)_{i=1}^t$ is at least $c_3 \log p$. In particular, s and t cannot be too

2000 *Mathematics Subject Classification*: 11A07, 11N69, 11B65.

P. Stănică is also associated to the Institute of Mathematics of Romanian Academy, Bucharest, Romania.

close to each other. Indeed, if p is such a prime and the maximum x_i allowed in (1) is at most $c_3 \log p$, then all the numbers in $P_{s,t}(p)$ will be quadratic residues modulo p, and in particular $P_{s,t}(p)$ cannot contain the entire $\mathbb{Z}_p^*$. On the other hand, when s is very large, for example when $s > p^{5/4+\varepsilon}$, then an immediate argument based on the known upper bounds for the size of the smallest primitive root modulo p shows that the union of $P_{s,t}(p)$ over all the allowable values of t does cover the entire $\mathbb{Z}_p^*$ when p is large. Thus, the question becomes interesting when we search for *small* values of both s and t for which $P_{s,t}(p)$ does cover the entire $\mathbb{Z}_p^*$.

This question was inspired by the paper [9] of the second author. In that paper, the problem investigated was to find the exponent at which a prime p divides some generalized Catalan numbers. However, the question of whether a certain subset of Catalan numbers, namely the numbers of the form

$$\frac{1}{p}\binom{p}{m_1, \ldots, m_t}, \tag{2}$$

covers the entire $\mathbb{Z}_p^*$ was not investigated in [9]. Here, the numbers appearing in (2) are all the nontrivial multinomial coefficients. In our notation, this question reduces to whether

$$\bigcup_{t \geq 2} P_{p,t}(p) \tag{3}$$

is the entire $\mathbb{Z}_p^*$. Allowing also $t = 1$ in (3) we deduce that even $0 \in \mathbb{Z}_p$ belongs to this set, and $s = p$ is the smallest value of s for which this can happen. As a byproduct of our results, we show that the set (3) is indeed the entire $\mathbb{Z}_p^*$ for $p \neq 5$.

Our main results are the following:

THEOREM 1. *Let $\varepsilon > 0$ be arbitrary. There exists a computable positive constant $p_0(\varepsilon)$ such that whenever $p > p_0(\varepsilon)$, then $P_{s,t}(p) = \mathbb{Z}_p^*$ for all t and s such that $t > p^{\varepsilon}$ and $s - t > p^{1/2+\varepsilon}$.*

The above result is certainly very far from best possible. We believe that the exponent $1/2$ can be replaced by a much smaller number, or even maybe that the conclusion remains true when $s - t > p^{2\varepsilon}$. However, we have not been able to prove that.

THEOREM 2. *If $p \neq 5$ is a prime, then the set* (3) *is the entire $\mathbb{Z}_p^*$.*

The trick in proving Theorem 2 is to detect a small value of p_0 such that the assertion of Theorem 2 holds for $p > p_0$, and then to test the claim for all primes p from 2 up to p_0.

Theorem 1 above shows, in particular, that the set (3) (even a very small subset of it) is the entire $\mathbb{Z}_p^*$ when p is large. As an example for Theorem 1, we can easily prove that if 2 is a primitive root modulo p, then $A\cup B$, where

$$A=\left\{2^u\left(\frac{p-1}{2}\right)!\;\middle|\;1\le u\le\frac{p-1}{2}\right\},$$
$$B=\left\{2^v\left(\frac{p-3}{2}\right)!\;\middle|\;0\le v\le\frac{p-3}{2}\right\},$$

covers the entire $\mathbb{Z}_p^*$. We see first that A and B each contain $(p-1)/2$ distinct residues modulo p. The intersection $A\cap B$ is empty when 2 is a primitive root modulo p. We omit the details. What is interesting is that, in general, we can cover easily all the even residues, and the odd residues from the first half of $\mathbb{Z}_p^*$, since

$$\frac{1}{p}\binom{p}{2,2k-1,p-2k-1}\equiv k \pmod p,$$
$$\frac{1}{p}\binom{p}{1,1,2k-1,p-2k-1}\equiv 2k \pmod p,$$

for any $1\le k\le (p-1)/2$.

Related to our work, we recall that the behavior of the sequence $n!$ $(\bmod\, p)$ was recently investigated in [2].

2. The proofs of the theorems. The main idea behind the proofs of both Theorems 1 and 2 is to find a suitable list $x_1,\ldots,x_t$ of many small numbers, each repeated a suitable number of times, such that we can modify (in a sense to be made precise below) a fixed element given by formula (1) for those $x_1,\ldots,x_t$ in sufficiently many ways (without, of course, getting outside $P_{s,t}(p)$) so as to obtain, in the end, all the congruence classes in $\mathbb{Z}_p^*$.

Here is the basic operation by which we can modify a fixed element, call it

$$F:=\prod_{i=1}^{t}x_i!,$$

to obtain, hopefully, new elements in $P_{s,t}(p)$:

(M) *Assume that $i_1<\ldots<i_j$ and $l_1<\ldots<l_j$ are two disjoint subsets of indices in $\{1,\ldots,t\}$. Then*

$$\Big(\prod_{s=1}^{j}(x_{l_s}+1)\Big)\Big(\prod_{s=1}^{j}x_{i_s}\Big)^{-1}\cdot F=x_1!\ldots(x_{l_1}+1)!\ldots(x_{i_1}-1)!\ldots x_t! \tag{4}$$
$$=F'\in P_{s,t}(p).$$

We shall always apply (4) with $x_{l_1}=\ldots=x_{l_j}=1$. With this convention, we may eliminate the initial number F, take inverses in (4), and then reformulate the question as follows:

QUESTION. *Is it true that for suitable integers t and s (satisfying, for example, the hypothesis of Theorem 1) we can find some positive integers $x_1, \dots, x_t$ summing to s such that every nonzero residue class modulo p can be represented by a number of the form*

$$\prod_{r=1}^{j} \frac{x_{i_r}}{2} \tag{5}$$

where $\{i_1, \dots, i_j\} \subset \{1, \dots, t\}$ can be any subset such that there exists another subset of j indices $\{l_1, \dots, l_j\}$ disjoint from $\{i_1, \dots, i_j\}$ for which $x_{l_r} = 1$ for all $r = 1, \dots, j$?

Proof of Theorem 1. All we have to show is that if the parameters s and t satisfy the hypothesis of Theorem 1, then we can construct $x_1, \dots, x_t$ for which the answer to the above question is affirmative. Fix $\varepsilon > 0$ and a positive integer k with $1/k < \varepsilon < 2/k$. From now on, all positive constants $c_1, c_2, \dots$ which will appear will be computable and will depend only on k. We shall show that if p is large enough with respect to k, then we can construct a good sublist of $x_1, \dots, x_t$ in the following manner:

(a) We first take and repeat exactly two times each of the prime numbers up to $p^{1/k}$.

(b) We then adjoin at most $c_1 \log\log p$ even numbers (counted with multiplicities), each smaller than $p^{1/2+1/k}$.

(c) The numbers of the form (5), where the x_i's are from the lists (a) and (b) and the maximum length j of a product in (5) is not more than $2k + 2c_1 \log\log p$, cover the entire $\mathbb{Z}_p^*$.

It is clear that if we can prove the existence of a list satisfying (a)–(c), then we are done. Indeed, we may first adjoin to the sublist resulting from (a) and (b) a number of 1's, about $2k + 2c_1 \log\log p$. The list obtained has no more than

$$c_2 \frac{p^{1/k}}{\log p} + 2k + 4c_1 \log\log p < p^{\varepsilon} - 1 < t - 1 \tag{6}$$

numbers while its sum is at most

$$c_3 \frac{p^{2/k}}{\log p} + 2k + 2c_1 \log\log p + 2c_1 p^{1/2+1/k} \log\log p < p^{1/2+\varepsilon} - 1 < s - t - 1, \tag{7}$$

for large p. At this step, we complete the list with some more 1's until we get a list with precisely $t-1$ numbers, which is possible by (6), and set the last number of the list to be

$$x_t := s - \sum_{i=1}^{t-1} x_i,$$

which is still positive by (7).

To show the existence of a sublist with properties (a)–(c) above, we start with the set

$$A := \{n \mid n < p^{1/k} \text{ and } n \text{ is prime}\}.$$

The numbers from A will form the sublist (a) but, so far, we take each of them exactly once. Let

$$B_1 := \left\{ \frac{n_1}{2} \dots \frac{n_k}{2} \;\middle|\; n_i \in A,\, n_i \neq n_j \text{ for } 1 \leq i \neq j \leq k \right\}.$$

We now show that $b_1 := \# B_1$ is large. Indeed, the set B_1 will certainly contain all the numbers of the form

$$\frac{p_1}{2} \dots \frac{p_k}{2} = 2^{-k} p_1 \dots p_k, \tag{8}$$

where p_i is an arbitrary prime subject to the condition

$$p_i \in \left(\frac{p^{1/k}}{2^i}, \frac{p^{1/k}}{2^{i-1}} \right) \quad \text{for } i = 1, \dots, k. \tag{9}$$

Moreover, notice that the residue classes modulo p of the elements of the form (8), where the primes p_i satisfy (9), are all distinct. Indeed, if two numbers of the form (8) coincided modulo p, then, after cancelling the factor of 2^{-k}, we would get two integers which coincide modulo p. Since they are both smaller than p, they must be, in fact, equal. But the elements (8) are all distinct since their prime divisors p_i satisfy (9).

Applying the Prime Number Theorem to estimate from below the number of primes in each one of the intervals in (9), we get

$$b_1 > c_4 \frac{p}{(\log p)^k} > \frac{p}{(\log p)^{k+1}}, \tag{10}$$

whenever $p > c_5$. We construct recursively a (finite) increasing sequence of subsets B_m for $m \geq 1$ in the following way:

Assume that B_m has been constructed and set $b_m := \# B_m$. Assume that $b_m < p - 1$ (that is, B_m is not the entire $\mathbb{Z}_p^*$ already). We then have the following trichotomy:

(i) If $b_m \geq p/2$, then we set $B_{m+1} := B_m \cdot B_m$ and notice that $B_{m+1} = \mathbb{Z}_p^*$, so we can stop.

(ii) If $b_m < p/2$ and there exists an even number $a < p^{1/2+1/k}$ such that $a/2 \notin B_m \cdot B_m^{-1}$, then we set $a_m := a$, add a to the list of the x_i's (on sublist (b)), and let

$$B_{m+1} := B_m \cup \frac{a_m}{2} \cdot B_m.$$

Notice that

$$b_{m+1} \geq 2 b_m. \tag{11}$$

(iii) If $b_m < p/2$ and all even numbers $a < p^{1/2+1/k}$ have the property that $a/2$ is already in $B_m \cdot B_m^{-1}$, we choose the even number a smaller than

$p^{1/2+1/k}$ for which the number of representations of $a/2$ of the form $x \cdot y^{-1}$ with $x, y \in B_m$ is minimal. We then set $a_m := a$, add a to the list of the x_i's (on sublist (b)), set

$$B_{m+1} := B_m \cup \frac{a_m}{2} \cdot B_m,$$

and notice that

$$b_{m+1} \geq 4b_m/3. \tag{12}$$

In (i)–(iii) above, if U and V are two subsets of $\mathbb{Z}_p^*$, we have denoted by $U \cdot V$ the set of all elements of $\mathbb{Z}_p^*$ of the form $u \cdot v$ with $u \in U$ and $v \in V$, and by U^{-1} the set of all elements u^{-1} for $u \in U$.

We have to justify that (i)–(iii) do indeed hold. Notice that (i) and (ii) are obvious. The only detail we have to prove is that inequality (12) holds in situation (iii). For this, we use the following result due to Sárkőzy (see [7]):

LEMMA 1. *Let p be a prime number, u, v, S, T be integers with $1 \leq u, v \leq p-1$, $1 \leq T \leq p$, and $C_1, \dots, C_u$ and $D_1, \dots, D_v$ be integers with*

$$C_i \not\equiv C_j \pmod{p} \quad \textit{for } 1 \leq i < j \leq u,$$
$$D_i \not\equiv D_j \pmod{p} \quad \textit{for } 1 \leq i < j \leq v.$$

For any integer n, let $f(n)$ denote the number of solutions of

$$C_x \cdot D_y \equiv n \pmod{p}, \qquad 1 \leq x \leq u,\ 1 \leq y \leq v.$$

Then

$$\left| \sum_{n=S+1}^{S+T} f(n) - \frac{uvT}{p} \right| < 2(puv)^{1/2} \log p. \tag{13}$$

We apply Lemma 1 with $u = v = b_m$, $C_1, \dots, C_u$ all the residue classes in B_m, and $D_1, \dots, D_u$ all the residue classes in B_m^{-1}. We also set $S = 0$ and T to be the largest integer smaller than $p^{1/2+1/k}/2$. Clearly, $T > p^{1/2+1/k}/3$. Since we are discussing situation (iii) above, we certainly have $f(n) \geq 1$ for all positive integers $n \leq T$. Let $M := \min\{f(n) \mid 1 \leq n \leq T\}$, and then $a_m := 2c$, where $f(c) = M$. Denote b_m by b. We apply inequality (13) to get

$$M < \frac{b^2}{p} + \frac{2b\sqrt{p}\log p}{T}. \tag{14}$$

We first show that

$$\frac{2b\sqrt{p}\log p}{T} < \frac{b^2}{3p}. \tag{15}$$

Indeed, since $T > p^{1/2+1/k}/3$ and $b = b_m \geq b_1 > p/(\log p)^{k+1}$ (by (10)), it follows that in order for (15) to hold, it suffices that

$$54(\log p)^{k+2} < p^{1/k},$$

which is certainly satisfied when $p > c_6$. Thus, inequalities (14) and (15) show that

$$M < \frac{4b^2}{3p} < \frac{2b}{3}, \tag{16}$$

where the last inequality follows from $b < p/2$. In particular,

$$b_{m+1} = \#(B_m \cup c \cdot B_m) \geq b_m + (b_m - M) \geq 2b - \frac{2b}{3} = \frac{4b}{3}, \tag{17}$$

which proves (12).

The combination of (10), (11) and (12) shows that

$$b_{m+1} > \left(\frac{4}{3}\right)^m b_1 > \left(\frac{4}{3}\right)^m \frac{p}{(\log p)^{k+1}} \tag{18}$$

if $b_m < p/2$. Now notice that

$$\left(\frac{4}{3}\right)^m > \frac{(\log p)^{k+1}}{2}$$

provided that $m > c_7 \log\log p$, where one can take $c_7 := (k+1)/\log(4/3)$, for example, and for such large m inequality (18) shows that $b_{m+1} > p/2$. In particular, situations (ii) or (iii) will not occur for more than $c_7 \log\log p$ steps after which we arrive at a point where we apply situation (i) to construct B_{m+1} and we are done. Clearly, (i)–(iii) and the above arguments prove the existence of a sublist of the x_i's satisfying conditions (a)–(c), which finishes the proof of Theorem 1.

Proof of Theorem 2. We follow the method outlined in the proof of Theorem 1. Thus, it suffices to find a list of positive integers, say $A := \{x_1, \ldots, x_s\}$, with

$$U := \sum_{i=1}^{s} x_i < p,$$

and such that for every $m \in \mathbb{Z}_p^*$ there exists a subset $I \subseteq \{1, \ldots, s\}$ for which

$$m \equiv \prod_{i \in I} x_i! \pmod{p}.$$

It is clear that once we show the existence of such an A, we can formally multiply the right hand side of the above congruence by an appropriate number of 1!'s so that the sum of the x_i for $i \in I$ and the 1's is precisely p.

Step 1. We start with a set A_1 of distinct positive integers such that

$$U_1 := \sum_{x \in A_1} x$$

is not too large, and set

$$B_1 := \left\{ \frac{n_1}{2} \cdot \frac{n_2}{2} \,\middle|\, n_1 < n_2 \text{ in } A_1 \right\} \pmod p.$$

For $m \geq 2$, we construct inductively the sets A_m and B_m by the method explained in the proof of Theorem 1. We set $b_m := \#B_m$, $s_m := b_m/p$, and we choose

$$T := 2\lfloor \lambda\sqrt{p}\log p \rfloor + 1,$$

where $\lambda > 2$ is some parameter, to be specified later, and $\lfloor x \rfloor$ is the largest integer $\leq x$. From the way the sets A_m and B_m are constructed for $m \geq 1$, it follows that as long as $s_m < 1/2$, A_{m+1} is obtained from A_m by adjoining to it just one element a_m of size no larger than T, and then B_{m+1} is taken to be $B_m \cup a_m \cdot B_m \pmod p$. Thus,

$$U_{m+1} := \sum_{x \in A_{m+1}} x \leq T + \sum_{x \in A_m} x = T + U_m \quad \text{for } m \geq 1,$$

and therefore

$$U_{m+1} \leq mT + U_1 \tag{19}$$

for all $m \geq 1$ as long as $s_m < 1/2$. However, by (14) and our choice of T, it follows that when constructing A_{m+1} from A_m, we choose the parameter M in such a way that

$$M < \frac{b_m^2}{p} + \frac{2b_m\sqrt{p}\log p}{T} < b_m\left(s_m + \frac{1}{\lambda}\right),$$

therefore inequality (17) now shows that

$$b_{m+1} \geq 2b_m - M > b_m\left(\left(2 - \frac{1}{\lambda}\right) - s_m\right).$$

Hence,

$$s_{m+1} > (\beta - s_m)s_m, \tag{20}$$

where

$$\beta := \beta(\lambda) := 2 - \frac{1}{\lambda} = \frac{2\lambda - 1}{\lambda}.$$

Of course, the above construction will be repeated only as long as $s_m < 1/2$. If we denote by n the largest positive integer such that $s_n < 1/2$, then $s_{n+1} \geq 1/2$, therefore the last set B_{n+2}, which is the entire $\mathbb{Z}_p^*$, is taken to be $B_{n+1} \cdot B_{n+1} \pmod p$, i.e., A_{n+2} is taken to be the list of all elements of A_{n+1}, but now each is repeated twice. Thus,

$$U_{n+2} \leq 2U_{n+1} \leq 2(nT + U_1).$$

From these arguments it follows that in order to ensure that U_{n+2} is not larger than $p - 1$, it suffices to check that

$$2(nT + U_1) < p. \tag{21}$$

The number U_1 can be easily computed in terms of A_1, therefore all we need in order to check that (21) holds is a good upper bound on n in terms of A_1. We recall that n is the largest positive integer with $s_n < 1/2$, where the sequence $(s_m)_{m\ge 1}$ has initial term $s_1 := b_1/p$ and satisfies the recurrence (20).

STEP 2. We give an upper bound on n. Since $\lambda > 2$, it follows that $\beta > 3/2$, therefore (20) shows that $s_{m+1} > s_m$ as long as $s_m < 1/2$. By (20), we also have

$$s_{k+1} > \beta s_k\left(1-\frac{s_k}{\beta}\right) \qquad \text{for } k = 1,\ldots,n,$$

therefore

$$s_{n+1} > \beta^n s_1 \prod_{k=1}^{n}\left(1-\frac{s_k}{\beta}\right).$$

Since $s_k < 1/2$ for $k = 1,\ldots,n$, it follows that

$$\frac{s_k}{\beta} < \frac{1}{2\beta} = \frac{\lambda}{2(2\lambda-1)}.$$

The inequality

$$1-x > e^{-\mu x} \tag{22}$$

holds for all $x \in (0, \lambda/(2(2\lambda-1)))$ with some value $\mu := \mu(\lambda)$, and the best value of μ is precisely

$$\mu := -\frac{\log(1-x)}{x}\bigg|_{x:=\frac{1}{2\beta}} = \frac{2(2\lambda-1)}{\lambda}\log\left(\frac{4\lambda-2}{3\lambda-2}\right), \tag{23}$$

because the function $x \to -\log(1-x)/x$ is decreasing in the interval $(0, 1/(2\beta)]$. Thus,

$$\begin{aligned} \log s_{n+1} &> n\log\beta + \log s_1 + \sum_{k=1}^{n}\log\left(1-\frac{s_k}{\beta}\right) \\ &> n\log\beta + \log s_1 - \frac{\mu}{\beta}\sum_{k=1}^{n} s_k. \end{aligned} \tag{24}$$

We now find an upper bound on $\sum_{k=1}^{n} s_k$. Notice that since $\lambda > 1/2$, it follows that whenever $s_m < 1/2$, one also has

$$s_{m+1} > (\beta - s_m)s_m > (1+\varrho)s_m,$$

where the best $\varrho := \varrho(\lambda)$ is given by

$$\beta - \frac{1}{2} = 1+\varrho,$$

or, equivalently,

$$\varrho := \beta - \frac{3}{2} = \frac{1}{2} - \frac{1}{\lambda} = \frac{\lambda-2}{2\lambda},$$

and
$$1+\varrho = \frac{3\lambda - 2}{2\lambda}.$$
In particular,
$$s_{n-1} < \frac{1}{1+\varrho}\, s_n,$$
and if k is any positive integer less than n, then
$$s_{n-k} < \left(\frac{1}{1+\varrho}\right)^k s_n.$$
Thus,
$$\sum_{k=1}^{n} s_k < s_n \sum_{k\geq 0}\left(\frac{1}{1+\varrho}\right)^k < \frac{1}{2}\,\frac{\varrho+1}{\varrho} = \frac{3\lambda-2}{2(\lambda-2)}.$$

The above calculations show that
$$\log s_{n+1} > n\log\beta + \log s_1 - \mu\,\frac{(3\lambda-2)\lambda}{2(2\lambda-1)(\lambda-2)} = n\log\beta + \log s_1 - \gamma,$$
where
$$\gamma := \gamma(\lambda) := \mu\,\frac{(3\lambda-2)\lambda}{2(2\lambda-1)(\lambda-2)} = \frac{3\lambda-2}{\lambda-2}\log\left(\frac{4\lambda-2)}{3\lambda-2}\right).$$
Thus, if we choose n such that
$$n\log\beta + \log s_1 - \gamma \geq \log(1/2), \tag{25}$$
then we are sure that $s_{n+1} > 1/2$. Inequality (25) is equivalent to
$$n\log\beta > -\log(2s_1) + \gamma,$$
hence to
$$n > \frac{1}{\log\beta}\,(-\log(2s_1)+\gamma).$$
Therefore, we may write
$$n_0 := 1 + \left\lfloor \frac{1}{\log\beta}\,(-\log(2s_1)+\gamma)\right\rfloor, \tag{26}$$
and conclude that $n \leq n_0$. Thus, inequality (21) will be satisfied provided that
$$n_0 T + U_1 < p/2, \tag{27}$$
where n_0 is given by (26).

Step 3. Here, we show that we can do the above construction for $p > 9\cdot 10^6$. From now on, we write $x := p$ and $y := \sqrt{x/2}$, and we assume that $x > 2\cdot 10^6$. In particular, $y > 10^3$. We choose
$$A_1 := \{q \mid q \text{ is prime and } q \leq y\},$$

and therefore

$$B_1 := \left\{ \frac{q_1}{2} \cdot \frac{q_2}{2} \;\middle|\; q_1 < q_2 \text{ and } q_1, q_2 \in A \right\}.$$

It is clear that the elements of B_1 are in distinct congruence classes in $\mathbb{Z}_p^*$, therefore we may consider B_1 as a subset of $\mathbb{Z}_p^*$ and its cardinality is precisely

$$b_1 := \binom{\pi(y)}{2} = \frac{\pi(y)(\pi(y)-1)}{2},$$

where $\pi(y)$ is the number of primes $\le y$. Thus,

$$\frac{1}{2s_1} = \frac{x}{\pi(y)(\pi(y)-1)}. \tag{28}$$

We next give an upper bound on U_1. We claim that

$$U_1 < \frac{1}{2}\pi(y)(\pi(y)+1)\left(\log\pi(y) + \log\log\pi(y) - 1 + 1.8\,\frac{\log\log\pi(y)}{\log\pi(y)}\right). \tag{29}$$

This follows almost immediately from inequality (v) in Théorème A of [4], which states that

$$p_m < m\left(\log m + \log\log m - 1 + \frac{1.8\log\log m}{\log m}\right) \quad \text{for all } m \ge 13. \tag{30}$$

Here p_m denotes the mth prime number. The function

$$t \mapsto \log t + \log\log t - 1 + 1.8\,\frac{\log\log t}{\log t} \tag{31}$$

is increasing for $t > 13$. Moreover, since $y > 10^3$, it follows that $N := \pi(y) \ge 168$,

$$\begin{aligned} &\log N + \log\log N - 1 + 1.8\,\frac{\log\log N}{\log N} \\ &\qquad \ge \log 168 + \log\log 168 - 1 + 1.8\,\frac{\log\log 168}{\log 168} \approx 6.33 > 6, \end{aligned} \tag{32}$$

and

$$p_m < 6m \quad \text{for } m = 1, \ldots, 13. \tag{33}$$

The combination of (30)–(33) shows that

$$\begin{aligned} U_1 = \sum_{p\le y} p &< \left(\log\pi(y) + \log\log\pi(y) - 1 + 1.8\,\frac{\log\log\pi(y)}{\log\pi(y)}\right)\sum_{k=1}^{N} k \\ &= \frac{1}{2}\,N(N+1)\left(\log\pi(y) + \log\log\pi(y) - 1 + 1.8\,\frac{\log\log\pi(y)}{\log\pi(y)}\right), \end{aligned}$$

which is precisely (29).

Having expressed s_1 in terms of $\pi(y)$ and having found an upper bound for U_1 in terms of $\pi(y)$, we now use the inequalities

$$\frac{t}{\log t - 0.5} < \pi(t) < \frac{t}{\log t - 1.5} \quad \text{for all } t > 67 \tag{34}$$

(see Theorem 2 of [6]). The lower bound of (34) together with (28) and (26) yields an upper bound for n_0 in terms of x; the upper bound of (34) gives an upper bound for U_1 in terms of x. Inserting both into (27), we get an inequality which is satisfied for all $x > 11 \cdot 10^6$ at $\lambda = 3$. We have used Mathematica (1) to check that this inequality is true for all $x > 10.3 \cdot 10^6$ (but it fails at $x = 10.2 \cdot 10^6$). Finally, we have checked, using Mathematica again, that (27) is true at $\lambda = 3$ for any prime $x := p$ in the interval $(9 \cdot 10^6, 11 \cdot 10^6)$. In fact, the largest prime $x := p$ for which (27) does not hold at $\lambda = 3$ is $p = 8269189$.

Step 4. It suffices to check that for all primes $5 < p < 9 \cdot 10^6$, the set

$$\Big\{ \prod_{i=1}^{t} m_i! \ \Big| \ \sum_{i=1}^{t} m_i = p - 1 \Big\} \tag{35}$$

covers the entire $\mathbb{Z}_p^*$. Here is a trick that works for p large enough.

Lemma 2. *Assume that $a > 1$ is a primitive root modulo p, and v and b are positive integers in the interval $(1, p-1)$ such that $b \equiv a^v \pmod{p}$ and*

$$v^2 a < p(v - b). \tag{36}$$

Then the set given by (35) covers $\mathbb{Z}_p^*$.

Proof. Take $w := \lfloor (p-1)/v \rfloor$, $t := (v-1)+w$, $m_i := a$ for $i = 1, \ldots, v-1$, and $m_i := b$ for $i = v, v+1, \ldots, t$. Notice first that

$$\sum_{i=1}^{t} m_i = (v-1)a + wb < va + \frac{p}{v} b < p,$$

where the last inequality follows from (36). Thus, we may complete the t-tuple $(m_1, \ldots, m_t)$ with 1's to get a longer vector summing to $p - 1$. Notice also that for each pair (λ, μ) of nonnegative integers with $\lambda \le v - 1$ and $\mu \le w$ we have

$$(a!)^{v-1}(b!)^w = a^\lambda b^\mu ((a-1)!^\lambda (b-1)!^\mu a!^r b!^s),$$

where $r = v - 1 - \lambda$ and $s = w - \mu$. Thus, it suffices to show that every congruence class in $\mathbb{Z}_p^*$ can be represented in the form $a^\lambda b^\mu$ for some nonnegative λ and μ with $\lambda \le v - 1$ and $\mu \le w$. But clearly, every such class is of the form a^t for some $t \in [1, p-1]$ because a is a primitive root modulo p.

(1) A trademark of Wolfram Research.

We may now apply division with remainder to write

$$t = \mu v + \lambda,$$

where $\lambda \le v - 1$ and $\mu := \lfloor t/v \rfloor$. Thus, $\mu \le w$ and

$$a^t = a^{\mu v + \lambda} = a^{\lambda}(a^v)^{\mu} = a^{\lambda} b^{\mu},$$

and the lemma is proved.

Before proceeding, one may ask whether for every sufficiently large prime p there exist positive integers a, b, and v satisfying the hypothesis of Lemma 2. We have been unable to find an unconditional proof of that, but it can be shown that this is indeed so under the Extended Riemann Hypothesis.

LEMMA 3. *Assuming the Extended Riemann Hypothesis, there exists a constant p_0 so that if $p > p_0$ is a prime then there exist integers $a, b, v \in (1, p-1)$ with a being a primitive root modulo p, $b \equiv a^v \pmod{p}$ and*

$$v^2 a < p(v - b). \tag{37}$$

Proof. The following proof is due to Igor Shparlinski. Let p be a sufficiently large prime and let H, K, M, N be positive numbers smaller than p. Let a be an arbitrary primitive root modulo p. It is then known that the number of numbers $v \in [H, H+K]$ such that $a^v \pmod{p} \in [M+1, M+N]$ is $KN/p + O(p^{1/2} \log^2 p)$, where the implied constant is absolute (see [5]). We take $H := 2p^{3/4} \log^{5/4} p$, $K := 2p^{3/4} \log^{5/4} p$, $M := 1$ and $N := p^{3/4} \log^{5/4} p$. Thus, if a is any primitive root modulo p, then the number of numbers $v \in [2p^{3/4} \log^{5/4} p, 4p^{3/4} \log^{5/4} p]$ for which $a^v \pmod{p} \in [1, p^{3/4} \log^{5/4} p]$ is

$$\frac{KN}{p} + O(p^{1/2} \log^2 p) = p^{1/2} \log^{5/2} p + O(p^{1/2} \log^2 p) > 0$$

for p sufficiently large. Thus, if p is large and a is fixed, then there exists an integer $v \in [2p^{3/4} \log^{5/4} p, 4p^{3/4} \log^{5/4} p]$ so that if $b \equiv a^v \pmod{p}$, then $b \in [1, p^{3/4} \log^{5/4} p]$. This is so for an arbitrary primitive root a modulo p. Under the Extended Riemann Hypothesis, it is known (see [8] and [10]) that the smallest primitive root modulo p, call it $g(p)$, satisfies $g(p) = O(\omega(p-1)^6 \log^2 p)$, where $\omega(p-1)$ is the number of distinct prime divisors of $p-1$. Since $\omega(p-1) = o(\log p)$, it follows that if p is large, then the interval $[1, \log^8 p]$ contains a primitive root modulo p. In fact, for our argument it suffices that $[1, p^{1/4}/\log^2 p]$ contains a primitive root a modulo p. With these choices of $a := g(p)$ and v, we have

$$av^2 \le \frac{p^{1/4}}{\log^2 p} (4p^{3/4} \log^{5/4} p)^2 = 16 p^{7/4} \log^{1/2} p, \tag{38}$$

while

$$p(v - b) \ge \frac{pv}{2} \ge p^{7/4} \log^{5/4} p, \tag{39}$$

and now the combination of (38) and (39) obviously shows that (37) holds with these choices of a and v when p is large.

It could be that Hildebrand's [3] improvements on Burgess's [1] character sum estimates could lead to the conclusion that for large p the inequality $g(p) \le p^{1/4}/\log^2 p$ does indeed hold, and if this were so then our Lemma 3 would be true unconditionally. We have been unable to decide this question.

STEP 5. We now return to the proof of Theorem 2 and explain how we did the computations for the remaining primes $p < 9 \cdot 10^6$. We first showed computationally that for every prime $p \in [7.6 \cdot 10^3, 9 \cdot 10^6]$ there exist integers a, b, and v satisfying the hypothesis of Lemma 2. For this, we took the first 25 odd primes and checked them against being primitive roots modulo p. It is clear that at least one of these primes will be a primitive root modulo p for most p in our range. We collected all those primes which are primitive roots modulo p in a set called $A(p)$. Then we tried to find a value for v. We could have looped over all possible values of v, but this would have resulted in a cycle of length $p-1$ for each p, and the computation would have taken too long. Instead, let v_0 be an initial value of v and set $b \equiv a^{v_0} \pmod p$. If v_0 is good, we are done. If not, we set the next v to be

$$v := v_0 + 1 + \left\lfloor \frac{\log p/b}{\log a} \right\rfloor.$$

In a sense, this is the smallest $v > v_0$ for which there is a chance for $a^v = a^{v_0} a^{v-v_0} = b a^{v-v_0}$ to be small modulo p. We kept on doing this for about $3\sqrt{p}$ times for each $a \in A$. If no good values of a and v were found, then we had the program put p in a list of "bad" primes. The computation was done with $v_0 := \lfloor \log p / \log a \rfloor$, but a different choice of v_0 might have given better results.

Now, $\pi(9 \cdot 10^6) = 602489 < 6.1 \cdot 10^5$. After the first run of the algorithm between the 100th and 610000th prime, we obtained a list of 1799 "bad" primes, the largest being 9112771.

In the second iteration, we increased the range for v to $40\sqrt{p}$ and the range of odd primes which may be primitive roots modulo p to 80, and we sieved the previous list. The list shortened to 27 "bad" primes, the first being 541 and the largest 7591. These primes were handled by a different method: we wrote a Mathematica program which showed that the union of the sets

$$A_p(s) = \left\{ 2^u \left(\frac{p-2s-1}{2} \right)! \;\middle|\; 0 \le u \le \left\lfloor \frac{p+2s+1}{4} \right\rfloor \right\}, \tag{40}$$

where $0 \le s \le (p-3)/2$, covers the entire $\mathbb{Z}_p^*$, for any p in the remaining set of "bad" primes. In fact, the above sets were shown to cover $\mathbb{Z}_p^*$ for all the primes < 1000 as well, except for $p = 5$. We conjecture that the union of (40) for all the possible values of s covers $\mathbb{Z}_p^*$ for any prime $p \neq 5$, but we have no idea of how to attack this question.

Acknowledgements. We thank Professors W. O. Nowell and Y. Song for their help with programming in C++ which double-checked our Mathematica computations, and Professor Igor Shparlinski for various useful suggestions as well as for supplying the proof of Lemma 3. We also thank the anonymous referee for suggestions which improved the quality of this paper and for pointing out some errors in a previous version. The first author was supported in part by grants ECOS-ANUIES M02-M01, SEP-CONACYT 37259-E and SEP-CONACYT 37260-E. The second author was partially supported by a research grant from the School of Sciences at his institution.

REFERENCES

[1] D. A. Burgess, *On character sums and primitive roots*, Proc. London Math. Soc. 12 (1962), 179–192.

[2] C. Cobeli, M. Vâjâitu and A. Zaharescu, *The sequence* $n! \pmod p$, J. Ramanujan Math. Soc. 15 (2000), 135–154.

[3] A. Hildebrand, *A note on Burgess' character sum estimate*, C. R. Math. Rep. Acad. Sci. Canada 8 (1986), 35–37.

[4] J. Massias et G. Robin, *Bornes effectives pour certaines fonctions concernant les nombres premiers*, J. Théor. Nombres Bordeaux 8 (1996), 215–242.

[5] H. L. Montgomery, *Distribution of small powers of a primitive root*, in: Advances in Number Theory (Kingston, ON, 1991), Oxford Univ. Press, 1993, 137–149.

[6] A. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. 6 (1962), 64–94.

[7] A. Sárkőzy, *On the distribution of residues of products of integers*, Acta Math. Hungar. 49 (1987), 397–401.

[8] V. Shoup, *Searching for primitive roots in finite fields*, Math. Comp. 58 (1992), 369–380.

[9] P. Stănică, p^q*-Catalan numbers and squarefree binomial coefficients*, J. Number Theory 100 (2003), 203–216.

[10] Y. Wang, *On the least primitive root of a prime*, Acta Math. Sinica 10 (1961), 1–14.

IMATE, UNAM
Ap. Postal 61-3 (Xangari), CP. 58 089
Morelia, Michoacán, Mexico
E-mail: fluca@matmor.unam.mx

Department of Mathematics
Auburn University Montgomery
Montgomery, AL 36124-4023, U.S.A.
E-mail: pstanica@mail.aum.edu

Received 22 January 2002;
revised 23 January 2003 (4158)