

# Moscow Journal

*of  
Combinatorics  
and  
Number Theory*



**Moscow Journal of Combinatorics and Number Theory.** 2011. Vol. 1. Iss. 2. 128 p.

*The journal was founded in 2010.*

*Published by the Moscow Institute of Physics and Technology  
with the support of Yandex and Microsoft.*

The aim of this journal is to publish original, high-quality research articles from a broad range of interests within combinatorics, number theory and allied areas. One volume of four issues is published annually.

**Website of our journal**

<http://mjcnt.phystech.edu>

**E-mail**

[mjcnt@phystech.edu](mailto:mjcnt@phystech.edu)

**Address of the Editorial Board**

Moscow institute of physics  
and technology (state university)  
Faculty of Innovations  
and High Technology,  
Laboratory Korpus, k. 209,  
9, Institutskii pereulok,  
Dolgoprudny,  
Moscow Region,  
Russia,  
141700

**Адрес редакции**

Московский физико-технический  
институт (государственный университет)  
Факультет инноваций  
и высоких технологий  
Лабораторный корпус, к. 209,  
Институтский переулок, д. 9,  
г. Долгопрудный,  
Московская область,  
Российская Федерация,  
141700

**URSS Publishers**

56, Nakhimovsky Prospekt,  
Moscow,  
Russia,  
117335

**Издательство «УРСС»**

Нахимовский пр-т, 56  
Москва,  
Российская Федерация,  
117335

Журнал зарегистрирован в Федеральной службе по надзору в сфере массовых коммуникаций, связи и охраны культурного наследия 3 сентября 2010 г. Свидетельство ПИ № ФС77-41900.

Формат 70 × 100/16. Печ. л. 8. Зак. № ЖО-01.

Отпечатано в ООО «ЛЕНАНД».

117312, Москва, пр-т Шестидесятилетия Октября, 11А, стр. 11.

**ISSN 2220-5438**

© УРСС, 2011

SCIENTIFIC LITERATURE  
AND TEXTBOOKS

E-mail: [URSS@URSS.ru](mailto:URSS@URSS.ru)  
Our catalogue on the Internet:

<http://URSS.ru>

Phone/fax: +7(499) 724 25 45,  
+34 (625) 37 87 73



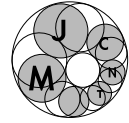
URSS

10838 ID 157682



9 785453 000241

All rights reserved. No part of this book may be used or reproduced in any manner whatsoever without written permission of the publisher.



# On an equation of Goormaghtigh

Florian Luca (Mexico)

**Abstract:** In this paper, we give an upper bound on the number of representations of a positive integer  $N$  as  $N = (q^x - 1)/(q - 1)$  with prime  $q$  and integer  $x$ . We use our result to give an upper bound on the number of  $n \leq N$  such that the sum of divisors function of  $n$  is a multiple of a fixed prime  $p$ .

**Keywords:** Applications of linear forms in logarithms, sieves

**AMS Subject classification:** Primary, 11D41; Secondary, 11D45, 11N25

**Received:** 25.10.2010; **revised:** 14.06.2011

## 1. The Goormaghtigh equation

Given  $N$  let

$$\mathcal{A}(N) = \left\{ (q, x) : N = \frac{q^x - 1}{q - 1}, q \text{ prime, } x \text{ integer} \right\}.$$

Note that  $\mathcal{A}(31) = \{(2, 5), (5, 3)\}$ . In [5], Loxton proved that

$$\#\mathcal{A}(N) \leq (\log N)^{1/2+o(1)}$$

as  $N \rightarrow \infty$ .

In fact, he proved that the above upper bound holds for all positive integer solutions  $(q, x)$  of the equation  $N = (q^x - 1)/(q - 1)$  as  $N \rightarrow \infty$ , without the assumption that  $q$  is prime. Here, we prove the following theorem.

THEOREM 1. *The estimate*

$$\#\mathcal{A}(N) \leq (\log N)^{1/4} \exp \left( (c_0 + o(1)) \frac{\log \log N}{\log \log \log N} \right)$$

holds as  $N \rightarrow \infty$  with  $c_0 = (\log 2)/2$ .

PROOF. Note that since  $N$  is fixed, if  $q$  is such that  $(q, x) \in \mathcal{A}(N)$  for some  $x$ , then  $q$  determines  $x$  uniquely and  $x$  determines  $q$  uniquely. In particular, if  $(q, x) \neq (q', x')$  are both in  $\mathcal{A}(N)$ , then  $q \neq q'$  and  $x \neq x'$ .

Any solution  $(q, r, x, y)$  in integers  $> 1$  of the equation

$$N = \frac{q^x - 1}{q - 1} = \frac{r^y - 1}{r - 1} \tag{1}$$

leads to the inequality

$$\left| y \log r - x \log q + \log \left( \frac{q - 1}{r - 1} \right) \right| < \frac{2}{N}. \tag{2}$$

For the proof, observe that (1) leads to

$$\left| \frac{q^x}{q - 1} - \frac{r^y}{r - 1} \right| \leq \max \left\{ \frac{1}{q - 1}, \frac{1}{r - 1} \right\} \leq 1.$$

Thus,

$$\left| \left( \frac{r^y}{r - 1} \right) \left( \frac{q - 1}{q^x} \right) - 1 \right| \leq \frac{q - 1}{q^x} < \frac{1}{N},$$

which implies (2), via the observation that if  $z$  and  $t$  are real numbers with  $t \in (0, 1/2]$  and  $|e^z - 1| < t$ , then  $|z| < 2t$ .

Assuming that  $(q_i, x_i)$  are in  $\mathcal{A}(N)$  for  $i = 1, \dots, n$  and are ordered in such a way that  $q_1 < q_2 < \dots < q_n$ , we see that inequality (2) with  $(q, r, x, y) = (q_1, q_i, x_1, x_i)$  for some  $i > 1$  is

$$\left| x_i \log q_i - x_1 \log q_1 + \log \left( \frac{q_1 - 1}{q_i - 1} \right) \right| < \frac{2}{N}.$$

For  $i = 2, \dots, n$ , let  $L_i(\mathbf{X})$  be the linear form in  $2n - 1$  variables

$$\mathbf{X} = (X_1, Y_2, \dots, Y_n, Z_2, \dots, Z_n),$$

given by

$$L_i(\mathbf{X}) = Y_i \log q_i - X_1 \log q_1 + Z_i \log \left( \frac{q_1 - 1}{q_i - 1} \right), \quad i = 2, \dots, n.$$

Since  $q_1 < \dots < q_i$  are distinct primes, they are multiplicatively independent. A result of Loxton [5] based on simultaneous linear forms in logarithms of rational numbers shows that there exist absolute constants  $c_1$  and  $c_2$  such that the inequality

$$x < (\log N)^{1/2} \exp \left( c_1 \sqrt{\log \log N \log \log \log N} \right) \tag{3}$$

holds for all solutions  $(q, x) \in \mathcal{A}(N)$  except for a subset  $\mathcal{A}_1(N)$  (which might be empty) of cardinality

$$\#\mathcal{A}_1(N) \leq c_2 \sqrt{\frac{\log \log N}{\log \log \log N}} \tag{4}$$

provided that  $N > e^e$ , which we assume. From now on, we work with pairs  $(q, x) \in \mathcal{A}_2(N) = \mathcal{A}(N) \setminus \mathcal{A}_1(N)$ .

LEMMA 1. *There exists a constant  $c_3$  such that if  $(q_1, x_1), (q_2, x_2)$  are in  $\mathcal{A}_2(N)$  and are such that  $x_1 > x_2$  and  $d = \gcd(x_1 - 1, x_2 - 1) > 1$ , then*

$$x_1 > c_3 (\log N)^{1/2} \tag{5}$$

and

$$d < \exp \left( 4c_1 \sqrt{\log \log N \log \log \log N} \right). \tag{6}$$

PROOF. Theorem 3 in [6] shows that if we write  $x_1 - 1 = dr$ ,  $x_2 - 1 = ds$ , then

$$q_1 \leq 9d^r \prod_{p|d} p^{\text{ord}_p(r!)},$$

where for a positive integer  $m$  and a prime number  $p$  we write  $\text{ord}_p(m)$  for the exact exponent of  $p$  in the factorization of  $m$ . Since

$$\text{ord}_p(r!) = \left\lfloor \frac{r}{p} \right\rfloor + \left\lfloor \frac{r}{p^2} \right\rfloor + \dots < \frac{r}{p} + \frac{r}{p^2} + \dots = \frac{r}{p-1},$$

we get that

$$q_1 \leq 9 \exp \left( r \log d + r \sum_{p|d} \frac{\log p}{p-1} \right).$$

Since certainly

$$\sum_{p|d} \frac{\log p}{p-1} \leq \sum_{i=1}^{\omega(d)} \frac{\log p_i}{p_i-1} \ll \log p_{\omega(d)} \ll \log(\omega(d)+1) \ll \log d,$$

(here,  $p_i$  is the  $i$ th prime number and  $\omega(d)$  is the number of distinct prime factors of  $d$ ), we get that

$$q_1 < \exp(c_4 r \log d),$$

where  $c_4$  is some absolute constant. Thus,

$$N = 1 + q_1 + \dots + q_1^{x_1-1} < 2q_1^{x_1-1} < 2 \exp(c_4 r^2 d \log d),$$

which gives  $c_4 r^2 d \log d > \log(N/2) > (\log N)/2$ . Hence

$$(x_1 - 1)^2 = (dr)^2 > c_3^2 \log N \frac{d}{\log d}, \tag{7}$$

where we put  $c_3 = (2c_4)^{-1/2}$ . Since  $\log d < d$ , we get inequality (5). Since also  $d/\log d > d^{1/2}$  holds for all  $d \geq 2$ , inequalities (3) and (7) lead to inequality (6).  $\square$

LEMMA 2. For each  $d \geq 12$  let

$$\mathcal{S}_d = \{(q, x) \in \mathcal{A}_2(N) : x \equiv 1 \pmod{d}\}.$$

Then

$$\#\mathcal{S}_d < c_5 \sqrt{\log \log N \log \log \log N},$$

where we can take  $c_5 = 2c_1(\log 1,1)^{-1}$  provided that  $N$  is sufficiently large.

PROOF. Let  $(q, r, x, y)$  be a solution of equation (1) with  $(q, x)$  and  $(r, y)$  in  $\mathcal{S}_d$ . Since

$$\gcd(x-1, y-1) \geq d \geq 12 > 4 \cdot 1,1 + 6 + \frac{1}{1,1},$$

Theorem 3 in [3] shows that  $(x-1)/(y-1) > 1,1$  provided that  $N$  is large enough. Assume that  $(q_1, x_1), \dots, (q_t, x_t)$  are all the elements of  $\mathcal{S}_d$  with  $x_1 < \dots < x_t$ .

Lemma 1 shows that  $x_i \geq c_3(\log N)^{1/2}$  for  $i = 2, \dots, t$ . Since also  $x_{i+1} \geq 1, 1x_i$  holds for all  $i = 1, \dots, t - 1$ , we get, by inequalities (3) and (5), that

$$(\log N)^{1/2} \exp(c_1 \sqrt{\log N \log \log N}) \geq x_t \geq x_2(1,1)^{t-2} \geq c_3(\log N)^{1/2}(1,1)^{t-2},$$

and therefore,

$$t \leq 2 + (\log 1,1)^{-1} c_1 \sqrt{\log \log N \log \log \log N} - (\log c_3)(\log 1,1)^{-1},$$

which implies the desired conclusion when  $N$  is sufficiently large. □

Now let  $T = \#\mathcal{A}_2(N)$ . We order the solutions  $(q_i, x_i) \in \mathcal{A}_2(N)$  in such a way that  $q_1 < q_2 < \dots < q_T$ . Whenever  $i > j$ , we put  $d_{i,j} = \ell_{q_j}(q_i)$  defined as the smallest positive integer  $\ell$  such that  $q_i \mid q_j^\ell - 1$ . Since  $q_i > q_j$ , we observe that  $d_{i,j} > 1$ . Furthermore, since  $\ell_{q_j}(q_i) \mid q_i - 1$ , we get that  $d_{i,j} \mid q_i - 1$ . Moreover, the relations

$$N - 1 = \frac{q_i(q_i^{x_i-1} - 1)}{q_i - 1} = \frac{q_j(q_j^{x_j-1} - 1)}{q_j - 1}$$

show that  $d_{i,j} \mid x_j - 1$ . We construct a square matrix  $D = (d_{i,j})_{1 \leq i, j \leq T}$  where we put  $d_{i,j} = 0$  provided that  $q_i \leq q_j$ . Clearly,  $D$  is lower triangular and contains  $\binom{T}{2}$  nonzero entries.

Put  $\mathcal{D} = \{d_{i,j} : 1 \leq i \leq j\} \setminus \{0\}$ . To count the nonzero elements in  $D$ , we proceed in the following way.

We first give an upper bound on the number of repetitions of an integer  $d$  in the same column say corresponding to  $j$ . Assume that  $d$  appears say  $c := c_j(d)$  times in column  $j$ . Then there exist  $c$  distinct primes  $q_{i_1}, \dots, q_{i_c}$  all larger than  $q_j$  such that  $q_{i_s} \mid q_j^d - 1$  for all  $s = 1, \dots, c$ . Hence

$$\prod_{s=1}^c q_{i_s} \mid q^d - 1,$$

showing that

$$q^d > q^d - 1 > \prod_{s=1}^c q_{i_s} > q^c,$$

so  $c < d$ . Thus,

$$c < d. \quad (8)$$

On the other hand, we have  $q_{i_s} \equiv 1 \pmod{d}$  for all  $s = 1, \dots, c$ . Thus,

$$N = q_{i_s}^{x_{i_s}-1} + \dots + 1 \equiv x_{i_s} - 1 \pmod{d}$$

for all  $s = 1, \dots, c$ . Hence we see that all  $x_{i_s}$  are in the same residue class modulo  $d$ . Since the  $x_{i_s}$ 's satisfy inequality (3), we get that

$$\begin{aligned} c &\leq \frac{(\log N)^{1/2}}{d} \exp(c_1 \sqrt{\log \log N \log \log \log N}) + 1 \leq \\ &\leq \frac{2(\log N)^{1/2}}{d} \exp(c_1 \sqrt{\log \log N \log \log \log N}). \end{aligned} \quad (9)$$

Comparing estimates (8) and (9), we get that

$$c_j(d) \leq (\log N)^{1/4} \exp(c_1 \sqrt{\log \log N \log \log \log N}) \quad (10)$$

provided that  $N$  is sufficiently large.

We are now ready to do some counting. Let  $D_1$  be the subset of the entries of  $D$  consisting of the values  $d$  which appear in only one column of the matrix  $D$ . Since whenever  $d \in \mathcal{D}$  appears, say as  $d = d_{i,j}$  in column  $j$ , we have  $d \mid x_j - 1$ , it follows in view of (10) that the number of such entries is

$$\begin{aligned} \#D_1 &\leq (\log N)^{1/4} \exp(c_1 \sqrt{\log \log N \log \log \log N}) \sum_{i=1}^T \tau(x_i - 1) \leq \\ &\leq T(\log N)^{1/4} \exp(c_1 \sqrt{\log \log N \log \log \log N}) \max_{1 \leq i \leq T} \{\tau(x_i - 1)\}, \end{aligned} \quad (11)$$

where for a positive integer  $m$  we write  $\tau(m)$  for the number of its divisors.

Let now  $D_2 = D \setminus D_1$ . If  $d \in \{2, \dots, 11\}$ , then  $d$  can take at most 10 values. Furthermore, by relation (8), each such  $d$  appears at most 10 times in each column. Thus, the number of such entries is  $\leq 100T$ .

Let now  $d \geq 12$  be in  $D_2$ . Lemma 2 shows that the element  $d$  can appear in at most  $c_5 \sqrt{\log \log N \log \log \log N}$  columns, while Lemma 1 shows that the largest

such  $d$  satisfies inequality (6). Thus, in view of (8), the number of entries in  $D_2$  with  $d \geq 12$  is at most

$$\sum_{d < \exp\left(4c_1 \sqrt{\log \log N \log \log \log N}\right)} c_5 d \sqrt{\log \log N \log \log \log N} < \exp\left(9c_1 \sqrt{\log \log N \log \log \log N}\right) \quad (12)$$

provided that  $N$  is sufficiently large. Noting that we have counted all the possible entries in  $D$ , we arrive at the inequality

$$\binom{T}{2} \leq 100T + T(\log N)^{1/4} \exp\left(c_1 \sqrt{\log \log N \log \log \log N}\right) \max_{1 \leq i \leq T} \{\tau(x_i - 1)\} + \exp\left(9c_1 \sqrt{\log \log N \log \log \log N}\right).$$

Since the  $x_i$ 's satisfy inequality (3), and since

$$\tau(m) \leq \exp((\log 2 + o(1)) \log m / \log \log m)$$

holds as  $m \rightarrow \infty$  (see Theorem 317 on page 345 in [4]), the desired inequality now follows.  $\square$

## 2. An application

Let  $\phi(n)$  and  $\sigma(n)$  be the Euler function and the sum of divisors function of positive integer  $n$ . It is known that if  $p$  is a prime and if we write

$$\mathcal{F}(p, N) = \{n \leq N : \phi(n) \equiv 0 \pmod{p}\},$$

then the inequality

$$\#\mathcal{F}(p, N) \ll \frac{N \log \log N}{p}$$

holds. We remind the reader how it is proved. We assume that  $p \leq N$  since otherwise  $\mathcal{F}(p; N)$  is empty. If  $n \in \mathcal{F}(p, N)$ , then either  $p^2 \mid n$ , or  $q \mid n$  for some prime  $q \equiv 1 \pmod{p}$ . Since for an integer  $a$  there are  $\lfloor N/a \rfloor$  positive integers  $n \leq x$  which are multiples of  $a$ , we then have

$$\#\mathcal{F}(p; N) \leq \frac{N}{p^2} + \sum_{\substack{q \leq N \\ q \equiv 1 \pmod{p}}} \frac{N}{q},$$

and the desired estimate for  $\#\mathcal{F}(p, N)$  follows from the known fact that

$$\sum_{\substack{p \leq N \\ p \equiv a \pmod{b}}} \frac{1}{p} \leq \frac{1}{p_{a,b}} + \frac{\log \log N}{\phi(b)} \quad (13)$$

uniformly for coprime integers  $1 \leq a \leq b \leq N$ , where  $p_{a,b}$  stands for the smallest prime congruent to  $a$  modulo  $b$  (note that  $p_{1,b} \geq b + 1$ ).

While in many respects the function  $\sigma(n)$  resembles the function  $\phi(n)$ , it turns out that it is not as easy to give a similar estimate for the cardinality of the set

$$\mathcal{S}(p, N) = \{n \leq N : \sigma(n) \equiv 0 \pmod{p}\}.$$

Here, we give a result showing that the difficulty in estimating  $\#\mathcal{S}(p, N)$  is given by precisely the number  $\#\mathcal{A}(p)$ .

**THEOREM 2.** *The estimate*

$$\#\mathcal{S}(p, N) \leq \frac{N\#\mathcal{A}(p)}{p} + O\left(\frac{N \log \log N}{p}\right)$$

holds uniformly in  $p \geq 2$  and  $N \geq 2$ .

Theorem 2 above together with Theorem 1 improve Theorem 1 of [2].

**PROOF.** We write  $c_0, c_1, \dots$  for computable constants. By the multiplicativity of the sum of divisors function, it follows that whenever  $n \in \mathcal{S}(p, N)$ , there is a prime power  $q^k$  such that  $p \mid \sigma(q^k)$  and  $q^k \parallel n$ . Recall that  $q^k \parallel n$  means that  $q^k \mid n$  and  $q^{k+1} \nmid n$ . Let  $\mathcal{S}_0$  be the set of  $n$  which are divisible by at least one such prime to exactly the first power. Then  $p \mid q + 1$ . Since  $q \mid n$ , it follows that  $q \leq N$ . Let  $q \equiv -1 \pmod{p}$  be fixed. The number of positive integers  $n \leq N$  such that  $q \mid n$  does not exceed  $N/q$ . Hence, summing up over all the possibilities for  $q$ , we get

$$\#\mathcal{S}_0 \leq \sum_{\substack{q \leq N \\ q \equiv -1 \pmod{p}}} \frac{N}{q} = N \sum_{\substack{q \leq N \\ q \equiv -1 \pmod{p}}} \frac{1}{q} \ll \frac{N \log \log N}{p}, \quad (14)$$

where in the above estimates we used again estimate (13) and the fact that the smallest prime  $q \equiv -1 \pmod{p}$  is  $\geq 2p - 1$ .

From now on, we assume that  $n$  is divisible by  $q^k$ , where  $k \geq 2$  and  $p \mid \sigma(q^k)$ . Note that  $2^k \leq q^k \leq N$ , therefore,  $k \leq c_0 \log N$ , where  $c_0 = (\log 2)^{-1}$ . For fixed  $p$  and  $k$ , the congruence

$$\sigma(q^k) = q^k + q^{k-1} + \dots + 1 \equiv 0 \pmod{p}$$

puts  $q$  into  $s_k \leq k$  arithmetical progressions  $q \equiv q_i \pmod{p}$ . Here, we write  $q_i$  for the smallest positive integer in the above progression and we assume that  $1 \leq q_1 < q_2 < \dots < q_{s_k} \leq p-1$ . We also write  $q_{q_i,p}$  for the smallest prime congruent to  $q_i$  modulo  $p$ . Note that

$$2q_i^k \geq q_i^k \left( 1 + \frac{1}{q_i} + \dots \right) \geq p,$$

therefore,  $q_i \gg p^{1/k}$ . Put  $\lambda_{k,i} = (q_i^{k+1} - 1)/(p(q_{q_i,p} - 1))$ . Note that  $\lambda_{k,i} < 2q_{q_i,p}^k/p$ , and that  $1 \leq \lambda_{k,1} < \dots < \lambda_{k,s_k}$ . Fix  $i \in \{1, \dots, k\}$  and write  $\mathcal{S}_{k,i}$  for the subset of  $\mathcal{S}(p, N)$  formed by the integers  $n \leq N$  such that  $q^k \mid n$  for some  $q \equiv q_i \pmod{p}$ . Given such a prime  $q$ , the number of the integers  $n \leq N$  which are multiples of  $q^k$  does not exceed  $N/q^k$ . Summing up the contributions to  $\mathcal{S}(p, N)$  over all such primes  $q$ , we get

$$\#\mathcal{S}_{k,i} \leq \sum_{\substack{q \leq N \\ q \equiv q_i \pmod{p}}} \frac{N}{q^k} \leq \frac{N}{q_{q_i,p}^k} + \sum_{\ell \geq 1} \frac{N}{(q_i + \ell p)^k} \leq \frac{N}{q_i^k} + \frac{N\zeta(2)}{p^k} \ll \frac{N}{p\lambda_{k,i}} + \frac{N}{p^k}. \tag{15}$$

Summing up the above inequality (15) over all the possibilities for  $i \leq s_k \leq k$ , we get that

$$\sum_{i=1}^{s_k} \#\mathcal{S}_{k,i} \ll \frac{N}{p} \sum_{i=1}^{s_k} \frac{1}{\lambda_{k,i}} + \frac{Nk}{p^k}.$$

Finally, summing up the latter inequality over all  $k \geq 2$ , we get

$$\sum_{2 \leq k \leq c_0 \log N} \sum_{i=1}^{s_k} \#\mathcal{S}_{k,i} \ll \frac{N}{p} \sum_{2 \leq k \leq c_0 \log N} \sum_{i=1}^{s_k} \frac{1}{\lambda_{k,i}} + N \sum_{k \geq 2} \frac{k}{p^k}. \tag{16}$$

It is clear that the second sum is  $O(N/p)$ . Thus, it remains to deal with the first double sum. We write

$$\mathcal{S} = \{ (k, q_i) : 2 \leq k \leq c_0 \log N \text{ and } 1 \leq i \leq s_k \}.$$

From now on, to simplify notation, we assume that  $q_i = q_{q_i, p}$  is prime for all  $(k, q_i) \in \mathcal{S}$ . We split  $\mathcal{S}$  into several subsets as follows.

Let  $\mathcal{S}_1$  be the set of pairs  $(k, q_i)$  such that  $\lambda_{k,i} > (\log N)^2$ . Note that

$$\sum_{(k, q_i) \in \mathcal{S}_1} \frac{1}{p\lambda_{k,i}} \leq \frac{1}{p(\log N)^2} \#\mathcal{S}_1 \leq \frac{1}{p(\log N)^2} \sum_{1 \leq k \leq c_0 \log N} s_k \ll \frac{1}{p}. \tag{17}$$

Let  $\mathcal{S}_2$  be the set of pairs  $(k, q_i) \in \mathcal{S} \setminus \mathcal{S}_1$  such that  $q_i^k \leq (\log N)^2$ . In this case, we get that  $k < c_1 \log \log N$  for some appropriate constant  $c_1$ . Thus,

$$\#\mathcal{S}_2 \leq \sum_{2 \leq k \leq c_1 \log \log N} s_k \leq (c_1 \log \log N)^2.$$

Furthermore, if  $\mathcal{S}_2$  is nonempty, then  $p \leq (\log N)^2$ , therefore,  $p\lambda_{k,i} \leq (\log N)^4$  for all  $(k, q_i) \in \mathcal{S}_2$ . Fixing  $\lambda = \lambda_{k,i}$  we see that the equation

$$\frac{q^{k+1} - 1}{q - 1} = p\lambda$$

has at most  $(\log(p\lambda))^{1/4+o(1)} < (\log \log N)^{1/3}$  solutions  $(q, k)$  if  $N$  is sufficiently large by Theorem 1. Since  $\lambda$  can take at most  $(c_1 \log \log N)^2$  distinct values, we get that

$$\begin{aligned} \sum_{(k, q_i) \in \mathcal{S}_2} \frac{1}{p\lambda_{k,i}} &\leq \frac{(\log \log N)^{1/3}}{p} \sum_{1 \leq \lambda \leq (c_1 \log \log N)^2} \frac{1}{\lambda} \ll \\ &\ll \frac{(\log \log N)^{1/3}}{p} \log \log \log N = O\left(\frac{\log \log N}{p}\right). \end{aligned} \tag{18}$$

Let  $\mathcal{S}_3$  be the set of pairs  $(k, q_i) \in \mathcal{S} \setminus (\mathcal{S}_1 \cup \mathcal{S}_2)$  such that  $q_i < (\log N)^2$ . Assume that  $\mathcal{S}_3$  has at least two elements in it, since otherwise we certainly have

$$\sum_{(k, q_i) \in \mathcal{S}_3} \frac{1}{p\lambda_{k,i}} \leq \frac{1}{p}.$$

Let  $(k_0, q_0)$  and  $(k_1, q_1)$  be two pairs of elements in  $\mathcal{S}_3$  with  $k_0$  maximal. Writing

$$p = \frac{q_0^{k_0+1} - 1}{\lambda_0(q_0 - 1)} = \frac{q_1^{k_1+1} - 1}{\lambda_1(q_1 - 1)},$$

we get that

$$(q_1 - 1)\lambda_1 q_0^{k_0+1} - (q_0 - 1)\lambda_0 q_1^{k_1+1} = \delta, \quad (19)$$

where  $\delta = (q_1 - 1)\lambda_1 - (q_0 - 1)\lambda_0$ .

Assume first that  $\delta = 0$ . Since  $q_0$  is coprime to  $q_0 - 1$ , it follows that either  $q_0 = q_1$ , or  $q_0^{k_0+1} \mid \lambda_0$ . The second situation is impossible since  $\lambda_0 \leq (\log N)^2 < q_0^{k_0+1}$ . Thus,  $q_1 = q_0$  and then  $k_1 < k_0$ . But

$$p(\log N)^2 \geq p\lambda_0 = \frac{q_0^{k_0+1} - 1}{q_0 - 1} > q_0^{k_0-k_1} \frac{q_0^{k_1+1} - 1}{q_0 - 1} = q_0^{k_0-k_1} p\lambda_1 \geq q_0^{k_0-k_1} p,$$

giving  $q_0^{k_0-k_1} \leq (\log N)^2$ . Since  $q_0^{k_0+1} \equiv q_0^{k_1+1} \equiv 1 \pmod{p}$ , we get that  $p \mid q_0^{k_0-k_1} - 1$ . Thus,  $p < q_0^{k_0-k_1} \leq (\log N)^2$ , giving  $q_0^{k_0} \leq 2p\lambda_0 \leq 2(\log N)^4$ , whence  $k_0 \leq c_2(\log \log N)$ .

Assume now that  $\delta \neq 0$ . We then rewrite equation (19) as

$$1 - \gamma q_1^{k_1+1} q_0^{-k_0-1} = \frac{\delta_1}{q_0^{k_0+1}},$$

where  $\gamma = \lambda_0 \lambda_1^{-1} (q_0 - 1)(q_1 - 1)^{-1}$  and  $\delta_1 = \delta \lambda_1^{-1} (q_1 - 1)^{-1}$ . Applying a linear form in logarithms à la Baker (see, for example, [1]), we get that

$$k_0 \log q_0 - \log(\delta_1) \ll \log H(\gamma) \log q_1 \log q_0 \log k_0,$$

where  $H(\gamma)$  is the height of the rational number  $\gamma$  (i. e., the maximum between its numerator and denominator). Since certainly  $H(\gamma) \ll (\log N)^4$  and  $q_0, q_1$  do not exceed  $(\log N)^2$ , we get that

$$k_0 \log q_0 \ll (\log \log N)^2 \log q_0 \log k_0 + \log \log N,$$

whence

$$k_0 \ll (\log \log N)^2 \log k_0,$$

leading to  $k_0 \ll (\log \log N)^2 \log \log \log N$ . Thus, in both cases, namely if  $\delta = 0$  or not in equation (19), we get that  $k_0 \leq c_2(\log \log N)^2 \log \log \log N$ . In particular,

$$\#\mathcal{S}_3 \leq \sum_{2 \leq k \leq c_2(\log \log N)^2(\log \log \log N)} s_k \leq c_2^2 (\log \log N)^4 (\log \log \log N)^2.$$

Since  $p < 2q_0^{k_0}$ , we get that  $\log p \ll (\log \log N)^3 (\log \log \log N)$ . Again, for any fixed value of  $\lambda \leq (\log N)^2$ , by Theorem 1, the equation

$$\frac{q^x - 1}{q - 1} = p\lambda$$

has at most  $(\log p + \log \lambda)^{1/4+o(1)} < (\log \log N)^{0.8}$  solutions  $(q, x)$  once  $N$  is sufficiently large. Thus,

$$\begin{aligned} \sum_{(k, q_i) \in \mathcal{S}_3} \frac{1}{p\lambda_{k,i}} &< \frac{(\log \log N)^{0.8}}{p} \sum_{1 \leq \lambda \leq c_2^2 (\log \log N)^3 (\log \log \log N)^2} \frac{1}{\lambda} \ll \\ &\ll \frac{(\log \log N)^{0.8}}{p} \log \log \log N \ll \frac{\log \log N}{p}. \end{aligned} \tag{20}$$

Let  $\mathcal{S}_4$  be the subset of  $\mathcal{S} \setminus (\mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3)$  such that  $k + 1$  is not prime. Then

$$p\lambda = \frac{q^{k+1} - 1}{q - 1} = \prod_{\substack{d|k+1 \\ d>1}} \Phi_d(q),$$

where  $\Phi_d(X) \in \mathbb{Z}[X]$  is the  $d$ th cyclotomic polynomial. Since  $p$  is prime, we get that  $p \mid \Phi_d(q)$  for some divisor  $d$  of  $k + 1$ . In particular,  $p \leq (q + 1)^{\phi(k+1)}$ , therefore,

$$(\log N)^2 \geq \lambda \geq \frac{q^{k+1} - 1}{p(q - 1)} > \frac{q^k}{q^{\phi(k+1)}} \left(1 + \frac{1}{q}\right)^{-\phi(k+1)}.$$

Since certainly  $\phi(k + 1) \leq k \ll \log N$ , but  $q > (\log N)^2$ , we get that

$$\left(1 + \frac{1}{q}\right)^{-\phi(k+1)} = 1 + O\left(\frac{\phi(k+1)}{q}\right) = 1 + O\left(\frac{1}{\log N}\right) > \frac{1}{2}$$

if  $N$  is sufficiently large. Thus,

$$2(\log N)^2 > q^{k-\phi(k+1)} > q^{\sqrt{k+1}-1},$$

where for the latter inequality we used the fact that  $k + 1 - \phi(k + 1) \geq \sqrt{k + 1}$ , since  $k + 1$  is not prime. Taking logarithms, we get

$$\sqrt{k + 1} \log q \ll \log \log N,$$

leading to  $k \log q \ll (\log \log N)^2$ . In particular,

$$k \leq c_3(\log \log N)^2,$$

therefore,

$$\#\mathcal{S}_4 \leq \sum_{2 \leq k \leq c_3(\log \log N)^2} s_k \leq c_3^2(\log \log N)^4,$$

and  $\log p \leq \log 2 + k \log 2 \ll (\log \log N)^2$ . Again, Theorem 1 tells us that for each fixed  $\lambda \leq (\log N)^2$  the equation

$$\frac{q^x - 1}{q - 1} = p\lambda$$

has at most  $(\log p + \log \lambda)^{1/4+o(1)} < (\log \log N)^{3/4}$  solutions  $(q, x)$  once  $N$  is large, therefore,

$$\begin{aligned} \sum_{(k, q_i) \in \mathcal{S}_4} \frac{1}{p\lambda_{k,i}} &< \frac{(\log \log N)^{3/4}}{p} \sum_{1 \leq \lambda \leq c_3^2(\log \log N)^4} \frac{1}{\lambda} \ll \\ &\ll \frac{(\log \log N)^{3/4}}{p} \log \log \log N \ll \frac{\log \log N}{p}. \end{aligned} \quad (21)$$

Let  $\mathcal{S}_5$  be the set of pairs in  $\mathcal{S} \setminus (\mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 \cup \mathcal{S}_4)$  such that  $\lambda \neq 1$ . If  $n \in \mathcal{S}_5$ , then  $k+1$  is a prime. It is known that all prime factors of  $(q^{k+1} - 1)/(q - 1)$  are congruent to 1 modulo  $k+1$ . In particular,  $k+1$  divides  $p-1$ . Let  $r$  be any fixed prime factor of  $p-1$ . Let us count the number of solutions  $(k, q_i)$  with  $k = r-1$ . Then

$$\frac{q_i^r - 1}{q - 1} = p\lambda_{r-1,i}$$

and since  $r$  is now fixed, it follows that  $\lambda_{r-1,i}$  are distinct as  $i = 1, \dots, s_{r-1}$ . All such solutions have the property that  $\lambda_{r-1,i} \equiv 1 \pmod{r}$ . Thus,

$$\sum_{i=1}^{s_{r+1}} \frac{1}{p\lambda_{r-1,i}} \leq \frac{1}{p} \sum_{i=0}^r \frac{1}{1+ir} \ll \frac{\log r}{pr}.$$

We now sum up this last relation over all possible prime factors  $r$  of  $p-1$ . Since  $\omega(p-1) < 2 \log N / \log \log N$  for large  $N$ , it follows by the Prime Number Theorem

that

$$\sum_{(k,q_i) \in \mathcal{S}_5} \frac{1}{p\lambda_{i,k}} \ll \frac{1}{p} \sum_{1 \leq r \leq 3 \log N} \frac{\log r}{r} \ll \frac{\log \log N}{p}. \quad (22)$$

It is clear that if  $\mathcal{S}_6$  denotes the set for remaining pairs  $(k, q_i)$ , then they all have  $\lambda_{k,i} = 1$ . In particular,  $\#\mathcal{S}_6 \leq \#\mathcal{A}(p)$ . Furthermore, for each such solution we have

$$q_i^k \left( 1 + O\left(\frac{1}{q}\right) \right) = p$$

and  $q_i > (\log N)^2$ , therefore,

$$\frac{1}{q^k} = \frac{1}{p} + O\left(\frac{1}{p(\log N)^2}\right).$$

Thus,

$$\sum_{(k,q_i) \in \mathcal{S}_6} \frac{1}{q_i^k} \leq \frac{\#\mathcal{A}(p)}{p} + O\left(\frac{\#\mathcal{A}(p)}{p(\log N)^2}\right) = \frac{\#\mathcal{A}(p)}{p} + O\left(\frac{1}{p}\right). \quad (23)$$

Theorem 2 now follows immediately from estimates (14), (15), (16), (17), (18), (20), (21), (22) and (23).  $\square$

**Remark.** If  $(q^x - 1)/(q - 1) = p$ , then  $q \mid p - 1$ . Thus,  $\#\mathcal{A}(p) \leq \omega(p - 1)$ . For most primes  $p$ ,  $\omega(p - 1) < 2 \log \log p$ . Hence we see by Theorem 2 that the estimate

$$\mathcal{S}(p, N) \ll \frac{\log \log N}{p}$$

holds for most primes  $p$ . Of course, since the believed conjectures are that  $\#\mathcal{A}(p) = O(1)$ , the above inequality should be true for all primes  $p$  in light of Theorem 2.

**Acknowledgements.** I thank Dr. Shanta Laishram and Professors T. N. Shorey and Igor Shparlinski for useful discussions. I also thank the referee and the editor for corrections and additions to the original manuscript. Work on this paper started during Summer of 2007 when I visited the Tata Institute for Fundamental Research in Mumbai, India with a TWAS Associateship, and ended during the academic year 2010–2011 when I was in sabbatical leave from the UNAM supported by a PASPA fellowship. I thank the people of TIFR for their hospitality and the TWAS and PASPA for support.

## Bibliography

1. **A. Baker, G. Wüstholz**, *Logarithmic forms and group varieties*, J. Reine Angew. Math. **442** (1993), 19–62.
2. **S. Balasurya, F. Luca**, *Character sums with the aliquot divisors function*, Uniform Distribution Theory **2** (2007), 121–138.
3. **Y. Bugeaud, T. N. Shorey**, *On the Diophantine equation  $\frac{x^m - 1}{x - 1} = \frac{y^n - 1}{y - 1}$* , Pacific J. Math. **207** (2002), 61–75.
4. **G. H. Hardy, E. M. Wright**, *An Introduction to the theory of numbers*, Sixth Edition. Revised by D. R. Heath-Brown and J. H. Silverman. With a foreword by Andrew Wiles. Oxford University Press, Oxford, 2008.
5. **J. H. Loxton**, *Some problem involving powers of integers*, Acta Arith. **46** (1986), 113–123.
6. **Yu. V. Nesterenko, T. N. Shorey**, *On an equation of Goormaghtigh*, Acta Arith. **83** (1998), 381–389.

FLORIAN LUCA

Instituto de Matemáticas  
Universidad Nacional Autónoma de México  
C. P. 58089, Morelia, Michoacán, México  
fluca@matmor.unam.mx