JOURNAL
de Théorie des Nombres
de BORDEAUX
anciennement Séminaire de Théorie des Nombres de Bordeaux

Joachim von Zur Gathen, Arnold Knopfmacher, Florian Luca, Lutz G. Lucht, Igor E. Shparlinski

Average order in cyclic groups

Tome 16, n° 1 (2004), p. 107-123.

<http://jtnb.cedram.org/item?id=JTNB_2004___16_1_107_0>

Journal de Théorie des Nombres
de Bordeaux **16** (2004), 107–123

Average order in cyclic groups

par Joachim von zur GATHEN, Arnold KNOPFMACHER,
Florian LUCA, Lutz G. LUCHT et Igor E. SHPARLINSKI

Résumé. Pour chaque entier naturel n, nous déterminons l'ordre moyen $\alpha(n)$ des éléments du groupe cyclique d'ordre n. Nous montrons que plus de la moitié de la contribution à $\alpha(n)$ provient des $\varphi(n)$ éléments primitifs d'ordre n. Il est par conséquent intéressant d'étudier également la fonction $\beta(n) = \alpha(n)/\varphi(n)$. Nous déterminons le comportement moyen de α, β, $1/\beta$ et considérons aussi ces fonctions dans le cas du groupe multiplicatif d'un corps fini.

Abstract. For each natural number n we determine the average order $\alpha(n)$ of the elements in a cyclic group of order n. We show that more than half of the contribution to $\alpha(n)$ comes from the $\varphi(n)$ primitive elements of order n. It is therefore of interest to study also the function $\beta(n) = \alpha(n)/\varphi(n)$. We determine the mean behavior of α, β, $1/\beta$, and also consider these functions in the multiplicative groups of finite fields.

Section 1. Introduction

For a positive integer n, we determine the average order $\alpha(n)$ of the elements in the additive cyclic group $\mathbb{Z}_n$ of order n. The major contribution to $\alpha(n)$ is from the $\varphi(n)$ primitive elements in $\mathbb{Z}_n$, each of order n. We show that, in fact, the other elements never contribute more than the primitive ones do.

More precisely, we consider the relative version $\beta(n) = \alpha(n)/\varphi(n)$. With

$$A = \frac{\zeta(2)\zeta(3)}{\zeta(6)} = \frac{315\,\zeta(3)}{2\pi^4} \approx 1.94359\ 64368$$

we have for $n \geq 2$:

$$1 = \liminf_{n\longrightarrow\infty} \beta(n) < \beta(n) < \limsup_{n\longrightarrow\infty} \beta(n) = A.$$

We also determine the mean behavior of α, β, and $1/\beta$, and discuss the average order of elements in the multiplicative groups of finite fields. The lower bounds for β are different for even and for odd characteristic.

Manuscrit reçu le 19 avril 2002.

The original motivation for this research was the usage of groups in cryptography. Here one looks for cyclic groups of large order (preferably a prime number). If we take a finite field and pick a random element from it, how large can we expect its order to be? Intuition says that one should avoid fields whose multiplicative group order is largely made up from small prime factors. The results of this paper put this intuition on a firm basis.

Section 2. The average order

For $a \in \mathbb{Z}_n$, we denote by $\mathrm{ord}(a)$ its order in the additive group $\mathbb{Z}_n$. Then $\mathrm{ord}(a)$ divides n, and for each divisor d of n, there are exactly $\varphi(d)$ elements in $\mathbb{Z}_n$ of order d. Thus the average order in $\mathbb{Z}_n$ is

$$\alpha(n) = \frac{1}{n} \sum_{a \in \mathbb{Z}_n} \mathrm{ord}(a) = \frac{1}{n} \sum_{d|n} d\varphi(d).$$

The main contribution is the term with $d = n$, and we normalize by it:

$$\beta(n) = \frac{\alpha(n)}{\varphi(n)}.$$

Since $1/n$ and $\varphi(n)$ are multiplicative functions of n, so is their Dirichlet convolution $\alpha(n)$ (see Apostol 1976, Theorem 2.14), and also $\beta(n)$. We first determine their values in the case of a prime power.

Lemma 2.1. *Let p be a prime and $k \geq 1$ an integer. Then*

$$\alpha(p^k) = \frac{p^{k+1}}{p+1} + \frac{1}{p^k(p+1)}, \quad \beta(p^k) = 1 + \frac{1}{p^2-1}\left(1 + \frac{1}{p^{2k-1}}\right).$$

In particular $\beta(1) = 1 < \beta(p^{k+1}) < \beta(p^k) \leq \beta(p) = 1 + 1/(p^2-p)$.

Proof. We have

$$\begin{aligned} \alpha(p^k) &= \frac{1}{p^k} \sum_{0 \leq i \leq k} p^i \varphi(p^i) = \frac{1}{p^k}\left(1 + \sum_{1 \leq i \leq k} (p-1) \cdot p^{2i-1}\right) \\ &= \frac{p^{2k+1}+1}{p^k(p+1)}, \\ \beta(p^k) &= \frac{p^{2k+1}+1}{p^k(p+1)(p-1)p^{k-1}} = 1 + \frac{1}{p^2-1}\left(1 + \frac{1}{p^{2k-1}}\right). \end{aligned}$$

□

Theorem 2.2. *For an integer $n \geq 2$, we have the following inequalities.*

(i) $1 \leq \prod_{p|n}\left(1 + \frac{1}{p^2-1}\right) < \beta(n) \leq \prod_{p|n}\left(1 + \frac{1}{p(p-1)}\right) < A.$

(ii) $1 = \liminf_{n \longrightarrow \infty} \beta(n) < \beta(n) < \limsup_{n \longrightarrow \infty} \beta(n) = A.$

Proof. We have

$$\prod_{p \text{ prime}} \left(1 + \frac{1}{p(p-1)}\right) = \prod_{p \text{ prime}} \frac{1 - p^{-1} + p^{-2}}{1 - p^{-1}}$$
$$= \prod_{p \text{ prime}} \frac{1 - p^{-6}}{(1 - p^{-2})(1 - p^{-3})} = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \;=\; A.$$

Claim (i) now follows from the multiplicativity of β and the lemma. For (ii), we clearly have $1 < \beta(n) < A$ for all $n \geq 2$. When n ranges through the primes, then $\beta(n) = 1 + \frac{1}{n(n-1)}$ tends to 1, and when n_k is the product of the first k primes, then $\lim_{k\to\infty} \beta(n_k) = A$. □

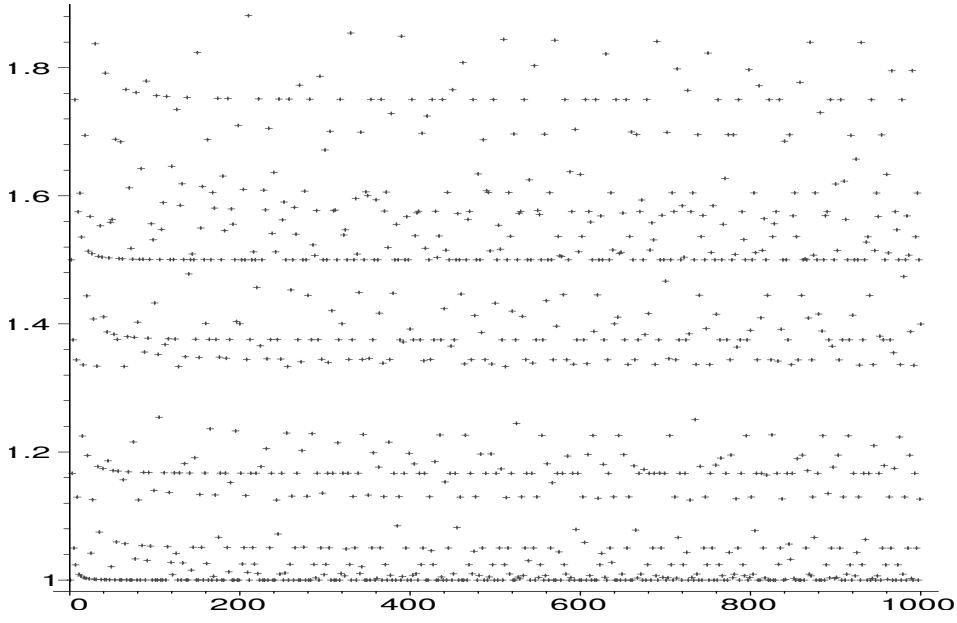

FIGURE 2.1. Relative average order $\beta(n)$ for $n \leq 1000$.

Figure 2.1 shows the behavior of $\beta(n)$ for $n \leq 1000$. The visible bands at $1 = \beta(1)$, $1.5 = \beta(2)$, $1.17 \approx \beta(3)$, for example, are created by numbers of the form $n = kp$ with small k and p either prime or having only large prime factors, namely $k = 1, 2, 3$ for the bands mentioned.

We have seen that $\alpha(n)$ is firmly wedged between $\varphi(n)$ and $A \cdot \varphi(n)$. Since $\liminf_{n\to\infty} \varphi(n)/n = 0$, we also have

$$\liminf_{n\to\infty} \alpha(n)/n = 0.$$

Theorem 4.4 below shows that this lower limit is even obtained on subsequences corresponding to the multiplicative groups of finite fields.

Our upper limit A occurs in several other contexts. Kendall & Rankin (1947), Section 3, consider the number of divisors of n that are divisible by the squarefree part of n, and show that its asymptotic mean value is A. Knopfmacher (1973) gives a more precise description of the mean value, and Knopfmacher (1972), Theorem 3.1 (vi), presents a generalization. The moments of this function are studied in Knopfmacher & Ridley (1974), Theorem 4.4. LeVeque (1977), Problem 6.5, determines A as the sum given in (3.7) below, and shows that the asymptotic mean of $1/\varphi$ is $Ax^{-1}\log x$. The constant A also appears in Bateman (1972), Montgomery (1970), and Riesel & Vaughan (1983).

Throughout the paper, $\log x$ is the natural logarithm of x.

Section 3. The mean average order

In this section, we determine the mean of the averaging functions α and β, and of $\gamma = 1/\beta$. A pleasant feature, due to double averaging, is that the error terms become rather small. We denote the average of an arithmetic function g by $\bar{g}$—not to be confused with complex conjugation:

$$\bar{g}(x) = \frac{1}{x} \sum_{1 \le n \le x} g(n)$$

for $x \ge 1$. There is a well-developed theory with many general results about the existence of means of arithmetic functions, see Elliott (1985); Indlekofer (1980, 1981); Postnikov (1988). However, those general results do not imply the specific statements of this work.

The average $\bar{\alpha}$ is connected to the constant

$$C_\alpha = \frac{\zeta(3)}{2\zeta(2)} = \frac{3\zeta(3)}{\pi^2} \approx 0.36538\,14847.$$

Theorem 3.1. *The mean $\bar{\alpha}$ of α satisfies*

$$\bar{\alpha}(x) = C_\alpha x + O((\log x)^{2/3}(\operatorname{loglog} x)^{4/3}) \text{ for } x \ge 3.$$

Proof. We have

$$\bar{\alpha}(x) = \frac{1}{x} \sum_{1 \le n \le x} \alpha(n) \;=\; \frac{1}{x} \sum_{1 \le n \le x} \frac{1}{n} \sum_{d \mid n} d\varphi(d) = \frac{1}{x} \sum_{1 \le k \le x} \frac{1}{k} \sum_{d \le x/k} \varphi(d).$$

Walfisz (1963), Chapter IV, proves via exponential sum estimates that

$$\left| \bar{\varphi}(x) - \frac{x}{2\zeta(2)} \right| \le c\,(\log x)^{2/3}(\operatorname{loglog} x)^{4/3} \text{ for } x \ge 3,$$

with some constant c. Now from

$$\sum_{x < k} \frac{1}{k^3} \le \int_{x-1}^{\infty} \frac{dt}{t^3} = \frac{1}{2(x-1)^2}$$

we obtain

$$\begin{aligned}|\bar\alpha(x) - C_\alpha x| &= \Big|\frac{1}{x}\sum_{1\le k\le x}\frac{1}{k}\sum_{1\le d\le x/k}\varphi(d) - \frac{1}{x}\sum_{1\le k\le x}\frac{1}{k}\cdot\frac{x^2}{2\zeta(2)k^2} \\ &\qquad + \frac{1}{x}\sum_{1\le k\le x}\frac{1}{k}\cdot\frac{x^2}{2\zeta(2)k^2} - \frac{\zeta(3)x}{2\zeta(2)}\Big| \\ &< \frac{1}{x}\sum_{1\le k\le x}\frac{1}{k}\cdot\frac{x}{k}\cdot c\,(\log x)^{2/3}(\log\log x)^{4/3} + \frac{x}{2\zeta(2)}\sum_{x<k}\frac{1}{k^3} \\ &\le c\,\zeta(2)(\log x)^{2/3}(\log\log x)^{4/3} + \frac{x}{4\zeta(2)(x-1)^2}. \qquad \square\end{aligned}$$

Montgomery (1987) has shown that the error in the estimate for $\bar\varphi(x)$ is not $O((\log\log x)^{1/2})$, and conjectured that its maximum order is $\log\log x$.

We also have an explicit but worse error bound, both for $\bar\varphi$ and for $\bar\alpha$.

Lemma 3.2. *For $x \ge 1$, we have*

(i) $\Big|\bar\varphi(x) - \dfrac{x}{2\zeta(2)}\Big| < 2 + \log x$,

(ii) $\big|\bar\alpha(x) - C_\alpha x\big| < 4 + \zeta(2)\log x$.

Proof. It is easily verified that (i) holds for $1 \le x < 2$. We let $x \ge 2$, and observe that

$$\sum_{1\le d\le x}\mu(d)\left\lfloor\frac{x}{d}\right\rfloor = 1$$

for $x \ge 1$, see (Apostol 1976, Theorem 3.12). It follows that

$$\begin{aligned}\bar\varphi(x) &= \frac{1}{2x}\sum_{1\le d\le x}\mu(d)\left(\left\lfloor\frac{x}{d}\right\rfloor^2 + \left\lfloor\frac{x}{d}\right\rfloor\right) = \frac{1}{2x}\sum_{1\le d\le x}\mu(d)\left\lfloor\frac{x}{d}\right\rfloor^2 + \frac{1}{2x} \\ &= \frac{x}{2}\sum_{1\le d\le x}\frac{\mu(d)}{d^2} - \frac{1}{2x}\sum_{1\le d\le x}\mu(d)\left(\frac{x}{d} - \left\lfloor\frac{x}{d}\right\rfloor\right)\left(\frac{x}{d} + \left\lfloor\frac{x}{d}\right\rfloor\right) + \frac{1}{2x}.\end{aligned}$$

Hence

$$\bar\varphi(x) - \frac{x}{2\zeta(2)} = R(x)$$

with

$$\begin{aligned}|R(x)| &= \left|-\frac{x}{2}\sum_{d>x}\frac{\mu(d)}{d^2} - \frac{1}{2x}\sum_{1\le d\le x}\mu(d)\left(\frac{x}{d} - \left\lfloor\frac{x}{d}\right\rfloor\right)\left(\frac{x}{d} + \left\lfloor\frac{x}{d}\right\rfloor\right) + \frac{1}{2x}\right| \\ &\le \frac{x}{2}\sum_{d>x}\frac{1}{d^2} + \sum_{1\le d\le x}\frac{1}{d} + \frac{1}{2x}.\end{aligned}$$

By inserting the estimates

$$\sum_{d>x} \frac{1}{d^2} \le \int_{\lfloor x \rfloor}^{\infty} \frac{dt}{t^2} = \frac{1}{\lfloor x \rfloor}, \quad \sum_{1 \le d \le x} \frac{1}{d} \le 1 + \int_1^x \frac{dt}{t} = 1 + \log x,$$

we see that for $x \ge 2$

$$|R(x)| \le 1 + \log x + \frac{1}{2}\left(\frac{x}{\lfloor x \rfloor} + \frac{1}{x}\right) < 2 + \log x.$$

This shows (i), and (ii) follows by inserting (i) into the proof of Theorem 3.1. □

For two arithmetic functions $f, g\colon \mathbb{N} \longrightarrow \mathbb{C}$, $f * g$ is their Dirichlet convolution, with

$$(f * g)(n) = \sum_{d|n} f(d) g(n/d)$$

for all n. Furthermore, we denote by $\mathbf{1}$ the constant function on $\mathbb{N}$ with value 1, and μ is the Möbius function.

Lemma 3.3. *Let f and g be arithmetic functions with $f = \mathbf{1} * g$, and consider the Dirichlet series*

$$\widetilde{g}(s) = \sum_{n \ge 1} \frac{g(n)}{n^s}.$$

(i) *If $\widetilde{g}(s)$ is absolutely convergent for $\Re s \ge 0$, then the mean of f is*

$$\bar{f}(x) = \widetilde{g}(1) + O\left(\frac{1}{x}\right),$$

and more precisely

$$|\bar{f}(x) - \widetilde{g}(1)| \le \frac{1}{x} \sum_{n \ge 1} |g(n)|$$

for $x \ge 1$.

(ii) *If f is multiplicative and $\widetilde{g}(s)$ converges absolutely for some s with $\Re s \ge 0$, then $\widetilde{g}(s)$ can be written as the Euler product*

$$\widetilde{g}(s) = \prod_{p \text{ prime}} \left(1 + \frac{g(p)}{p^s} + \frac{g(p^2)}{p^{2s}} + \cdots\right).$$

The absolute convergence of $\widetilde{g}(s)$ is equivalent to

$$\sum_{\substack{p \text{ prime} \\ k \ge 1}} \frac{|f(p^k) - f(p^{k-1})|}{p^{ks}} < \infty\,.$$

Proof. For $x \geq 1$, we have

$$\begin{aligned}\left|\sum_{n\leq x} f(n) - x\widetilde{g}(1)\right| &= \left|\sum_{d\leq x} g(d)\left\lfloor \frac{x}{d}\right\rfloor - x\widetilde{g}(1)\right| \\ &\leq \left|x\sum_{d\leq x}\frac{g(d)}{d} - x\widetilde{g}(1)\right| + \sum_{d\leq x}|g(d)| \\ &\leq x\sum_{d>x}\frac{|g(d)|}{d} + \sum_{d\leq x}|g(d)| \leq \sum_{1\leq d}|g(d)|,\end{aligned}$$

which implies (i).

If f is multiplicative, then so is $g = \mu * f$, and $g(p^k) = f(p^k) - f(p^{k-1})$ for all primes p and $k \in \mathbb{N}$. Now the Euler product representation of $\widetilde{g}(s)$ follows from the unique factorization theorem. If $\widetilde{g}(s)$ is absolutely convergent, then so is the partial series $\sum g(p^k)\,p^{-ks}$ taken over all prime powers p^k. Conversely, absolute convergence of the latter series implies that for any $x \geq 1$

$$\begin{aligned}\sum_{n\leq x}\frac{|g(n)|}{n^{\Re s}} &\leq \prod_{\substack{p\leq x\\ p\text{ prime}}}\left(1+\sum_{k\geq 1}\frac{|g(p^k)|}{p^{k\Re s}}\right) \\ &\leq \prod_{\substack{p\leq x\\ p\text{ prime}}}\exp\left(\sum_{k\geq 1}\frac{|g(p^k)|}{p^{k\Re s}}\right) \leq \exp\left(\sum_{\substack{p\text{ prime}\\ k\geq 1}}\frac{|g(p^k)|}{p^{k\Re s}}\right) < \infty\,.\end{aligned}$$

Thus $\widetilde{g}(s)$ converges absolutely, which finishes the proof of (ii). □

The mean of β is connected to the constant

$$C_\beta = \frac{\zeta(3)\zeta(4)}{\zeta(8)} = \frac{105\,\zeta(3)}{\pi^4} \approx 1.29573\,09579.$$

Theorem 3.4. *The average value $\bar\beta$ of β equals $C_\beta + O(x^{-1})$, and more precisely*

$$|\bar\beta(x) - C_\beta| < x^{-1}\prod_{p\,\text{prime}}\left(1+\frac{p+2}{p^3-p}\right)$$

for $x \geq 1$.

Proof. We use Lemma 3.3 with $f = \beta$ and $g = \mu * \beta$. Thus

$$g(p^k) = \beta(p^k) - \beta(p^{k-1}) = \begin{cases}\dfrac{1}{p(p-1)} & \text{for } k = 1,\\ -\dfrac{1}{p^{2k-1}} & \text{for } k \geq 2,\end{cases}$$

for a prime p and an integer $k \geq 1$. Due to

$$\sum_{k\geq 1} |g(p^k)| = \frac{1}{p\,(p-1)} + \sum_{k\geq 2} \frac{1}{p^{2k-1}} = \frac{p+2}{p^3-p}, \tag{3.5}$$

the series $\sum |g(p^k)|\, p^{-k\Re s}$ taken over all prime powers p^k converges for $\Re s \geq 0$, and Lemma 3.3 (ii) implies the absolute convergence of the Dirichlet series $\widetilde{g}(s)$. In particular, we obtain

$$\begin{aligned}\widetilde{g}(1) &= \prod_{p \text{ prime}} \left(1 + \frac{1}{p^2(p-1)} - \sum_{k\geq 2} \frac{1}{p^{3k-1}}\right) = \prod_{p \text{ prime}} \left(1 + \frac{p+1}{p\,(p^3-1)}\right) \\ &= \prod_{p \text{ prime}} \frac{1+p^{-4}}{1-p^{-3}} = \prod_{p \text{ prime}} \frac{1-p^{-8}}{(1-p^{-3})(1-p^{-4})} = \frac{\zeta(3)\,\zeta(4)}{\zeta(8)} = C_\beta\,.\end{aligned}$$

Finally, Lemma 3.3 combined with (3.5) yields

$$\begin{aligned}|\bar{\beta}(x) - C_\beta| &< x^{-1} \sum_{d\geq 1} |g(d)| = x^{-1} \prod_{p \text{ prime}} \Big(1 + \sum_{k\geq 1} |g(p^k)|\Big) \\ &= x^{-1} \prod_{p \text{ prime}} \left(1 + \frac{p+2}{p^3-p}\right) \approx 2.26507\,69892 \cdot x^{-1},\end{aligned}$$

which completes the proof. □

It is interesting to compare the behavior of $\bar{\beta}(x) \approx \zeta(3)\zeta(4)/\zeta(8)$ with its naive "prediction" $\bar{\alpha}(x)/\bar{\varphi}(x) \approx \zeta(3)$, see Theorem 3.1 and Lemma 3.2. We have $\zeta(4)/\zeta(8) \approx 1.07792\,81367$.

Figure 3.1 shows the behavior of

$$(\bar{\beta}(x) - C_\beta) \cdot x \cdot \prod_{p \text{ prime}} \left(1 + \frac{p+2}{p^3-p}\right)^{-1} \tag{3.6}$$

for integer $x \leq 1000$. Theorem 3.4 says that this quantity is absolutely smaller than 1.

We can also express our constants A and C_β as sums of Dirichlet series via the Euler product decomposition

$$\sum_{1\leq n} f(n) = \prod_{p \text{ prime}} (1 + f(p) + f(p^2) + \cdots),$$

which is valid for a multiplicative function f in the case of absolute convergence. Now

$$1 + \frac{1}{p(p-1)} = \sum_{0\leq k} \frac{(\mu(p^k))^2}{p^k \varphi(p^k)}, \quad 1 + \frac{1}{p(p^3-1)} = \sum_{0\leq k} \frac{(\mu(p^k))^2}{p^k \varphi(p^k) \sigma(p^{2k})}$$

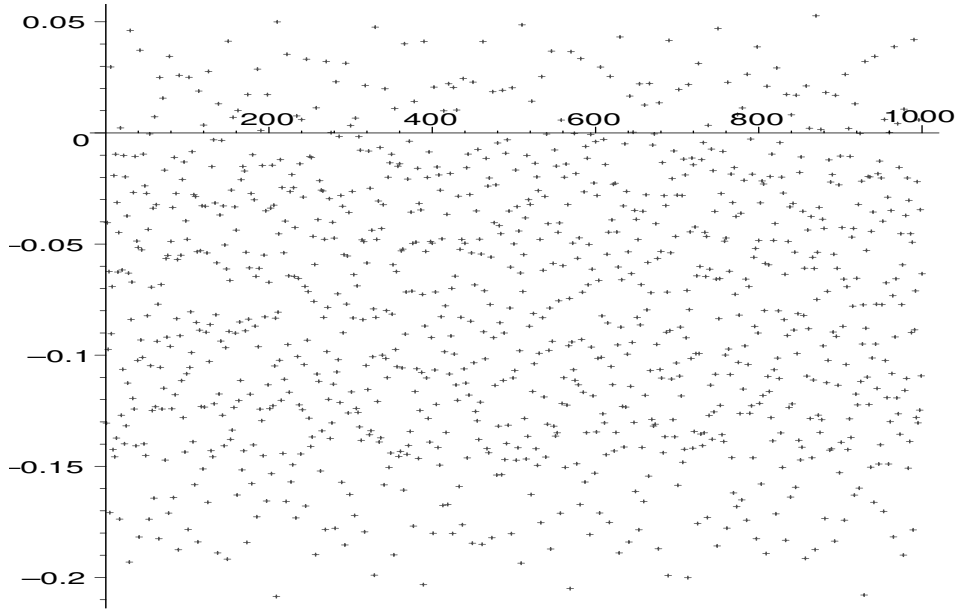

FIGURE 3.1. The average of β normalized as in (3.6).

imply that

$$A = \sum_{1\le n} \frac{(\mu(n))^2}{n\varphi(n)}, \quad C_\beta = \sum_{1\le n} \frac{(\mu(n))^2}{n\varphi(n)\sigma(n^2)}. \tag{3.7}$$

Both series seem to converge much slower than the product representations.

The mean of the function $\gamma = 1/\beta = \varphi/\alpha$ is connected to the constant

$$C_\gamma = \prod_{p \text{ prime}} \left(1 - \frac{1}{p}\right)\left(1 + \left(1 - \frac{1}{p^2}\right)\sum_{1\le k} \frac{1}{p^k + p^{-k-1}}\right) \approx 0.80146\,96934.$$

Theorem 3.8. *The mean $\bar\gamma$ of γ satisfies $\bar\gamma(x) = C_\gamma + O(x^{-1})$, and more precisely,*

$$|\bar\gamma(x) - C_\gamma| \le Dx^{-1}$$

for a constant D which is explicitly given in the proof below.

Proof. Again, we use Lemma 3.3, with the multiplicative function $f = \gamma$. For a prime p and $k \ge 1$, we have

$$f(p^k) = \frac{1}{\beta(p^k)} = \frac{1 - p^{-2}}{1 + p^{-2k-1}},$$

by Lemma 2.1. For the multiplicative function $g = \mu * f$ we find

$$g(p^k) = f(p^k) - f(p^{k-1}) = \begin{cases} -\dfrac{1+p^{-1}}{p^2(1+p^{-3})} & \text{if } k = 1, \\ \dfrac{1}{p^{2k-1}(1+p^{-2k+1})(1+p^{-2k-1})} & \text{if } k \ge 2. \end{cases}$$

Thus the Dirichlet series $\widetilde{g}(s)$ is absolutely convergent for $\Re s \ge 0$. We have

$$\widetilde{g}(1) = \sum_{1 \le n} \frac{g(n)}{n} = \prod_{p \text{ prime}} \left(1 + \sum_{1 \le k} \frac{g(p^k)}{p^k}\right).$$

For a prime p, the factor in this product equals

$$\begin{aligned} 1 + \sum_{1 \le k} \frac{g(p^k)}{p^k} &= 1 + \sum_{1 \le k} \frac{f(p^k) - f(p^{k-1})}{p^k} = 1 - \frac{1}{p} + \sum_{1 \le k} \left(1 - \frac{1}{p}\right) \frac{f(p^k)}{p^k} \\ &= \left(1 - \frac{1}{p}\right)\left(1 + \left(1 - \frac{1}{p^2}\right) \sum_{1 \le k} \frac{1}{p^k + p^{-k-1}}\right). \end{aligned}$$

Lemma 3.3 now implies that

$$|\bar{\gamma}(x) - C_\gamma| \le \frac{1}{x} \sum_{1 \le n} |g(n)|,$$

and the claim follows from the absolute convergence of $\widetilde{g}(0)$. A numerical evaluation of the error term gives

$$D = \sum_{1 \le n} |g(n)| = \prod_{p \text{ prime}} \left(1 + \sum_{1 \le k} |g(p^k)|\right) \approx 1.96531. \qquad \square$$

We have $C_\beta \cdot C_\gamma \approx 1.03848\,90929$.

Section 4. Finite fields

Our original motivation for this work was to study the average order in the (cyclic) multiplicative group $\mathbb{F}_q^\times = \mathbb{F}_q \backslash \{0\}$ of a finite field $\mathbb{F}_q$. We first show that for the two families $q = 2^k$ and q a prime, $\alpha(q-1)/(q-1)$ is on average between two positive constants, and also exhibit subfamilies for which this quotient tends to zero. We also obtain several results for $\beta(q-1)$.

Theorem 4.1. *There are two absolute constants $A_2 \ge A_1 > 0$ such that for all $K \ge 1$*

$$A_1 \le \frac{1}{K} \sum_{1 \le k \le K} \frac{\alpha(2^k - 1)}{2^k - 1} \le A_2.$$

Proof. We use the asymptotic formula from Shparlinski (1990)

$$\frac{1}{K}\sum_{1\le k\le K}\frac{\varphi(2^k-1)}{2^k-1}=\eta+O(K^{-1}\log K), \tag{4.2}$$

with η given by the absolutely convergent series

$$\eta=\sum_{d\ \mathrm{odd}}\frac{\mu(d)}{dt_d}\approx 0.73192,$$

where t_d is the multiplicative order of 2 modulo d. The claim follows from (4.2) and Theorem 2.2. □

The proof of Theorem 4.1 implies that for any constant $c<\eta$, $\alpha(2^k-1)\ge c\cdot(2^k-1)$ for infinitely many integers k. We may, of course, take $A_2=1$ in Theorem 4.1; it is not clear whether Theorem 4.1 holds with a smaller value of A_2. We also see that for any $\varepsilon>0$ and sufficiently large values of K, Theorem 4.1 holds with $A_1=\eta-\varepsilon$.

Stephens (1969) shows in his Lemma 1 that

$$\sum_{p\le x}\frac{\varphi(p-1)}{p}=\kappa\,\mathrm{li}x+O(x/(\log x)^D),$$

where

$$\kappa=\sum_{d\ge 1}\frac{\mu(d)}{d\varphi(d)}=\prod_p\left(1-\frac{1}{p(p-1)}\right)\approx 0.37397$$

is Artin's constant and $D>1$ is arbitrary. The sum does not change by much if we replace p by $p-1$ in the denominator, since

$$\sum_{p\le x}\frac{\varphi(p-1)}{p-1}=\sum_{p\le x}\frac{\varphi(p-1)}{p}+O(\log\log x).$$

Using the bounds of Theorem 2.2 on $\beta=\alpha/\varphi$, the fact that $p-1$ is even for $p\ge 3$, and $\beta(2)=3/2$, we find that

$$\frac{3\kappa}{2}-\epsilon\le\frac{1}{x}\sum_{p\le x}\frac{\alpha(p-1)}{p-1}\le A\kappa+\epsilon, \tag{4.3}$$

for any $\epsilon>0$ and sufficiently large x.

Thus there is an infinite sequence of fields of characteristic 2, and also one of prime fields, in which the average order is close to its largest possible value. Now we show that $\alpha(2^k-1)$ and $\alpha(p-1)$ infinitely often take relatively small values, just as $\varphi(2^k-1)$ and $\varphi(p-1)$ do.

Theorem 4.4. *For infinitely many integers $k \geq 3$ and for infinitely many primes p, we have*

$$\frac{\alpha(2^k-1)}{2^k-1} = O\left(\frac{1}{\operatorname{loglog} k}\right) \qquad \text{and} \qquad \frac{\alpha(p-1)}{p-1} = O\left(\frac{1}{\operatorname{loglog} p}\right).$$

Proof. Let p_i denote the ith prime. For an integer $r \geq 1$, we put

$$k_r = (p_2-1)\cdots(p_r-1) \qquad \text{and} \qquad m_r = p_2\cdots p_r = n_r/2.$$

Then m_r divides $2^{k_r}-1$, and therefore

$$\alpha\left(2^{k_r}-1\right) < A\cdot\varphi\left(2^{k_r}-1\right) \leq \frac{A\cdot(2^{k_r}-1)\varphi(m_r)}{m_r}.$$

Using the bound

$$\varphi(m_r) = O\left(\frac{m_r}{\log(\log m_r+1)}\right)$$

(see Hardy & Wright (1962), Theorem 328) and $k_r < m_r$, we obtain the first statement.

To prove the second bound, we select q_r as the smallest prime number in the arithmetic progression 1 mod m_r. Then

$$\alpha(q_r-1) < A\cdot\varphi(q_r-1) \leq \frac{A\cdot(q_r-1)\varphi(m_r)}{m_r} = O\left(\frac{q_r}{\log(\log m_r+1)}\right).$$

From Linnik's Theorem on the smallest prime number in an arithmetic progression, we have $\log q_r = O(\log m_r)$, and the result follows. □

In particular,

$$\liminf_{k\to\infty} \alpha(2^k-1)/(2^k-1) = \liminf_{q \text{ prime}} \alpha(q-1)/(q-1) = 0.$$

Open Question. *Obtain analogs of (4.2) and (4.3) for the sums*

$$\sum_{1\leq k\leq K} \frac{\alpha(2^k-1)}{2^k-1}, \qquad \sum_{1\leq k\leq K} \beta(2^k-1), \qquad \sum_{p\leq K} \frac{\alpha(p-1)}{p-1}, \qquad \sum_{p\leq K} \beta(p-1).$$

In the above we considered only $\alpha(2^k-1)$. Similar results also hold for $\alpha(p^k-1)$ for any fixed p and growing k.

The convergence to zero of $\alpha(q_r-1)/(q_r-1)$ as above seems rather slow. For the largest known "primorial prime" $q = n_{33\,237}+1$, where as before n_k is the product of the first k primes (see Caldwell & Gallot 2000), with 169 966 digits and the largest prime factor $p_{33\,237} = 392\,113$ of $q-1$, we have $\alpha(q-1)/(q-1) \approx 0.0847$. Also, $\beta(q-1) \approx 1.94359\,608$ is close to A.

Concerning lower bounds for β, the situation is quite different between characteristic 2 and odd characteristic.

In a finite field $\mathbb{F}_{2^k}$ of characteristic 2, the group of units is cyclic with $2^k - 1$ elements. For a Mersenne prime $M_k = 2^k - 1$, we have $\beta(M_k) = 1 + (M_k^2 - M_k)^{-1}$. If there are infinitely many of them, then $\liminf \beta(2^k - 1) = 1$. For the current world record $k = 69\,72593$ (see Chris Caldwell's web site `http://www.utm.edu/research/primes`), we have $\beta(M_k) \approx 1 + 0.52 \cdot 10^{-41\,97919}$.

For a field $\mathbb{F}_q$ of odd characteristic, 2 divides $q - 1 = \#\mathbb{F}_q^\times$ and thus $\beta(q-1) > 4/3$, by Lemma 2.1. For a prime $q = m2^k + 1$ with m odd, we have

$$\beta(q-1) = \frac{4}{3}\left(1 + \frac{1}{2^{2k+1}}\right) \cdot \beta(m).$$

As an example, with the prime $m = 10^{500} + 961$ and $k = 3103$, q is indeed prime (Keller 2000), and

$$\beta(q-1) \approx \frac{4}{3}(1 + 10^{-1000}).$$

We now prove the limits indicated by these experimental results.

Theorem 4.5. *We have*

(i) $\liminf_{p \text{ prime}} \beta(p-1) = 4/3$,

(ii) $\limsup_{p \text{ prime}} \beta(p-1) = A$,

(iii) $\liminf_{k\to\infty} \beta(2^k - 1) = 1$.

Proof. To show that the limit in (i) is at least 4/3, we notice that if $p \geq 3$, then $p - 1 = 2^k m$ with some $k \geq 1$ and some odd integer m, and therefore

$$\beta(p-1) = \beta(2^k m) = \frac{4}{3}\Big(1 + \frac{1}{2^{2k+1}}\Big)\beta(m) > \frac{4}{3}.$$

For the equality in (i), we use a theorem of Chen (see Chen (1973), or Lemma 1.2 in Ford (1999), or Chapter 11 of Halberstam & Richert (1974)) which says that for each even natural number n there exists x_0 such that for every $x \geq x_0$ there exists a prime number $p \in (x/2, x]$ with $p \equiv 1 \bmod n$ such that $(p-1)/n$ has at most two prime factors, and each of them exceeds $x^{1/10}$.

We now choose a positive integer k and apply Chen's Theorem with $n = 2^k$ to conclude that there exist infinitely many prime numbers p such that $p - 1 = 2^k m$, where m has at most two prime factors, and each of them exceeds $p^{1/10}$. For such prime numbers p, we have

$$\beta(p-1) = \beta(2^k m) = \beta(2^k)\beta(m) = \frac{4}{3}\Big(1 + \frac{1}{2^{2k+1}}\Big)\beta(m).$$

If m is prime, then

$$\beta(m) = 1 + \frac{1}{m(m-1)},$$

if $m = r^2$ is a square of a prime, we have

$$\beta(m) = 1 + \frac{r^3+1}{r^3(r^2-1)},$$

while if $m = rs$ is a product of two distinct primes, then

$$\beta(m) = \Big(1 + \frac{1}{r(r-1)}\Big)\Big(1 + \frac{1}{s(s-1)}\Big),$$

by Lemma 1. At any rate, with k fixed and p tending to infinity through prime numbers of the above form, we get that the number $\frac{4}{3}\Big(1 + \frac{1}{2^{2k+1}}\Big)$ is a cluster point for the set $B = \{\beta(p-1)\colon p \text{ prime}\}$. Since this is true for all positive integers k, we get that $4/3$ is also a cluster point for B, which takes care of (i).

For (ii), Theorem 1 says that the limit in (ii) is at most A. To show equality, we let x be a large positive real number, write

$$P_x = \prod_{p \le x} p,$$

and let q_x be the smallest prime number in the arithmetic progression $P_x + 1 \bmod P_x^2$, which exists by Dirichlet's Theorem, since $P_x + 1$ is coprime to P_x^2. We have $q_x - 1 \equiv P_x \bmod P_x^2$ and may write

$$q_x - 1 = P_x m_x,$$

where each prime factor of m_x is larger than x. Thus

$$\begin{aligned} \beta(q_x - 1) &= \beta(P_x)\beta(m_x) = \prod_{p \le x}\Big(1 + \frac{1}{p(p-1)}\Big) \cdot \beta(m_x), \\ \frac{\beta(q_x-1)}{A} &= \beta(m_x) \cdot \prod_{p > x}\Big(1 + \frac{1}{p(p-1)}\Big)^{-1}. \end{aligned} \tag{4.6}$$

We now consider the prime factorization

$$m_x = p_1^{e_1} \cdots p_k^{e_k}$$

of m_x, where $p_1, \dots, p_k > x$ are distinct primes and $e_1, \dots, e_k$ are positive integers. Then

$$\begin{aligned} 1 < \beta(m_x) &\le \prod_{1 \le i \le k} \beta(p_i) = \prod_{1 \le i \le k}\Big(1 + \frac{1}{p_i(p_i-1)}\Big) \\ &< \prod_{x < p}\Big(1 + \frac{1}{p(p-1)}\Big) < \exp\Big(\sum_{x<p} \frac{1}{p(p-1)}\Big) = 1 + O(x^{-1}). \end{aligned} \tag{4.7}$$

Now (4.6) and (4.7) imply that

$$\lim_{x\to\infty} \frac{\beta(q_x - 1)}{A} = 1,$$

which takes care of (ii). To prove (iii), we show that

$$\lim_{p \text{ prime}} \beta(2^p - 1) = 1. \tag{4.8}$$

If d and n are positive integers with d dividing n, then $\varphi(d) \le \varphi(n)$. Hence

$$1 < \beta(n) = \frac{1}{n} \sum_{d|n} d\frac{\varphi(d)}{\varphi(n)} \le \frac{1}{n} \sum_{d|n} d = \frac{\sigma(n)}{n} < \frac{n}{\varphi(n)}.$$

Let p be any prime number and consider the prime factorization

$$2^p - 1 = p_1^{e_1} \cdots p_k^{e_k}$$

of $2^p - 1$. For any $i \le k$, we have $2^p \equiv 1 \bmod p_i$, so that the order of 2 modulo p_i divides p. Since p is prime, it equals this order, and hence $p_i \equiv 1 \bmod p$. In particular, $p_i > p$, and therefore

$$2^p > 2^p - 1 \ge p_1 \cdots p_k > p^k,$$

so that $k < p/\log_2 p$. Thus

$$\begin{aligned} 1 < \beta(2^p - 1) &< \frac{2^p - 1}{\varphi(2^p - 1)} = \prod_{i=1}^{k} \Big(1 + \frac{1}{p_i - 1}\Big) \\ &\le \Big(1 + \frac{1}{p}\Big)^k < \exp\Big(\frac{k}{p}\Big) < \exp\Big(\frac{1}{\log_2 p}\Big) = 1 + o(1), \end{aligned}$$

which proves (4.8). □

Acknowledgements

We are grateful to Wilfrid Keller for help with large primes, to Helmut Prodinger for pointing out a reference, and to Karl-Heinz Indlekofer for useful discussions.

The first author thanks the *John Knopfmacher Centre for Applicable Analysis and Number Theory* for arranging his visit there during which most of this research was done.

References

T. M. Apostol (1976), *Introduction to Analytic Number Theory.* Springer-Verlag, New York.

P. T. Bateman (1972) . *The distribution of values of the Euler function.* Acta Arithmetica **21**, 329–345.

C. K. Caldwell & Y. Gallot (2000), *Some results for* $n! \pm 1$ *and* $2 \cdot 3 \cdot 5 \cdots p \pm 1$. Preprint.

J. R. Chen (1973), *On the representation of a large even integer as a sum of a prime and a product of at most two primes.* Scientia Sinica **16**, 157–176.
P. D. T. A. Elliott (1985), *Arithmetic functions and integer products*, volume **272** of Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, New York.
K. Ford (1999), *The number of solutions of $\phi(x) = m$.* Annals of Mathematics **150**, 1–29.
H. Halberstam & H.E. Richert (1974), *Sieve Methods.* Academic Press.
G. H. Hardy & E. M. Wright (1962), *An introduction to the theory of numbers.* Clarendon Press, Oxford. 1st edition 1938.
K.-H. Indlekofer (1980), *A mean-value theorem for multiplicative functions.* Mathematische Zeitschrift **172**, 255–271.
K.-H. Indlekofer (1981), *Limiting distributions and mean-values of multiplicative arithmetical functions.* Journal für die reine und angewandte Mathematik **328**, 116–127.
W. Keller (2000). Private communication.
D. G. Kendall & R. A. Rankin (1947), *On the number of Abelian groups of a given order.* Quarterly Journal of Mathematics **18**, 197–208.
J. Knopfmacher (1972), *Arithmetical properties of finite rings and algebras, and analytic number theory. II.* Journal für die reine und angewandte Mathematik **254**, 74–99.
J. Knopfmacher (1973), *A prime divisor function.* Proceedings of the American Mathematical Society **40**, 373–377.
J. Knopfmacher & J. N. Ridley (1974), *Prime-Independent Arithmetical Functions.* Annali di Matematica **101**(4), 153–169.
W. LeVeque (1977), *Fundamentals of Number Theory.* Addison-Wesley.
H. L. Montgomery (1970), *Primes in arithmetic progressions.* Michigan Mathematical Journal **17**, 33–39.
H. L. Montgomery (1987), *fluctuations in the mean of Euler's phi function.* Proceedings of the Indian Academy of Sciences (Mathematical Sciences) **97**(1-3), 239–245.
A. G. Postnikov (1988), *Introduction to analytic number theory.* Volume **68** of Translations of Mathematical Monographs. American Mathematical Society, Providence, RI.
H. Riesel & R. C. Vaughan (1983), *On sums of primes.* Arkiv for Matematik **21**(1), 46–74.
I. E. Shparlinski (1990), *Some arithmetic properties of recurrence sequences.* Matematicheskie Zametki **47**(6), 124–131. (in Russian); English translation in Mathematical Notes **47**, (1990), 612–617.
P. J. Stephens (1969), *An Average Result for Artin's Conjecture.* Mathematika **16**(31), 178–188.
A. Walfisz (1963), *Weylsche Exponentialsummen in der neueren Zahlentheorie.* Number **15** in Mathematische Forschungsberichte. VEB Deutscher Verlag der Wissenschaften, Berlin.

Joachim VON ZUR GATHEN
Fakultät für Elektrotechnik, Informatik und Mathematik
Universität Paderborn, 33095 Paderborn, Germany
E-mail : gathen@upb.de
URL: http://www-math.upb.de/~aggathen/

Arnold KNOPFMACHER
The John Knopfmacher Centre
for Applicable Analysis and Number Theory
University of the Witwatersrand
P.O. Wits 2050, South Africa
E-mail : arnoldk@cam.wits.ac.za
URL: http://www.wits.ac.za/science/number_theory/arnold.htm

Florian LUCA
Instituto de Matemáticas
Universidad Nacional Autónoma de México
C.P. 58180, Morelia, Michoacán, México
E-mail : fluca@matmor.unam.mx

Lutz G. LUCHT
Institut für Mathematik
TU Clausthal, Erzstraße 1
38678 Clausthal-Zellerfeld, Germany
E-mail : lucht@math.tu-clausthal.de
URL: http://www.math.tu-clausthal.de/personen/lucht.html

Igor E. SHPARLINSKI
Department of Computing
Macquarie University
Sydney, NSW 2109, Australia
E-mail : igor@comp.mq.edu.au
URL: http://www.comp.mq.edu.au/~igor