

CENTRAL MOMENTS FOR ARITHMETIC FUNCTIONS

JOSEPH E. COLLISON

The only central moment considered in probabilistic number theory up until now has been the "variance" of an arithmetic function. This paper considers the case of higher central moments for such functions. It will be shown that if f is an additive complex valued arithmetic function then

$$\sum_{m \leq n} |f(m) - A(n)|^{2K} = O(n(\log \log n)^{2K-2} \sum_{p^\alpha \leq n} |f(p^\alpha)|^{2K} p^{-\alpha})$$

where K is a positive integer and

$$A(n) = \sum_{p^\alpha \leq n} f(p^\alpha) p^{-\alpha}.$$

It will also be shown that if f is an additive real valued arithmetic function and K is an odd positive integer, then

$$\sum_{m \leq n} (f(m) - A(n))^K = O(n(\log \log n)^{K-2+1/K} \sum_{p^\alpha \leq n} |f(p^\alpha)|^K p^{-\alpha}).$$

1. Preliminaries. Given a fixed positive integer K let X be a K -tuple of prime powers p^α , where the primes need not be distinct. Y is defined similarly. Next we define

$$\|X\| = \text{Max} \{p^\alpha: p^\alpha \text{ is a component of } X\}$$

and $|X| = \prod p^\alpha$ where the product is over those p^α which are components of X . By X_j we shall mean the j -tuple consisting of the first j components of X , and \tilde{X}_j shall denote the $K - j$ -tuple consisting of the last $K - j$ components of X . $X_j Y_k$ shall denote the first j components of X followed by the first k components of Y . By $X_j \|m$ we shall mean that $p^\alpha \|m$ for all the components of X_j . If f is an arithmetic function, then we define $F(X)$ to be $\prod f(p^\alpha)$ where the product is over all the components p^α of X .

LEMMA 1. Given the M distinct prime powers $P_i = p_i^{\alpha_i}$, $i = 1, \dots, M$, and the positive integer n ,

$$W(M, n) = n^{-1} \sum_{\substack{k \leq n \\ P_i | k, i \leq M}} 1 = \prod_{i=1}^M P_i^{-1} (1 - p_i^{-1}) + O(n^{-1})$$

where $|O(n^{-1})| \leq (3 \cdot 2^M - 1)n^{-1}$.

Proof. Let $N = L \prod_{i=1}^M P_i$ for any positive integer L . We will now show by induction on M that for all such N

$$(1.1) \quad W(M, N) = \prod_{i=1}^M P_i^{-1} (1 - p_i^{-1}) + O(N^{-1})$$

where $|O(N^{-1})| \leq 3(2^M - 1)N^{-1}$. We have

$$W(1, N) = N^{-1}([N/P_1] - [N/P_1 p_1]) = P_1^{-1} - N^{-1}[N/P_1 p_1]$$

so that the result holds for $M = 1$. Letting

$$K = ([L/p_1] + 1) \prod_{i=2}^M P_i$$

we see that for

$$W'(M, n) = n^{-1} \sum_{\substack{k \leq n \\ P_i | k, 1 < i \leq M}} 1$$

we have

$$W(M, N) = P_1^{-1}W'(M, N/P_1) - (K/N)W'(M, K) + R$$

where

$$R = N^{-1} \sum_{\substack{[N/P_1 p_1] < k \leq K \\ P_i | k, 1 < i \leq M}} 1 \leq N^{-1}(K - [N/P_1 p_1]) \prod_{i=2}^M P_i^{-1} + N^{-1} \leq 2/N.$$

Using estimates provided by the induction hypothesis we see

$$P_1^{-1}W'(M, N/P_1) = P_1^{-1} \prod_{i=2}^M P_i^{-1}(1 - p_i^{-1}) + V_1(N)$$

where $|V_1(N)| \leq 3(2^{M-1} - 1)/N$, and

$$(K/N)W'(M, K) = (P_1^{-1}p_1^{-1} + KN^{-1} - P_1^{-1}p_1^{-1}) \prod_{i=2}^M P_i^{-1}(1 - p_i^{-1}) + V_2(N)$$

where $|V_2(N)| \leq 3(2^{M-1} - 1)/N$. Since

$$0 \leq (KN^{-1} - P_1^{-1}p_1^{-1}) \prod_{i=2}^M P_i^{-1}(1 - p_i^{-1}) \leq L^{-1}P_1^{-1} \prod_{i=2}^M P_i^{-1} = N^{-1}$$

(1.1) now follows.

Let $N = ([n \prod_{i=1}^M P_i^{-1}] + 1) \prod_{i=1}^M P_i$, so that the first part of the proof applies to $W(M, N)$. Then

$$\begin{aligned} |W(M, n) - W(M, N)| &\leq \sum_{j=n}^{N-1} |W(M, j) - W(M, j+1)| \\ &\leq \sum_{j=n}^{N-1} \left| \frac{1}{j} - \frac{1}{j+1} \right| \sum_{\substack{k \leq j \\ P_i | k, i > 0}} 1 + \frac{1}{j+1} \sum_{\substack{j < k \leq j+1 \\ P_i | k, i > 0}} 1 \\ &\leq \sum_{j=n}^{N-1} (j+1)^{-1} \prod_{i=1}^M P_i^{-1} + \sum_{\substack{n < j \leq N \\ P_i | j, i > 0}} j^{-1} \\ &\leq \left(\prod_{i=1}^M P_i^{-1} \right) \log(N/n) + n^{-1} \left(\left[N \prod_{i=1}^M P_i^{-1} \right] - \left[n \prod_{i=1}^M P_i^{-1} \right] \right) \\ &\leq n^{-1} \log \left(1 + n^{-1} \prod_{i=1}^M P_i \right)^{n \prod_{i=1}^M P_i^{-1}} + n^{-1} \leq 2n^{-1} \end{aligned}$$

which provides the desired result.

LEMMA 2. For $M \geq 2$ and letting $P = p^\alpha$ represent the power of a prime

$$\sum_{\substack{P_1 \dots P_M > n \\ P_i \leq n, i \leq M}} P_1^{-1} \dots P_M^{-1} \leq C_1 M^4 \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{M-2}$$

for some absolute constant C_1 .

Proof. Separating the two largest prime powers from the rest we see

$$\sum_{\substack{P_1 \dots P_M > n \\ P_i \leq n}} P_1^{-1} \dots P_M^{-1} \leq M(M-1) \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{M-2} (R_1 + R_2)$$

where

$$R_1 = \sum_{\substack{p^\alpha q^\beta > n \\ p^\alpha \leq n, q^\beta \leq n}} p^{-\alpha} q^{-\beta}$$

is known to be bounded [2; P. 35], and

$$R_2 = \sum_{n^{1/M} < p^\alpha \leq n} p^{-\alpha} \sum_{(n/p^\alpha)^{1/(M-1)} < q^\beta \leq n/p^\alpha} q^{-\beta}.$$

With regard to R_2 , we note that for $np^{-\alpha} \geq 3^M$ the second sum is equal to

$$\begin{aligned} & \log \log [n/p^\alpha] - \log \log [(n/p^\alpha)^{1/(M-1)}] + O(1) \\ & \leq \log \log (n/p^\alpha) - \log \log (n/p^\alpha)^{1/2M} + O(1) \\ & = \log 2M + O(1). \end{aligned}$$

For $np^{-\alpha} < 3^M$ we have $q^\beta < 3^M$ and so the second sum in R_2 is bounded by $\log M + O(1)$ in this case. In a similar manner it can be shown that

$$\sum_{n^{1/M} < p^\alpha \leq n} p^{-\alpha} \leq \log 2M + O(1).$$

Thus there are constants C_3 and C_4 for which

$$\begin{aligned} R_1 + R_2 & \leq (\log 2M)^2 + C_3 \log 2M + C_4 \\ & \leq M^2 + C_3 M + C_4. \end{aligned}$$

Letting $C_1 = 1 + C_3/2 + C_4/4$ we obtain the desired result.

2. Even central moments. Now we shall show that

$$\sum_{m \leq n} |f(m) - A(n)|^{2K} = O\left(n (\log \log n)^{2K-2} \sum_{p^\alpha \leq n} |f(p^\alpha)|^{2K} p^{-\alpha}\right).$$

THEOREM 1. *Let f be an additive complex valued arithmetic function and let K be a fixed positive integer. Then for $n \geq 4$*

$$(2.1) \quad \begin{aligned} M_{2K}(n) &= \sum_{m \leq n} |f(m) - A(n)|^{2K} \\ &\leq (2K)! 1024^K K^{6K} C_2 n \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{2K-2} \sum_{p^\alpha \leq n} |f(p^\alpha)|^{2K} p^{-\alpha}. \end{aligned}$$

Proof. First we will show that

$$(2.2) \quad M_{2K}(n) = n \sum_{\substack{||X|| \leq n, \\ ||Y|| \leq n}} F(X) \overline{F(Y)} \overline{T(X, Y, n)}$$

where

$$\begin{aligned} T(X, Y, n) &= \sum_{j=0}^K \sum_{k=0}^K (-1)^{j+k} \binom{K}{j} \binom{K}{k} |\tilde{X}_j \tilde{Y}_k|^{-1} n^{-1} \sum_{\substack{m \leq n \\ X_j | m, Y_k | m}} 1. \\ M_{2K}(n) &= \sum_{j=0}^K \sum_{k=0}^K (-1)^{j+k} \binom{K}{j} \binom{K}{k} A^{K-j}(n) \overline{A^{K-k}(n)} \\ &\quad \cdot \sum_{m \leq n} \left(\sum_{\substack{p^\alpha \leq n \\ p^\alpha | m}} f(p^\alpha) \right)^j \left(\sum_{\substack{q^\beta \leq n \\ q^\beta | m}} \overline{f(q^\beta)} \right)^k \\ &= \sum_{j=0}^K \sum_{k=0}^K (-1)^{j+k} \binom{K}{j} \binom{K}{k} \left(\sum_{||\tilde{X}_j|| \leq n} F(\tilde{X}_j) |\tilde{X}_j|^{-1} \right) \\ &\quad \cdot \left(\sum_{||\tilde{Y}_k|| \leq n} \overline{F(\tilde{Y}_k)} |\tilde{Y}_k|^{-1} \right) \sum_{\substack{||X_j|| \leq n, \\ ||Y_k|| \leq n}} F(X_j) \overline{F(Y_k)} \sum_{\substack{m \leq n \\ X_j | m, Y_k | m}} 1 \end{aligned}$$

which equals the right side of (2.2).

Now let $M_{2K}(n, t)$ denote the restriction of the sum in (2.2) to those X and Y such that exactly t distinct primes occur in the factorization of $|XY|$. By virtue of the fact that

$$n^{-1} \sum_{\substack{m \leq n \\ X_j | m, Y_k | m}} 1 \leq P^{-1}(X_j, Y_k)$$

where $P(X_j, Y_k)$ is a product of the distinct prime powers p^α in $X_j Y_k$ with α being the highest power of p in $X_j Y_k$, an examination of $T(X, Y, n)$ reveals the fact that in an upper bound of the (j, k) term either $|XY|$ appears in the denominator or at least one prime is repeated in $X_j Y_k$. In the latter case, in order for the (j, k) term to be nonzero, a repeated prime must have the same power everywhere it occurs in $X_j Y_k$. So if r_1, \dots, r_t , where $r_1 + \dots + r_t = 2K$, provide the respective number of times the distinct primes p_1, \dots, p_t are repeated in XY , and $s(i, 1), \dots, s(i, u)$, where $s(i, 1) + \dots + s(i, u) = r_i$, provide the respective number of times the distinct powers $\alpha(i, 1),$

$\dots, \alpha(i, u)$ of p_i occur in XY , then as a result of the above discussion we see that

$$|F(X)F(Y)T(X, Y, n)| \leq \left(\sum_{j=0}^K \binom{K}{j} \right)^2 \prod_{i=1}^t \prod_{k=1}^u |f(p_i^{\alpha(i, k)})|^{s(i, k)} p_i^{-\alpha(i, k)}.$$

Thus we see from (2.2) and the last result that for $t < 2K$

$$|M_{2K}(n, t)| \leq n \sum_{\substack{r_1 + \dots + r_t = 2K \\ r_i > 0, i=1, \dots, t}} \frac{(2K)!}{r_1! \dots r_t!} \left(\sum_{j=0}^K \binom{K}{j} \right)^2 \cdot \prod_{i=1}^t \sum_{p \leq n} \sum_{u=1}^{r_i} \sum_{\substack{s_1 + \dots + s_u = r_i \\ s_k > 0, k=1, \dots, u}} \frac{r_i!}{s_1! \dots s_u!} \prod_{k=1}^u \sum_{\alpha \leq [\log n / \log p]} |f(p^\alpha)|^{s_k} p^{-\alpha}.$$

Since $\sum p^{-\alpha}$ summed over all positive α is bounded by 1, it follows by induction on u that

$$\prod_{k=1}^u \left(\sum_{\alpha \leq [\log n / \log p]} |f(p^\alpha)|^{s_k} p^{-\alpha} \right) \leq 2^{u-1} \sum_{\alpha \leq [\log n / \log p]} |f(p^\alpha)|^{r_i} p^{-\alpha}.$$

Hence

$$|M_{2K}(n, t)| \leq (2K)! 4^K n \sum_{\substack{r_1 + \dots + r_t = 2K \\ r_i > 0, i=1, \dots, t}} \prod_{i=1}^t \left(\sum_{p^\alpha \leq n} |f(p^\alpha)|^{r_i} p^{-\alpha} \right) \cdot \sum_{u=1}^{r_i} 2^{u-1} \sum_{\substack{s_1 + \dots + s_u = r_i \\ s_k > 0, k=1, \dots, u}} 1.$$

Using Hölder's inequality and the fact that the last sum is bounded by r_i^u , we see

$$|M_{2K}(n, t)| \leq (2K)! 4^K n \sum_{\substack{r_1 + \dots + r_t = 2K \\ r_i > 0, i=1, \dots, t}} \prod_{i=1}^t 2^{r_i} r_i^{r_i+1} \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{1-r_i/2K} \cdot \left(\sum_{p^\alpha \leq n} |f(p^\alpha)|^{2K} p^{-\alpha} \right)^{r_i/2K} \leq (2K)! 16^K (2K - t + 1)^{2K+2t} n \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{t-1} \sum_{p^\alpha \leq n} |f(p^\alpha)|^{2K} p^{-\alpha}.$$

That is, for $t < 2K$

$$(2.3) \quad |M_{2K}(n, t)| \leq (2K)! 64^K K^{2K} (4K^2)^t n \cdot \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{2K-2} \sum_{p^\alpha \leq n} |f(p^\alpha)|^{2K} p^{-\alpha}.$$

Next we shall consider the case where $t = 2K$. To do this we shall first show that if p_X is the smallest prime in X then

$$(2.4) \quad |T(X, Y, n)| \leq K^2 4^K (|XY| p_X q_Y)^{-1} + 3^{2K+1} n^{-1}$$

when all the primes in XY are distinct. By Lemma 1 we see that

$$T(X, Y, n) = |XY|^{-1}R(X)R(Y) + O(n^{-1})$$

where

$$R(X) = \sum_{j=0}^K (-1)^j \binom{K}{j} \prod_{i=1}^j (1 - p_i^{-1})$$

and $|O(n^{-1})| \leq 3^{2K+1}n^{-1}$. Now induction shows

$$\prod_{i=1}^j (1 - p_i^{-1}) = 1 - \sum_{s=1}^j p_s^{-1} \prod_{i=s+1}^j (1 - p_i^{-1})$$

and hence $|R(X)| \leq K2^K p_X^{-1}$. A similar result holds for $R(Y)$, and hence we have (2.4). Therefore, keeping in mind all primes in XY are distinct, and using Lemma 2 and Hölder's inequality, we see

$$\begin{aligned} |M_{2K}(n, 2K)| &\leq n \sum_{|XY| \leq n} |F(X)F(Y)T(X, Y, n)| \\ &\quad + n \sum_{\substack{||X|| \leq n, ||Y|| \leq n \\ |XY| > n}} |F(X)F(Y)T(X, Y, n)| \\ &\leq nK^4 4^K \left(\sum_{||X|| \leq n} |F(X)| p_1^{-1} |X|^{-1} \right)^2 + 3^{2K+1} \sum_{|XY| \leq n} |F(X)F(Y)| \\ &\quad + 4^K n \sum_{\substack{||X|| \leq n, ||Y|| \leq n \\ |XY| > n}} |F(X)F(Y)| |XY|^{-1} \\ &\leq nK^4 4^K \left(\sum_{||X|| \leq n} |F(X)|^{2K} |X|^{-1} \right)^{1/K} \left(\sum_{||X|| \leq n} p_1^{-2K/(2K-1)} |X|^{-1} \right)^{2-1/K} \\ &\quad + 3^{2K+1} \left(\sum_{|XY| \leq n} |F(X)F(Y)|^2 |XY|^{-1} \right)^{1/2} \left((2K)! \sum_{j \leq n} j \right)^{1/2} \\ &\quad + n4^K \left(\sum_{\substack{||X|| \leq n, ||Y|| \leq n \\ |XY| > n}} |F(X)F(Y)|^2 |XY|^{-1} \right)^{1/2} \\ &\quad \cdot \left(\sum_{\substack{||X|| < n, ||Y|| \leq n \\ |XY| > n}} |XY|^{-1} \right)^{1/2} \\ &\leq nK^4 4^K \left(\sum_{p^\alpha \leq n} p^{-\alpha-1} \right)^2 \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{2K-3+1/K} \sum_{p^\alpha \leq n} |f(p^\alpha)|^{2K} p^{-\alpha} \\ &\quad + (3^{2K+1}(2K)! + 4^K C_1^{1/2} (2K)^2) n \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{K-1} \\ &\quad \cdot \left(\sum_{p^\alpha \leq n} |f(p^\alpha)|^2 p^{-\alpha} \right)^K \\ &\leq (K^4 4^{K+1} + 3^{2K+1}(2K)! + 4^{K+1} C_1^{1/2} K^2) \\ &\quad \cdot n \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{2K-2} \sum_{p^\alpha \leq n} |f(p^\alpha)|^{2K} p^{-\alpha} \\ &\leq C_5 (2K)! 9^K K^4 n \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{2K-2} \sum_{p^\alpha \leq n} |f(p^\alpha)|^{2K} p^{-\alpha} \end{aligned}$$

where $C_5 = 4 + C_1^{1/2}$.

Combining this last result with (2.3) we now have

$$|M_{2K}(n)| \leq (2K)! \left(C_5 9^K K^4 + 64^K K^{2K} \sum_{t=1}^{2K-1} (4K^2)^t \right) \cdot n \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{2K-2} \sum_{p^\alpha \leq n} |f(p^\alpha)|^{2K} p^{-\alpha}$$

which yields (2.1) for $C_2 = 1/3 + 9C_5/1024$. This finishes the proof.

3. **Odd central moments.** If we wish to consider odd central moments, then we must restrict ourselves to additive real valued arithmetic functions. Using the proof of the previous theorem it can be seen that this simplifies matters insofar as double summations become single summations. For example, for odd K and such functions (2.2) becomes

$$M_K(n) = \sum_{m \leq n} (f(m) - A(n))^K = n \sum_{\|X\| \leq n} F(X) T(X, n)$$

where

$$T(X, n) = \sum_{j=0}^K (-1)^j \binom{K}{j} |\tilde{X}_j|^{-1} n^{-1} \sum_{\substack{m \leq n \\ X_j | m}} 1.$$

If the rest of the proof of the theorem is carried out essentially as it is with minor modifications, it can be seen that for $t < K$

$$M_K(n, t) = O\left(n (\log \log n)^{K-2} \sum_{p^\alpha \leq n} |f(p^\alpha)|^K p^{-\alpha} \right)$$

as before, and

$$(3.1) \quad |M_K(n, K)| \leq n K^2 2^K \sum_{\|X\| \leq n} |F(X)| p_1^{-1} |X|^{-1} + O\left(n (\log \log n)^{K-2} \sum_{p^\alpha \leq n} |f(p^\alpha)|^K p^{-\alpha} \right).$$

Now Hölder's inequality shows that

$$\begin{aligned} \sum_{\|X\| \leq n} |F(X)| p_1^{-1} |X|^{-1} &\leq \left(\sum_{\|X\| \leq n} |F(X)|^K |X|^{-1} \right)^{1/K} \\ &\quad \cdot \left(\sum_{\|X\| \leq n} p_1^{-K/(K-1)} |X|^{-1} \right)^{1-1/K} \\ &\leq 1.3 \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{K-2+1/K} \sum_{p^\alpha \leq n} |f(p^\alpha)|^K p^{-\alpha} \end{aligned}$$

since $\sum p^{-\alpha-1} \leq 1.3$. Hence we have:

THEOREM 2. *If f is an additive real valued function and K is an odd integer, then*

$$\sum_{m \leq n} (f(m) - A(n))^K = B_K(n)n \left(\sum_{p^\alpha \leq n} p^{-\alpha} \right)^{K-2+1/K} \sum_{p^\alpha \leq n} |f(p^\alpha)|^K p^{-\alpha}$$

where $\overline{\lim} B_K(n) \leq 1.3K^2 2^K$.

This increases the exponent of $\sum p^{-\alpha}$ by $1/K$ relative to Theorem 1, but in general it cannot be avoided as the following argument shows. It is known [3; p. 201] that

$$\sum_{p \leq n} g(p) \sim \int_9^n g(x)(\log x)^{-1} dx$$

provided $g(x)/\log x$ for $x \geq 9$ is positive, nonincreasing, and has the limit 0 as $x \rightarrow \infty$,

$$\int_9^\infty g(x)(\log x)^{-1} dx \text{ diverges,}$$

and

$$\int_9^\infty g(x)(\log x)^{-1} e^{-(\log x)^{1.14}} dx \text{ converges.}$$

These conditions are satisfied by $g_1(p) = p^{-1} |\log \log p|^{-1/K}$ and $g_2(p) = p^{-1} |\log \log p|^{-1}$. Hence, for $f(p) = (\log \log p)^{-1/K}$ and $f(p^\alpha) = 0$ for $\alpha > 1$, we see that

$$\sum_{\|X\| \leq n} |F(X)| p^{-1} |X|^{-1} \geq C_6 \left(\sum_{p \leq n} g_1(p) \right)^{K-1} \sim C_6 \left(\frac{K}{K-1} \right)^{K-1} (\log \log n)^{K-2+1/K}$$

and

$$\sum_{p^\alpha \leq n} |f(p^\alpha)|^K p^{-\alpha} = \sum_{p \leq n} g_2(p) \sim \log \log \log n.$$

In the light of (3.1) this shows the desired result.

REFERENCES

1. J. Collison, *A variance property for arithmetic functions*, Pacific J. Math., **63** (1976), 347-355.
2. J. Kubilius, *Probabilistic Methods in the Theory of Numbers*, Translations of Mathematical Monographs, Amer. Math. Soc. Vol. 11, Providence, R. I., 1964.
3. E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Vol. 1, Leipzig: Teubner Verlagsgesellschaft, 1909. Reprinted, Chelsea Publishing Company, New York, 1974.

Received May 19, 1977.

BARUCH COLLEGE, CITY UNIVERSITY OF NEW YORK
NEW YORK, NY 10010