Докл. Акад. Наук СССР
Том 298 (1988), № 6

Soviet Math. Dokl.
Vol. 37 (1988), No. 1

# AN INVARIANCE PRINCIPLE FOR ADDITIVE ARITHMETIC FUNCTIONS

UDC 519.216.24

E. MANSTAVICHYUS [MANSTAVIČIUS]

Consider a sequence of probability spaces $\{\Omega_n, \mathscr{F}_n, \nu_n\}$, where $\Omega_n = \{1, \dots, n\}$, $\mathscr{F}_n = 2^{\Omega_n}$, and $\nu_n(\cdots) = n^{-1}\#\{m \le n, \cdots\}$ is the frequency of natural numbers $m \le n$ which satisfy whatever condition is substituted for the three dots. We define a sequence of arithmetic processes

$$H_n = H_n(m,t) = \sum_{\substack{p^\alpha || m \\ p \le z_n(t)}} h_n(p^\alpha) - \alpha_n(t), \qquad m \in \Omega_n,\ t \in [0,T]$$

where $h_n(p^\alpha) \in \mathbf{R}$, $p$ is a prime number, $\alpha \in \mathbf{N}$, and $p^\alpha || m$ means that $p^\alpha | m$ ($p^\alpha$ divides $m$), but $p^{\alpha+1} \nmid m$. Let $\alpha_n(\cdot) \in \mathbf{D}[0,T]$ and let $z_n(\cdot)\colon [0,T] \to \{1, \dots, n\}$ be a monotonically increasing transformation. We will assume, without mentioning it again, that $z_n([0,T]) = \{1 = k_{n1} < \cdots < k_{nj_n} = n\}$, $\max_{1 \le j \le j_n}(k_{n,j+1} - k_{nj}) = o(n^\varepsilon)$ for any $\varepsilon > 0$, and $\max_{1 \le j \le j_n} \operatorname{meas} z_n^{-1}(k_{nj}) = o(1)$. Here and in what follows the passage to the limit as $n \to \infty$ is not explicitly indicated.

The aim of this paper is to investigate weak convergence of the measures $\nu_n \circ H_n^{-1}$, corresponding to the processes $H_n$, in the space $\mathbf{D}[0,T]$ with the Skorokhod topology. The theorem given here establishes a weak invariance principle for dependent random variables $h_n^{(p)}(m)$, $p \le n$, where $\nu_n(h_n^{(p)}(m) = h_n(p^\alpha)) = n^{-1}([np^{-\alpha}] - [np^{-\alpha-1}])$. Here $[u]$ is the integer part of the number $u$. It is known [1] that a relatively weak dependence of the quantities $h_n^{(p)}(m)$, for $p \le r$, $\ln r = o(\ln n)$, increases with the growth of $p$. However, as it was shown in that paper, in the case of limiting processes with independent increments, the influence of "large" primes is eliminated. This effect, for particular choices of times and $h_n(p^\alpha) = h(p^\alpha)/\beta(n)$, where $\beta(n) \to \infty$, was observed in [2]–[5]. One can also find there bibliographical references to the history of the problem. Passage to the scheme of series extends the class of the limiting processes. The method of proof becomes considerably more complicated.

In what follows, $X = X(t)$ is a stochastically continuous process, given on some probability space $\{\Omega, \mathscr{F}, \mathbf{P}\}$, with independent increments and trajectories from $\mathbf{D}[0,T]$. We will write its characteristic function in the form

$$E\exp\{i\lambda X(t)\} = \exp\left\{ i\lambda\gamma(t) + \int_{-\infty}^{\infty} (e^{i\lambda u} - 1 - i\lambda u^*) u^{*^{-2}}\, d\psi_t(u) \right\},$$

where

$$u^* = \begin{cases} u, & \text{if } |u| < 1, \\ \operatorname{sgn} u, & \text{if } |u| \ge 1. \end{cases}$$

Here $\gamma(t)$ is continuous in $t$, while $\psi_t(u)$ is a bounded function, continuous in $t$ and nondecreasing in $t$ and $u$, so that for fixed $0 \le \tau < t \le T$ the difference $\psi_t(u) - \psi_\tau(u)$ is also nondecreasing.

---

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 11K65; Secondary 11Q15.
This translation incorporates corrections submitted by the author.

Put

$$A_n(t) = \sum_{p \le z_n(t)} \frac{h_n^*(p)}{p},$$

$$B(u,n) = \sum_{p \le u} \frac{h_n^{*^2}(p)}{p},$$

$$\psi_t^n(u) = \sum_{\substack{p \le z_n(t) \\ h_n(p) < u}} \frac{h_n^{*^2}(p)}{p}.$$

The symbols $\Rightarrow$ and $\rightrightarrows$ will denote the weak convergence of measures and the convergence of completely nondecreasing functions, respectively.

THEOREM. *Suppose that $h_n(p^\alpha) = o(1)$ for any fixed $p^\alpha$. Then for $\nu_n \circ H_n^{-1} \Rightarrow P \circ X^{-1}$ it is necessary and sufficient that the following conditions hold:*
(I) $B(n,n) - B(n^\varepsilon, n) = o(1)$ *for any* $\varepsilon$, $0 < \varepsilon < 1$;
(II) $\psi_t^n(u) \rightrightarrows \psi_t(u)$,
(III) $\alpha_n(t) = A_n(t) + \gamma(t) + o(1)$
*for any* $t$, $0 \le t \le T$.

The method of proving the sufficiency is generally well known (see, for example, [4]). We remark here that with the help of Kubilius's method and using condition (I) we can pass to the processes

$$\sum_{p \le z_n(t)} \xi_{np} - \alpha_n(t),$$

where the $\xi_{np}$, $p \le n$, are independent random variables, $\mathbf{P}(\xi_{np} = h_n(p)) = 1/p$, $\mathbf{P}(\xi_{np} = 0) = 1 - 1/p$, and then use Prokhorov's theorem ([6], Theorem 3.2).

*Necessity.* The main difficulty is in deriving condition (I). For $0 < c \le 1$ we set $t_n(c) = \sup\{t \in [0,T];\ z_n(t) \le n^c\}$. Let $i_{-1} = 1$, $i_0 = \frac{3}{4}$, $i_1 = \frac{1}{2}$, $i_2 = \frac{5}{12}$, and $i_k = 1/k$ when $k \ge 3$, and let $J_k = (i_k, i_{k+1}]$. From the conditions imposed on $z_n(t)$ we obtain $z_n(t_n(c)) \ge n^{i_{k+1}}$ if $c \in J_k$ and $n$ is sufficiently large. We will show by induction that

$$B(n,n) - B(n^c, n) = o(1), \tag{1}$$

when $c \in J_k$, $k \ge -1$. As in [4], the condition $h_n(p^\alpha) = o(1)$ allows us to pass to the case $h_n(p^\alpha) = h(p)$ for an arbitrary prime number $p$ and $\alpha \ge 1$. We single out the principal term of the characteristic function $\varphi_{\tau t}^n(\lambda)$ of the quantity $H_n(m,t) - H_n(m,\tau)$, when $\tau < t \le T$ and $z_n(\tau) \ge n^{i_{k+1}}$. Letting

$$f_n(m) = \exp\left\{ i\lambda \sum_{\substack{p|m \\ z_n(\tau) < p \le z_n(t)}} h_n(p) \right\} = \sum_{d|m} g_n(d),$$

we obtain

$$\varphi_{\tau t}^n(\lambda) = \exp\{-i\lambda(\alpha_n(t) - \alpha_n(\tau))\} \frac{1}{n} \sum_{d \le n} g_n(d) \left[\frac{n}{d}\right]. \tag{2}$$

For a strongly multiplicative function $f_n(m)$, the multiplicative function $g_n(d)$ has the following properties: $g_n(1) = 1$, $g_n(p) = f_n(p) - 1$, and $g_n(p^\alpha) = 0$ if $\alpha \ge 2$. Thus, in the sum in (2), it is enough to consider only square-free $d \le n$. Moreover, $g_n(p) = 0$ when $p \notin (z_n(\tau), z_n(t)] = \mathscr{P}_{\tau t}$, and the nonzero terms correspond only to $d = 1$ and $d = p_1 \cdots p_q$, where $p_1 < \cdots < p_q$ and $p_i \in \mathscr{P}_{\tau t}$, $1 \le i \le q$, and $q$ runs through the values $1, \dots, s$. The maximal number of prime factors is determined by the inequalities

$p_1 \cdots p_s \le n$ and $p_1 > n^{i_{k+1}}$; thus $s = s(k) < i_{k+1}^{-1}$. We have $s(-1) = s(0) = 1$, $s(1) = 2$, and $s(k) = k$ for $k \ge 2$, and also another form of (2):

$$\begin{aligned}(3)\qquad \varphi^n_{\tau t}(\lambda) &= \exp\{-i\lambda(\alpha_n(t) - \alpha_n(\tau))\} \\ &\times \left(1 + \sum_{p \in \mathscr{P}_{\tau t}} \frac{g_n(p)}{p} + \sum_{\substack{p_1 < p_2,\ p_1 p_2 \le n \\ p_1, p_2 \in \mathscr{P}_{\tau t}}} \frac{g_n(p_1) g_n(p_2)}{p_1 p_2} + \cdots \right. \\ &\qquad \left. + \sum_{\substack{p_1 < \cdots < p_s, p_1 \cdots p_s \le n \\ p_1, \dots, p_s \in \mathscr{P}_{\tau t}}} \frac{g_n(p_1) \cdots g_n(p_s)}{p_1 \cdots p_s} \right) \\ &+ O\left( \frac{1}{n} \sum_{q=1}^{s} 2^q \sum_{\substack{p_1 \cdots p_q \le n \\ p_1, \dots, p_q > n^{i_{k+1}}}} 1 \right).\end{aligned}$$

From the prime number theorem, it is easy to obtain a bound $o(1)$ for the last term.

For brevity, we leave the cases $k = -1, 0, 1, 2$ to the reader. The main ideas can be seen in the following inductive step of the proof. Suppose estimate (1) has been proved when $c \in \bigcup_{-1}^{k-1} J_j$. The relation $B(n^{c+o(1)}, n) - B(n^c, n) = o(1)$ allows us to include also the point $c = i_k$. Suppose further that $c \in J_k$ and $k \ge 3$.

Because of the inductive hypothesis, we may pass to the case in which $h_n(p) = 0$ for $n^{1/k} < p \le n$. Indeed, if for any $\alpha \ge 1$ we put

$$b_n(p^\alpha) = \begin{cases} h_n(p), & \text{when } p \le n^{1/k}, \\ 0, & \text{when } p > n^{1/k} \end{cases}$$

and denote by $G_n(m, t)$ the arithmetic process obtained from $H_n$ by replacing $h_n(p^\alpha)$ with $b_n(p^\alpha)$, then it follows from (1) that, for any $\delta > 0$ and sufficiently large $n$, we have

$$\max_{n^{1/n} \le u \le n} \left| \sum_{n^{1/k} < p \le u} \frac{h_n^*(p)}{p} \right| = O((B(n,n) - B(n^{1/k}, n))^{1/2}) \le \frac{\delta}{2},$$

$$\begin{aligned}&\nu_n \left( \sup_t |H_n(m,t) - G_n(m,t)| \ge \delta \right) \\ &\le \nu_n \left( \max_{n^{1/k} < u \le n} \left| \sum_{\substack{p \mid m \\ n^{1/k} < p \le u}} h_n(p) - \sum_{n^{1/k} < p \le u} \frac{h_n^*(p)}{p} \right| \ge \frac{\delta}{2} \right) + o(1) = o(1).\end{aligned}$$

The last estimate is obtained from an analogue of the Kolmogorov inequality (see [4], inequality (11)). Thus, using the convergence $\nu_n \circ G_n^{-1} \Rightarrow P \circ X^{-1}$, we could consider the values of $b_n(p^\alpha)$.

Without changing the notation, and supposing that $h_n(p) = 0$ for $p > n^{1/k}$, we simplify (3). Since $g_n(p) = 0$ for $p > n^{1/k}$, then, by considering $p \le n^{1/n}$, the conditions $p_1 \cdots p_q \le n$, $q = 2, \dots, k$, can be omitted from the sums. Using the symmetry of the terms for $\tau < t \le T$ when $z_n(\tau) \ge n^{1/k+1}$, we obtain

$$(4)\qquad \varphi^n_{\tau t}(\lambda) = \exp\{-i\lambda(\alpha_n(t) - \alpha_n(\tau))\} \left( 1 + \sum_{q=1}^{k} (\sigma^n_{\tau t}(\lambda))^q / q! \right) + o(1)$$

uniformly in $\lambda \in \mathbf{R}$. Here

$$\sigma_{\tau t}^n(\lambda) = \sum_{p \in \mathscr{P}_{\tau t}} \frac{\exp\{i\lambda h_n(p)\} - 1}{p}.$$

Well-known properties of prime numbers give

$$(5) \qquad \begin{aligned} |\sigma_{\tau t}^n(\lambda)| &\le 2 \sum_{n^{1/k+1} < p \le n^{1/k}} p^{-1} \le 2\ln\left(1 + \frac{1}{k}\right) + o(1), \\ \sum_{q=1}^{k} \frac{|\sigma_{\tau t}^n(\lambda)|^q}{q!} &\le (1 + o(1))\left(\left(1 + \frac{1}{k}\right)^2 - 1\right) \le \frac{8}{9}, \end{aligned}$$

if $n$ is sufficiently large and $k \ge 3$ is fixed.

We now verify that the convergence $t_n(c) \to t' < T$ along some subsequence $n = n' \to \infty$ (we omit the primes from now on) implies the equality

$$(6) \qquad \psi_T(+\infty) - \psi_t(-\infty) = 0$$

for any $t, t' \le t \le T$. Supposing the contrary, because of the continuity and monotonicity in $t$ of the function $\psi_t(+\infty)$, we select fixed moments of time $t' = t_1 < \cdots < t_{2k+1} = T$ so that $\psi_{t_l}(+\infty) - \psi_{t_j}(+\infty) > 0$, $1 \le j < l \le 2k+1$. We have altogether $K = k(2k+1)$ such inequalities. Using the property $z_n(t_1) \ge n^{1/k+1}$, convergence of the characteristic functions of the increments, and the asymptotic independence of these increments, we obtain from (4)

$$(7) \qquad Q(\sigma_{t_1 t_j}^n(\lambda)) Q(\sigma_{t_j T}^n(\lambda)) = Q(\sigma_{t_1 T}^n(\lambda)) + o(1)$$

uniformly in $|\lambda| \le M$, for any $M > 0$ and $j = 2, \ldots, 2k+1$. Here and in what follows $Q(z) = 1 + z + z^2/2 + \cdots + z^k/k!$. Moreover, if $\tau < t$ is an arbitrary pair of points $t_j$, $1 \le j \le 2k+1$, and $\kappa_{\tau t}(\lambda) = E\exp\{i\lambda(X(t) - X(\tau))\}$, then

$$(8) \qquad \exp\{-i\lambda(\alpha_n(t) - \alpha_n(\tau))\} Q(\sigma_{\tau t}^n(\lambda)) = \kappa_{\tau t}(\lambda) + o(1)$$

for $|\lambda| \le M$. Because of (5) we can take logarithms in a nontrivial neighborhood of $\lambda = 0$ and ascertain that $\alpha_n(t) - \alpha_n(\tau)$ is bounded. Further, let $n = n'' \to \infty$ be a subsequence of the sequence $\{n'\}$ for which $\alpha_n(t) - \alpha_n(\tau) = a_{\tau t} + o(1)$ simultaneously for any pair $\tau$ and $t$. Set $\kappa(\lambda) = \kappa_{\tau t}(\lambda)\exp\{i\lambda a_{\tau t}\}$.

Note that for the sequence $n = n'' \to \infty$, equality (8) implies the existence of the limit $\lim_{n\to\infty} \sigma_{\tau t}^n(\lambda)$ in an interval $|\lambda| \le \lambda_k$ in which $|1 - \kappa(\lambda)| < 1/k$. To establish this we have to verify that the polynomial $Q(z) - \kappa(\lambda)$ has only one root in the disk $|z| \le 2\ln(1 + 1/k)$. Since on the boundary of this disk

$$|1 - \kappa(\lambda) + z^2/2 + \cdots + z^k/k!| < (1 + 1/k)^2 - 1 - 2\ln(1 + 1/k) \le 2\ln(1 + 1/k) = |z|,$$

the uniqueness of this root follows from Rouché's theorem.

Thus, we have the limiting relations $\sigma_{t_j t_l}^n(\lambda) = \sigma_{jl}(\lambda) + o(1)$, when $n = n'' \to \infty$ for all $j$ and $l$, $1 \le j < l \le 2k+1$. For a given $\lambda \in [-\lambda_k, \lambda_k]$, there must be zeros among the limits $\sigma_{jl}(\lambda)$. If not, in view of (7) and the inequality $\operatorname{Re}\sigma_{jl} < 0$, we would have $2k+1$ different roots $0, \sigma_{12}(\lambda), \ldots, \sigma_{1,2k+1} = \sigma_{12}(\lambda) + \cdots + \sigma_{2k,2k+1}(\lambda)$ of the polynomial $Q(z)Q(\sigma_{1,2k+1}(\lambda) - z) - Q(\sigma_{1,2k+1}(\lambda))$ of degree at most $2k$. Moreover, the set $\Lambda_{jl} = \{\lambda \in [-\lambda_k, \lambda_k];\ \sigma_{jl}(\lambda) = 0\}$ for some pair of indices $j$, $l$ has its Lebesgue measure $\operatorname{meas}\Lambda_{jl} \ge 2\lambda_k/K$. Since $\Lambda_{jl} \pm \Lambda_{jl} \subset \Lambda_{jl}$, we conclude, on account of Steinhaus' lemma, that $\Lambda_{jl} = [-\lambda_k, \lambda_k]$. From this and from (8) it follows that $|\kappa_{t_j t_l}(\lambda)| \equiv 1$ for $|\lambda| \le \lambda_k$, which leads to $\psi_{t_l}(+\infty) - \psi_{t_j}(+\infty) = 0$. This contradiction proves (6).

If (6) holds, then $\kappa_{\tau T}(\lambda) = \exp\{i\lambda(\gamma(T) - \gamma(t))\}$, and on account of (5) it is easy to derive from (8) that $\alpha_n(T) - \alpha_n(t) = \gamma(T) - \gamma(t) + o(1)$, and further, $\sigma_{tT}^n(\lambda) = o(1)$

for any $|\lambda| \leq M$. This last result implies the estimate $o(1) = B(n,n) - B(z_n(t'),n) \geq B(n,n) - B(n^{c+\varepsilon},n)$ with an arbitrary $\varepsilon > 0$ and $c \in J_k$. Since $\varepsilon$ is arbitrary in case of (6), assertion (1) follows.

When $t_n(c) \to T$, it is enough to use a consequence of the denseness of the family of the measure $\nu_n \circ H_n^{-1}$:

$$\nu_n(|H_n(m,T) - H_n(m,t_n(c))| \geq \delta) = o(1),$$

where $\delta > 0$ is arbitrary, and equality (8) which follows from it, in which $\kappa_{\tau t}(\lambda) = 1$, $t = T$, and $\tau = t_n(c)$. Estimate (1) follows again, if $c \in J_k$.

From (1), with arbitrary $c \in J_k$ and $k \geq -1$, condition (I) follows. In what follows it is enough to repeat arguments used in the proof of Theorem 1 from [4]. Theorem is proved.

The author would like to thank J. Kubilius and the members of his seminar for very helpful remarks.

Vilnius State University

Received 31/JULY/86

BIBLIOGRAPHY

1. I. P. Kubilyus [J. P. Kubilius], *Probabilistic methods in the theory of numbers*, rev. ed., Gos. Izdat. Politich. i Nauchn. Lit. Litovsk. SSR, Vilnius, 1962; English transl., Amer. Math. Soc., Providence, R. I., 1964.

2. N. M. Timofeev and Kh. Kh. Usmanov, Dokl. Akad. Nauk Tadzhik. SSR **25** (1982), 207–211. (Russian)

3. ——, Dokl. Akad. Nauk Tadzhik. SSR **27** (1984), 556–559. (Russian)

4. È. Manstavichyus [E. Manstavičius], Litovsk. Mat. Sb. **24** (1984), no. 3, 148–161; English transl. in Lithuanian Math. J. **24** (1984).

5. ——, Litovsk. Mat. Sb. **25** (1985), no. 1, 97–109; English transl. in Lithuanian Math. J. **25** (1985).

6. Yu. V. Prokhorov, Teor. Veroyatnost. i Primenen. **1** (1956), 177–238; English transl. in Theory Probab. Appl. **1** (1956).

Translated by V. DROBOT