Z. Wahrscheinlichkeitstheorie verw. Geb. 25, 103–108 (1973)

Log Log Law for Gaussian Processes

Raoul D. LePage

If $X_1, X_2, \ldots$ are mutually independent and identically distributed Gaussian random variables with means zero and variances one (i.i.d. standard Gaussian r.v.) and for $n \geqq 1$ we let $S_n = X_1 + \cdots + X_n$, Hartman and Wintner's law of the iterated logarithm (refer to [5]) applies to say that almost everywhere (a.e.) the log log($\ell\,\ell$) sequence $\{S_n/(2n\,\ell\,\ell\,n)^{\frac{1}{2}},\ n \geqq 3\}$ converges to the closed interval $[-1, 1]$ and clusters at every point thereof. Some explanation of these terms is in order. In a metric space, a sequence converges to a set if it eventually stays within every ε-neighborhood of that set, and clusters at a point if that point is the limit of at least one convergent subsequence.

The situation in higher dimensions is somewhat surprising. Suppose $k \geqq 1$ and $\{S_{1n}, n \geqq 3\}, \{S_{2n}, n \geqq 3\}, \ldots, \{S_{kn}, n \geqq 3\}$ are independent copies of the sequence $\{S_n,\ n \geqq 3\}$. That is, there is mutual independence between sequences and each sequence has the probability law of the sequence $\{S_n,\ n \geqq 3\}$. What are the convergence properties, this time in euclidean space of k dimensions (k-space), of the random vector sequence $\{(S_{1n}, \ldots, S_{kn})/(2n\,\ell\,\ell\,n)^{\frac{1}{2}}, n \geqq 3\}$? Inasmuch as each coordinate independently roams $[-1, 1]$, the set of cluster points a.e. cannot exceed the cube $[-1, 1]^k$. Interestingly enough,

Lemma $\mathscr{S}$. *The above vector sequence a.e. converges to the closed unit ball taken around the origin of k-space and clusters at every point of this ball.*

The conclusion of Lemma $\mathscr{S}$ may be put the following way: if $X_1, X_2, \ldots$ are i.i.d. Gaussian random k-vectors possessing mutually independent standard Gaussian coordinates then $\{(X_1 + \cdots + X_n)/(2n\,\ell\,\ell\,n)^{\frac{1}{2}},\ n \geqq 3\}$ a.e. tends to the unit ball and clusters at every point thereof.

Now suppose B with norm $\| \ \|$ is a real and separable Banach space on whose sigma algebra $\mathscr{B}$ of Borel subsets is situated a Gaussian measure μ. That is, μ is a probability measure on $\mathscr{B}$ and each $x^* \in B^*$ (the topological dual of B) induces a zero-mean Gaussian probability law on the real line. One example of such a structure is $B = C[0, 1]$ with μ = Wiener measure. If $X_1, X_2, \ldots$ are i.i.d. B-valued r.v. defined on a probability space $(\Omega, \mathscr{F}, P)$ and $P \circ X_n^{-1} = \mu$ on $\mathscr{B}$, for every $n \geqq 1$, what are the convergence properties of the sequence $\xi_n = (X_1 + \cdots + X_n)/(2n\,\ell\,\ell\,n)^{\frac{1}{2}},\ n \geqq 3$, in B? This will be shown to be directly analogous to the finite dimensional case. Using results obtained in [4] there is a coordinate system imposed by μ a.e. on B in terms of which $X_1, X_2, \ldots$ have mutually independent standard Gaussian coordinates. It will be proved that the sequence $\{\xi_n,\ n \geqq 3\}$ a.e. P tends in B-norm to the unit ball in this coordinate system and clusters at every point. What makes this result interesting is the fact that this coordinate system may be any complete and orthonormal (c.o.n.) sequence for the reproducing kernel Hilbert space determined by μ (see [4]) and that the aforementioned ball has coordinate-free description as the unit ball in the reproducing kernel Hilbert space.

Proof of Lemma $\mathscr{S}$. This is a corollary to Strassen's law of the iterated logarithm for k-dimensional Brownian motion [5]. Let $C_k[0,1]$ denote the space of continuous maps from $[0,1]$ to k-space, endowed with supremum norm using euclidean norm in k-space. By $C_k[0,\infty)$ we indicate the set of k-space valued maps defined and continuous on $[0,\infty)$. The k-dimensional Brownian motion considered by Strassen is a Gaussian process $\{\zeta_{(i)}(t)\colon t\geqq 0,\ 1\leqq i\leqq k\}$ with means zero and covariances $\operatorname{Cov}\{\zeta_{(i)}(s),\zeta_{(j)}(t)\}=\delta_{ij}\min\{s,t\}$ for $s,t\geqq 0$, $1\leqq i,j\leqq k$. Strassen has proved that the $C_k[0,\infty)$ version, which exists for this process, satisfies the following law of the iterated logarithm in $C_k[0,1]$: the sequence $\{\zeta_n,\ n\geqq 3\}$ defined by $\zeta_n(t)=(\zeta_{(1)}(nt),\ldots,\zeta_{(k)}(nt))/(2n\ell\ell n)^{\frac{1}{2}}$ for every $t\in[0,1]$, $n\geqq 3$, a.e. converges in $C_k[0,1]$ to a set K and clusters at every point of K. The set K is defined by: $f\in K$ if and only if $f=(f_1,\ldots,f_k)\in C_k[0,1]$, f is coordinatewise absolutely continuous with respect to Lebesgue measure, $f(0)=0$, and $\sum_1^k\int_0^1 \dot f_i^2\,dt\leqq 1$. Because the random sequence of k-space vectors $\{\zeta_n(1),\ n\geqq 3\}$ has precisely the probability law of $\{(S_{1n},\ldots,S_{kn})/(2n\ell\ell n)^{\frac{1}{2}},\ n\geqq 3\}$, and the mapping defined by $f\to f(1)$ is continuous from $C_k[0,1]$ to k-space, our lemma is proved if the set K goes into the unit ball under this map. If $f\in K$ then $\sum_1^k f_i^2(1)\leqq\sum_1^k\int_0^1 \dot f_i^2\,dt\leqq 1$. Conversely, if $\sum_1^k r_i^2\leqq 1$ then $(r_1,\ldots,r_k)$ is the image under evaluation at $t=1$ of $f\in K$ defined by $f_i(t)=r_i t$, $t\in[0,1]$, $1\leqq i\leqq k$. □

Lemma $\mathscr{L}$. *For $(B,\mathscr{B},\mu)$ as above, if $\{x_i^*,\ i\geqq 1\}\subset B^*$ is c.o.n. for the closure $\mathscr{L}$ of B^* in $L_2(B,\mathscr{B},\mu)$, then for each $i\geqq 1$, x_i defined by $\int_B x_i^*(x)\,x\,\mu(dx)$ exists as a B-convergent Bochner integral and $\{x_i,\ i\geqq 1\}$ are c.o.n. for a Hilbert space $H\subset B$ whose elements are B-convergent series of the form $x=\sum_1^\infty x_i^*(x)\,x_i$ with $\sum_1^\infty (x_i^*(x))^2<\infty$. On H the two norms are related by the inequality $\|x\|\leqq\|x\|_H\,\|\mu\|$ where $\|\mu\|^2=\int_B\|x\|^2\,\mu(dx)<\infty$. H may be identified as the reproducing kernel Hilbert space of the kernel defined by $R(x^*,y^*)=\int_B x^*(x)\,y^*(x)\,\mu(dx)$, $(x^*,y^*)\in B^*\times B^*$, provided elements of B are interpreted as functions on B^*. For a.e. $x\in B$, $\left\|x-\sum_1^k x_i^*(x)\,x_i\right\|\to 0$ as $k\to\infty$. The closure $\bar H$ of H in B is the topological support of μ.*

Proof. The proof of these results appears in [4] together with references to related papers by N. Jain and G. Kallianpur, and J. Kuelbs. □

Let $K=\left\{\sum_1^\infty r_i x_i\colon \sum_1^\infty r_i^2\leqq 1\right\}$ be the unit ball of H. We have frequent need of the map $x\to x^{(k)}=\sum_1^k x_i^*(x)\,x_i$ which for each $k\geqq 1$ is by Lemma $\mathscr{L}$ defined on B and continuous into H.

Lemma 0. *The set K is closed in B.*

Proof. Suppose $\{y_n,\ n\geqq 1\}\subset K$ and $y_n\to y$ in B-norm. For every $k>0$, $\|y^{(k)}\|_H^2=\sum_1^k (x_i^*(y))^2=\lim_n \sum_1^k (x_i^*(y_n))^2\leqq 1$. Therefore $y\in K$. $\square$

Lemma 1. *For every $n\geqq 3$, $k\geqq 1$, $\xi_n^{(k)}$ is well defined a.e. P, and a.e. $\{\xi_n^{(k)},\ n\geqq 3\}$ converges to $K^{(k)}=\left\{\sum_1^k r_i x_i\colon \sum_1^k r_i^2\leqq 1\right\}$ in the sense of H-norm and clusters at every point thereof.*

Proof. For every $n\geqq 3$, $A\in\mathscr{B}$, $P(\sqrt{2\ell\ell n}\,\xi_n\in A)=\mu(A)$. Therefore a.e. P, $\xi_n^{(k)}$ are well defined for $n\geqq 3$, $k\geqq 1$, and reside in H. The k-vector of coefficients defines a random sequence in k-space $\left\{\left(\sum_1^n x_1^*(X_i),\dots,\sum_1^n x_k^*(X_i)\right)\Big/(2n\ell\ell n)^{\frac{1}{2}},\ n\geqq 3\right\}$ which because of mutual independence of $X_1, X_2,\dots$ and orthonormality of $\{x_i^*,\ i\geqq 1\}$ has precisely the probability law of $\{(S_{1n},\dots,S_{kn})/(2n\ell\ell n)^{\frac{1}{2}},\ n\geqq 3\}$ of Lemma $\mathscr{S}$. $\square$

In what follows, the reader will find our notation facilitates comparison with [5].

Lemma 2. *For each $\varepsilon>0$, $h\in H$, $x\in B$, $r_0>1$, the statements $(r_0-1)\,\|\mu\|<\varepsilon/2$, $\frac{1}{r_0}h\in K$, $\left\|\frac{1}{r_0}h-x\right\|\geqq\varepsilon$ together imply $\|h-x\|\geqq\varepsilon/2$.*

Proof.
$$\|h-x\|=\left\|\frac{1}{r_0}h-x+\frac{r_0-1}{r_0}h\right\|\geqq\left\|\frac{1}{r_0}h-x\right\|-(r_0-1)\left\|\frac{1}{r_0}h\right\|.$$

If $\left\|\frac{1}{r_0}h-x\right\|\geqq\varepsilon$ the right side is $\geqq\varepsilon-(r_0-1)\left\|\frac{1}{r_0}h\right\|$. By Lemma $\mathscr{L}$,
$$\left\|\frac{1}{r_0}h\right\|\leqq\left\|\frac{1}{r_0}h\right\|_H\|\mu\|\leqq\|\mu\|\qquad\text{if } \frac{1}{r_0}h\in K.$$
Hence the right side is $\geqq\varepsilon-(r_0-1)\,\|\mu\|$. $\square$

For convenience let X denote X_1. Henceforth E will indicate "expectation", i.e. integration over Ω with respect to P.

Lemma 3. *For each $\varepsilon>0$, $r>1$, there exists k sufficiently large so that $P(\|X-X^{(k)}\|\geqq\varepsilon\sqrt{2\ell\ell n})\leqq e^{-r^2\ell\ell n}$ for all sufficiently large n.*

Proof. If $\gamma>0$ then for each $k\geqq 1$, $n\geqq 3$,
$$P(\|X-X^{(k)}\|\geqq\varepsilon\sqrt{2\ell\ell n})\leqq E\,e^{\gamma\|X-X^{(k)}\|^2}\,e^{-2\gamma\varepsilon^2\ell\ell n}.$$
By a result of Fernique [2], $E\,e^{\gamma\|X-X^{(k)}\|^2}<\infty$ for at least those γ such that a $t>0$ may be found for which $\gamma<(24t^2)^{-1}\log(P(\|X-X^{(k)}\|\leqq t)/P(\|X-X^{(k)}\|>t))$. Since by Lemma $\mathscr{L}$, $\|X-X^{(k)}\|\to 0$ a.e. P as $k\to\infty$, we may choose any $t>0$, $r>1$, and for k sufficiently large be assured of $\gamma>r^2/2\varepsilon^2$ satisfying Fernique's inequality. For such a choice of k and γ, $E\,e^{\gamma\|X-X^{(k)}\|^2}\,e^{-2\gamma\varepsilon^2\ell\ell n}<e^{-r^2\ell\ell n}$ for all sufficiently large n.

Lemma 4. *For each* $k \geqq 1$, $r_0 > 1$, *and* $r^2 < r_0^2$, $P\left(\frac{1}{r_0} X^{(k)}/\sqrt{2\ell\ell n} \notin K\right) \leqq e^{-r^2 \ell\ell n}$ *for all sufficiently large n.*

Proof. For $k \geqq 1$, $r > 1$, $n \geqq 3$, $P\left(\frac{1}{r_0} X^{(k)}/\sqrt{2\ell\ell n} \notin K\right) = P(\chi_k^2 > 2 r_0^2 \ell\ell n)$. The latter is asymptotically $(r_0^2 \ell\ell n)^{k-1} e^{-r_0^2 \ell\ell n}/(k-1)!$ as $n \to \infty$. For $r < r_0$ this is less than $e^{-r^2 \ell\ell n}$ for all sufficiently large n. □

For $\varepsilon > 0$ let $K_\varepsilon = \{x: x \in B, \|x - K\| < \varepsilon\}$, the open ε-neighborhood of the closed set K in B. Write $r \sim 1$ for "r sufficiently close to one".

Proposition 1. *For* $\varepsilon > 0$, $r \sim 1$, $P(X/\sqrt{2\ell\ell n} \notin K_\varepsilon) \leqq e^{-r^2 \ell\ell n}$ *for all sufficiently large n.*

Proof. Suppose $\varepsilon > 0$, $r > 1$, $k \geqq 1$, $n \geqq 3$. For $r_0 > r$,

$$P(X/\sqrt{2\ell\ell n} \notin K_\varepsilon) \leqq P\left(\frac{1}{r_0} X^{(k)}/\sqrt{2\ell\ell n} \notin K\right) + P\left(\frac{1}{r_0} X^{(k)}/\sqrt{2\ell\ell n} \in K, \left\|\frac{1}{r_0} X^{(k)} - X\right\| \geqq \varepsilon \sqrt{2\ell\ell n}\right). \tag{1}$$

By Lemma 2, since $X^{(k)} \in H$, the second term on the right of (1) is for $r_0 \sim 1$

$$\leqq P\left(\|X - X^{(k)}\| \geqq \frac{\varepsilon}{2} \sqrt{2\ell\ell n}\right).$$

By Lemma 3 we may choose k sufficiently large so the latter is $\leqq e^{-2r^2 \ell\ell n} \leqq \frac{1}{2} e^{-r^2 \ell\ell n}$ for all sufficiently large n. For this choice of k, by Lemma 4, the first term on the right side of (1) is $\leqq e^{-(\frac{1}{2}r + \frac{1}{2}r_0)^2 \ell\ell n} \leqq \frac{1}{2} e^{-r^2 \ell\ell n}$ for all sufficiently large n. □

As used below, square bracket [ ] will indicate the taking of integer part.

Proposition 2. *For* $\varepsilon > 0$, $r > 1$, *it is possible to choose* $c > 1$ *sufficiently close to one so that* $P(\max\{\|\xi_i - \xi_{[c^n]}\| : [c^n] \leqq i < [c^{n+1}]\} \geqq \varepsilon) \leqq e^{-r^2 \ell\ell [c^n]}$ *for all sufficiently large n.*

Proof. For $\varepsilon > 0$, $r > 1$, $c > 1$, $[c^n] \geqq 3$, $[c^n] < [c^{n+1}]$, let events U, V, W (depending on ε, c, n) be defined through the following relations with $I = \{i: [c^n] \leqq i < [c^{n+1}]\}$:

$$U = \{\max_{i \in I} \|\xi_i - \xi_{[c^n]}\| \geqq \varepsilon\},$$

$$V = \left\{\max_{i \in I} \|X_{[c^n]} + \dots + X_i\| \geqq \frac{\varepsilon}{2} (2[c^n] \ell\ell [c^n])^{\frac{1}{2}}\right\},$$

$$W = \{\max_{i \in I} |(i \ell\ell i/[c^n] \ell\ell [c^n])^{\frac{1}{2}} - 1| \, \|\xi_{[c^n]}\| > \varepsilon/2\}.$$

Then $P(U) \leqq P(V) + P(W)$. For all sufficiently large n, W is contained in the event $\xi_{[c^n]} \notin K_\varepsilon$, as K_ε is norm bounded in B. By Proposition 1, for $r \sim 1$, $P(\xi_{[c^n]} \notin K_\varepsilon) = P(X/\sqrt{2\ell\ell [c^n]} \notin K_\varepsilon) \leqq \frac{1}{2} e^{-r^2 \ell\ell [c^n]}$ for all sufficiently large n. To bound $P(V)$ we need the inequality, $P(\max_{1 \leqq j \leqq J} \|X_1 + \dots + X_j\| \geqq a) \leqq 2P(\|X\| \geqq a J^{-\frac{1}{2}} - \|\mu\| \sqrt{2})$ valid for all $J \geqq 1$, $a > 0$. The proof closely follows that of the specialization to the Gaussian case of a well-known r.v. inequality. For example, see [3], Lemma 2,

p. 45, where there is also a nice historical sketch of the development of the ordinary log log law for r.v. The proof strongly uses two things: mutual independence of $X_1, X_2, \ldots$ and the fact that for each $n \geqq 1$ the probability law of $(X_1 + \cdots + X_n)/\sqrt{n}$ is precisely that of X. Applying this inequality,

$$P(V) \leqq 2P\left(\|X\| \geqq \frac{\varepsilon}{2}(2\ell\,\ell[c^n]/(c^2-1))^{\frac{1}{2}} - \|\mu\|\sqrt{2}\right)$$
$$\leqq 2P\left(\|X\| \geqq \frac{\varepsilon}{2}(\ell\,\ell[c^n]/(c^2-1))^{\frac{1}{2}}\right)$$

for all sufficiently large n, using $([c^{n+1}]-[c^n])/[c^n] \leqq c^2-1$ for all sufficiently large n. By Fernique's result, there exists $\gamma > 0$ for which $E\,e^{\gamma\|X\|^2} < \infty$. For this γ, our last upper bound on $P(V)$ is $\leqq 2E\,e^{\gamma\|X\|^2}\,e^{-(\varepsilon^2/4)\ell\ell[c^n]/(c^2-1)}$. For c so close to one that $\varepsilon^2 > 4r^2(c^2-1)$ the latter is $\leqq \frac{1}{2}e^{-r^2\ell\ell[c^n]}$ for all sufficiently large n. Combining this with the bound previously obtained for $P(W)$ gives the result. □

Theorem. *The sequence* $\{\xi_n, n \geqq 3\}$ *a.e. P converges in B to K and clusters at every point of K in the sense of B-norm.*

Proof. Suppose $\varepsilon > 0$, $r > 1$. For $c > 1$, $[c^n] \geqq 3$, $[c^n] < [c^{n+1}]$,

$$\begin{aligned} &P(i\colon [c^n] \leqq i < [c^{n+1}],\ \xi_i \notin K_\varepsilon) \\ &\quad \leqq P(\xi_{[c^n]} \notin K_{\varepsilon/2}) + P(\max\{\|\xi_i - \xi_{[c^n]}\| : [c^n] \leqq i < [c^{n+1}]\} \geqq \varepsilon/2) \\ &\quad \leqq e^{-r^2\ell\ell[c^n]} \end{aligned}$$

by Propositions 1 and 2, for r, c sufficiently close to one and all sufficiently large n. Since $\sum_n e^{-r^2\ell\ell[c^n]} < \infty$ we may apply the Borel-Cantelli lemma to conclude $P(\xi_i \in K_\varepsilon$ eventually as $i \to \infty) = 1$. Since this is true for all rational $\varepsilon > 0$, the sequence $\{\xi_i, i \geqq 3\}$ is a.e. convergent to K. By Lemma 0 the set K is closed in B, which together with the previous sentence implies the set of cluster points a.e. cannot exceed K. To prove the set of cluster points a.e. contains K it is enough, because of separability of K in B, to prove that for each $x \in K$, $\varepsilon > 0$, $P(\|\xi_i - x\| < \varepsilon$ i.o.$) = 1$. Here i.o. means infinitely often as $i \to \infty$. Suppose $\varepsilon > 0$, $x \in K$. Choose $r \sim 1$, k so that $\|x - x^{(k)}\| < \varepsilon/5$ and so that, by Lemma 3,

$$P\left(\|X - X^{(k)}\| \geqq \frac{\varepsilon}{5}\sqrt{2\ell\,\ell\,n}\right) \leqq e^{-r^2\ell\ell n}$$

for all sufficiently large n. For this k, by Proposition 2, choose $c > 1$ sufficiently close to one so that the conclusion of Proposition 2 holds for $\{\xi_i, i \geqq 3\}$ and for $\{\xi_i^{(k)}, i \geqq 3\}$. From the Borel-Cantelli lemma applied to these estimates, a.e. P, defining $n = n(i)$ by $[c^n] \leqq i < [c^n]$, for all sufficiently large i,

$$\|\xi_i - \xi_{[c^n]}\| < \frac{\varepsilon}{5}, \qquad \|\xi_{[c^n]} - \xi_{[c^n]}^{(k)}\| < \frac{\varepsilon}{5}, \qquad \|\xi_{[c^n]}^{(k)} - \xi_i^{(k)}\| < \frac{\varepsilon}{5}.$$

By Lemma 1, a.e. P,

$$\|\xi_i^{(k)} - x^{(k)}\| \leqq \|\xi_i^{(k)} - x^{(k)}\|_H\,\|\mu\| < \frac{\varepsilon}{5}$$

i. o. Therefore, a. e. P, i. o.

$$\|\xi_i - x\| \leqq \|\xi_i - \xi_{[c^n]}\| + \|\xi_{[c^n]} - \xi_{[c^n]}^{(k)}\| + \|\xi_{[c^n]}^{(k)} - \xi_i^{(k)}\| + \|\xi_i^{(k)} - x^{(k)}\| + \|x^{(k)} - x\| < \varepsilon.$$

This completes the proof. □

This theorem, as applied to $B = C_k[0, 1]$, $\mu = k$-dimensional Brownian motion measure, differs slightly from Strassen's theorem quoted in the proof of Lemma $\mathscr{S}$. Both $\{\xi_n, n \geqq 3\}$ and $\{\zeta_n, n \geqq 3\}$ are zero-mean Gaussian sequences in $C_k[0, 1]$. Their coordinatewise covariances are: $\min\{n, m\} \min\{s, t\}$ for the former, versus $\min\{n s, m t\}$ for the latter, for $n, m \geqq 3$, $0 \leqq s, t \leqq 1$. Does Strassen's result follow as a perturbation of our theorem? This question is presently under study.

I wish to thank V. Mandrekar and also J. Williamson for helpful discussions of Strassen's theorem. The paper has also benefitted from good refereeing.

References

1. Chover, J.: On Strassen's version of the log log law. Z. Wahrscheinlichkeitstheorie verw. Gebiete **8**, 83–90 (1967).
2. Fernique, X.: Integrabilité des Vecteurs Gaussiens. C. r. Acad. Sci., Paris, Sér. A **270**, 1698–1699 (1970).
3. Lamperti, J.: Probability. New York: Benjamin 1966.
4. LePage, R.: Note relating Bochner integrals and reproducing kernels to series expansions on a Gaussian Banach space. Proc. Amer. Math. Soc. **32**, 285–288 (1972).
5. Strassen, V.: An invariance principle for the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie verw. Gebiete **3**, 211–226 (1964).

Raoul D. LePage
Department of Statistics and Probability
Wells Hall
Michigan State University
East Lansing, Michigan 48823
USA

(Received March 3, 1972)