ON THE RANDOM WALK AND BROWNIAN MOTION

BY

FRANK B. KNIGHT

Introduction. Consider on the one hand a separable Brownian motion (Wiener process) \( X(t) \), \( 0 \leq t < \infty \), with \( X(0) = 0 \), and on the other a classical random walk \( S(n) = \sum_{i=1}^{n} X_i \), \( 1 \leq n < \infty \), where \( X_1, X_2, \ldots \) is a sequence of Bernoulli trials with probability \( 1/2 \) for \( X_i = +1 \) and for \( X_i = -1 \). It is well known that for the sequence of processes \( R_k(t) \), \( 1 \leq k < \infty \), defined by the formula

\[
R_k(t) = 2^{-k}S([2^{k}t]), \quad 0 \leq t < \infty,
\]

the joint distributions converge weakly in \( k \) to the corresponding joint distributions of \( X(t) \). It is also well known \([7]\) that the difference equation satisfied by the transition function of \( R_k(t) \) goes over into the diffusion equation satisfied by the transition function of Brownian motion.

Connections of the type illustrated here between the random walk and the Brownian motion involve the derivation of analytical properties of the latter as limits of analogous properties of the former, independently of the probability spaces on which the processes are defined. However, it is also possible to define, as in the present paper, approximating sequences of random walks together with a limit Brownian motion process on a single probability space, in such a way that the customary limit theorems appear as consequences of probability 1 convergence on this common space, and this method can have certain advantages. For example, when the type of convergence on this space is uniform convergence of path functions in finite time intervals with probability 1, as it is to be here, it follows that the random walks provide a constructive definition of the Brownian motion (that is, of its space of path functions, field of measurable sets, and probability measure).

This paper is derived from part of the author’s Ph.D. thesis, written with the help and guidance of Professor William Feller. Sincere thanks are rendered to him and to Professor H. Trotter for their assistance. The mappings \( M_k \), on which the paper is based, are due to them.

The first stage in the construction is to define a sequence of random walks \( R_k(t) \) of the type mentioned in the introduction for each \( k \), but on separate probability spaces, and to combine these spaces in a single over-all space. We shall use “a.s.” to abbreviate “almost surely” or “with probability one.”

Definition 1. Let \( \alpha_k = 2^{-2k} \), \( \beta_k = 2^{-k} \), where \( k \) is a non-negative integer. The set \( \Omega_k \) of \( R_k \)-paths is the set of all functions \( w_k \) from \( \{ m\alpha_k : m \text{ a non-negative integer} \} \) onto \( \{ n\beta_k : n \text{ an integer} \} \), such that \( w_k(0) = 0 \) and \( w_k((m+1)\alpha_k) \)

Received by the editors September 30, 1960 and, in revised form, June 16, 1961.
equals \( w_k(m\alpha_k) + \beta_k \) or \( w_k(m\alpha_k) - \beta_k \) for each \( k \geq 0 \).

The sets \( \Omega_k \) are related by means of mappings \( M_k \) from \( \Omega_k \) onto \( \Omega_{k-1} \), \( k \geq 1 \).

**Definition 2.** \( M_kw_k \) is defined as follows: it has the sequence of values obtained by first eliminating from the sequence \( w_k(0), w_k(\alpha_k), w_k(2\alpha_k), \cdots \) all terms except those of the form \( n\beta_{k-1} \), and then eliminating all of the remaining terms which are equal to their immediate predecessors.

It is clear that \( M_kw_k \) is always an \( R_{k-1} \)-path. An inverse to any \( R_{k-1} \)-path may easily be constructed, but this inverse is never unique.

We next define on \( \Omega_k \) a \( \sigma \)-field \( F_k \) of subsets, and a probability measure \( P_k \), with which \( \Omega_k \) may be interpreted as the space of elementary events (sample space) of a discrete random walk \( R_k \) by means of the definition \( R_k(n\beta_k) = w_k(n\beta_k), 0 \leq n < \infty \). The random walk \( R_k \) starts at 0 and has for all \( n \) the stationary one-step transition probabilities

\[
\begin{align*}
\rho(n\beta_k, (n-1)\beta_k) &= \rho(n\beta_k, (n+1)\beta_k) = \frac{1}{2}.
\end{align*}
\]

The underlying \( \sigma \)-field \( F_k \) is defined to consist of the sets \( S \cap \Omega_k \) for all \( S \) in the \( \sigma \)-field generated by the open sets in the product space

\[
\prod_{m=0}^{\infty} \{ n\beta_k, -m \leq n \leq m \}
\]

of countably many discrete, compact, spaces. The probability measure is defined on this field by iteration of the transition probabilities, beginning with the sets

\[
\begin{align*}
E &= \{ w_k : w_k(0) = 0, w_k(\alpha_k) = \epsilon_1\beta_k, w_k(2\alpha_k) = (\epsilon_1 + \epsilon_2)\beta_k, \cdots, w_k(m\alpha_k) \\
&= (\epsilon_1 + \epsilon_2 + \cdots + \epsilon_m)\beta_k \} \quad \text{where} \quad \epsilon_i = \pm 1,
\end{align*}
\]

for which \( P_k(E) = 2^{-m} \). The extension of \( P_k \) to the least \( \sigma \)-field containing the sets \( E \) now follows by a standard method. This \( \sigma \)-field is precisely \( F_k \), as proved below.

**Theorem 1.** The spaces \((\Omega_k, F_k, P_k)\) together with the mappings \( M_{k+1}M_{k+2} \cdots M_r \) from \( \Omega_r \) to \( \Omega_k, r > k \), are a stochastic process \((\Omega_\infty, F_\infty, P_\infty)\) in the sense of Bochner [1, p. 119].

Before stating the meaning of this theorem, we prove a lemma.

**Lemma 1.** For each \( k > 0 \), the total inverse mapping \( M_k^{-1} \) of \( M_k \) is measure preserving from \((\Omega_{k-1}, F_{k-1}, P_{k-1})\) onto \((\Omega_k, F_k, P_k)\). In other words, the mapping of \( R_k \) defined by \( M_kR_k(n\beta_k) = M_kw_k(n\beta_k) \) carries the process \( R_k \) onto the process \( R_{k-1} \).

**Proof.** We first of all note that for \( 0 \leq m < \infty \) and \(-m \leq n \leq m \) there holds the following identity of conditional probabilities:
This follows from (1) in an evident way. Now (3) implies, by the Markov property for \( R_k \), that with \( \varepsilon_1 \) as in (2) and \( k > 0 \),
\[
P_k \{ w_k : M_{k-1}(w) = (\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_k) \beta_{k-1} \} = 2^{-m}.
\]

The lemma thus holds for the sets \( E \subseteq F_{k-1} \) defined as in (2) with \( k - 1 \) in place of \( k \). It therefore holds for all \( E \subseteq F_{k-1} \) by the standard extension.

Recalling, now, that the “projective limit” \( \Omega_\infty \) of the spaces \( \Omega_k \) under the mappings \( M_k \) is the set of all sequences \( (w_0, w_1, \ldots, w_n, \ldots) \) such that \( M_kw_k = w_{k-1} \) for all \( k > 0 \), and that each set \( B_k \) in \( F_k \) is the “projection” onto \( \Omega_k \) of the set of all elements of \( \Omega_\infty \) whose \( k \)th components are in \( B_k \), the theorem means that the finitely additive set function \( P \) induced on \( \Omega_\infty \) by the projection inverses of all \( B_k \) in \( F_k \), with measures \( P_k \), \( 0 \leq k < \infty \) (which are consistent by the lemma) can be extended to a countably additive measure \( P_\infty \) on the smallest \( \sigma \)-field \( F_\infty \) containing the projection inverses of all \( B_k \). To prove the possibility of this extension, it is sufficient to show that the spaces \( \Omega_k \) are “topological” [Bochner, 1, Theorem 5.1.1] in the sense that any measurable set may be approximated in measure by compact subsets, and that the mappings \( M_k \) are continuous. The property of being approximated in measure by compact subsets is clearly preserved under the operations of countable set union and intersection, and it holds for the cylinder sets of \( \Omega_k \) generated by subsets of the coordinate spaces \( \{ n\beta_k, -m \leq n \leq m \} \). The open sets of \( \Omega_k \) are defined by intersection of \( \Omega_k \) with the open sets of the product space \( \prod_{-m}^m \{ n\beta_k, -m \leq n \leq m \} \). By considering the coordinates of a point in the product space as defining the position of a corresponding real number in \( (0, 1] \) by successive subdivision of \( (0, 1] \) into semiclosed subintervals of length \( 1, 1/2, 1/6, \ldots, 1/n!, \ldots \), it is seen from properties of the real numbers that all of these open sets may be obtained as countable unions of finite intersections from the subbase consisting of the cylinder sets generated by subsets of the coordinate spaces, i.e. as disjoint unions of semiclosed intervals open at the left. It is immaterial whether the operations are performed before or after the intersection with \( \Omega_k \). Thus the open sets are approximated in the required manner. The closed sets of \( \Omega_k \) are compact and therefore trivially approximated. Finally, to extend from the class of open and closed sets to the least \( \sigma \)-field containing it, only the set operations of countable union and intersection are involved. Thus the spaces (\( \Omega_k, F_k, P_k \)) are topological\(^{(1)} \).

\(^{(1)}\) It is clear, in fact, that \( F_k \) is the field of Baire sets, and hence that \( P_k \) is a regular measure.
mappings $M_k$ are continuous since the discrete character of the coordinate spaces makes convergence degenerate at each coordinate. This completes the requirements of [1, Theorem 5.1.1], and proves the theorem.

We next introduce a continuous time parameter in each random walk $R_k$ by setting $R_k(t) = R_k(mak)$ for $mak \leq t < (m+1)ak$, $m \geq 0$. It is to be shown that for points $(w_0, w_1, w_2, \ldots, w_n, \ldots) \in \Omega_\infty$ the corresponding sequences $(R_0(t), R_1(t), R_2(t), \ldots, R_n(t), \ldots)$ converge uniformly in $(0, T)$, for each $T > 0$, with $P_\omega$-probability 1. Accordingly, it is necessary to investigate the relation induced by the mapping $M_k$ between the time parameters of $R_k(t)$ and $R_{k-1}(t)$.

**Definition 3.** For each $m \geq 0$, $k > 0$, let $E_{k-1}(m)$ be the (random) least integer such that for $Mkwk = wk^{(1)}$ the ordered set
\[
\{wk(ak), wk(2ak), \ldots, wk(E_{k-1}(m)ak)\}
\]
contains the ordered set $\{wk^{(1)}(ak-1), wk^{(1)}(2ak-1), \ldots, wk^{(1)}(mak-1)\}$ as an ordered subset. Then the "time lag" $L_{k,r}(t)$, $k \leq r$, of the process $R_r(t)$ with respect to $R_k(t)$, is defined as follows:

\[
L_{k,k}(t) = 0;
\]
\[
L_{k,r}(mak) = (E_{r-1} \cdots E_{k+1}E_k(m))ak - mak, \quad k < r.
\]
\[
L_{k,r}(t) = L_{k,r}(mak), \quad mak \leq t < (m+1)ak.
\]

The following relation for $r > k + 1$ is immediate(4):

\begin{align*}
L_{k,r}(mak) &= L_{k,k+1}(mak) + L_{k+1,k+2}(E_k(m)ak+1) + \cdots \\
&\quad + L_{r-1,r}(E_{r-2}E_{r-3} \cdots E_k(m)ak) \\
&= L_{k,r-1}(mak) + L_{r-1,r}(mak + L_{k,r-1}(mak)).
\end{align*}

The main theorem concerning the time lag will now be proved.

**Theorem 2.** For each $\delta > 0$, we have

\[
P_\omega \left\{ \max_{0 \leq t \leq mak} \left| L_{k,r}(t) \right| > \delta \right\} < \frac{2mak^2}{3\delta^2}.
\]

**Proof.** We first note that $E(E_k(1)) = 4$ and that $\sigma^2(E_k(1)) = 8$ (where $E$ and $\sigma^2$ denote expectation and variance). In fact, $E_k(1)$ has the same distribution as the number of tosses of a "fair coin" required until $|\text{number of heads} - \text{number of tails}| = 2$, and therefore $E_k(1) = 2(X + 1)$ where $X$ has a geometric distribution with $p = 1/2$. Since $E(X) = 1$ and $\sigma^2(X) = 4$, the statement about $E_k(1)$ is easily verified. Consequently,

\[
E(L_{k,k+1}(ak)) = 0 \quad \text{and} \quad \sigma^2(L_{k,k+1}(ak)) = 2^{-ak-1}.
\]

Now $L_{k,k+1}(mak)$ is clearly the sum of $m$ independent variables each with

(4) The referee has kindly suggested the last expression.
the same distribution as \( L_{k,i+1}(\alpha_k) \). Together with (5) and (6) this implies that 
\[ E(L_{k,i+1}(\alpha_k)) = 0. \]
We next calculate the conditional variance of \( L_{k,i+1}(\alpha_k) - L_{k,i}(\alpha_k) \) given \( L_{k,i}(\alpha_k) \), \( k < i < r \). If \( L_{k,i}(\alpha_k) \) is given, we see from (5) that \( L_{k,i+1}(\alpha_k) - L_{k,i}(\alpha_k) \) is the sum of \( \sum_{i=1}^{r-k} \) independent variables distributed like \( L_{k,i+1}(\alpha_k) \), and from Definition 3 we have
\[ E_i (L_{k,i+1}(\alpha_k) - L_{k,i}(\alpha_k)) = \alpha_i^{-1}(L_{k,i}(\alpha_k) + \alpha_k). \]
Therefore by (6),
\[ E((L_{k,i+1}(\alpha_k) - L_{k,i}(\alpha_k))^2 | L_{k,i}(\alpha_k)) = \alpha_i^{-1}(L_{k,i}(\alpha_k) + \alpha_k)2^{-i-1} = (L_{k,i}(\alpha_k) + \alpha_k)2^{-i-1}. \]
On the other hand, this conditional variance is the same as that of \( L_{k,i+1}(\alpha_k) \) given \( L_{k,i}(\alpha_k) \), which is
\[ E(L_{k,i+1}(\alpha_k) | L_{k,i}(\alpha_k)) = E(L_{k,i+1}(\alpha_k) | L_{k,i}(\alpha_k)). \]
But the expected value of (8) is \( 2^{-2(k+1)} - 1 \), while that of (9) is \( E(L_{k,i+1}(\alpha_k)) - E(L_{k,i}(\alpha_k)) \). Equating these two expectations and adding from \( i = k \) to \( i = r - 1 \), we have
\[ E(\sum_{i=k}^{r-1} 2^{-2(k+i)-1} \) \[ = 2^{-4k-1} \sum_{i=0}^{r-k} 2^{-2i} \< \frac{2^{-4k+1}}{3}. \]
Theorem 2 now follows from (10) by an application of the Kolmogorov inequality.

We shall require an additional uniformity in \( r \), given by the following corollary.

**Corollary 1.** For fixed \( k, m, 1 \leq i \leq m/2, \) and \( \delta = 2i\alpha_k, \)
\[ P_{\infty} \{ \text{l.u.b.} \ | L_{k,r}(t) | > \delta \} < \frac{8m\alpha_k}{3\delta^2} \left( 1 - \frac{8(m + \delta\alpha_k)^{-1}}{3\delta^2} \right)^{-1}. \]

**Proof.** It is easily seen from (5) that for \( k < r \leq s \)
\[ L_{k,s}(m\alpha_k) = L_{k,r}(m\alpha_k) + L_{r,s}(m\alpha_k + L_{k,r}(m\alpha_k)). \]
Let \( L_{k,n}(T) \) denote \( \max_{0 \leq t \leq T} | L_{k,n}(t) | \), let \( k \) and \( s \) be fixed with \( k < s \), and for \( \delta = 2i\alpha_k \) define \( r(\delta) \) to be the least integer \( r, k < r \leq s \), such that \( L_{k,r}(m\alpha_k) \geq \delta \) when such exists, or \( k \) otherwise. Since for given \( t \) and \( t' \) the variables \( L_{k,r}(t) \) and \( L_{r,s}(t') \) are independent, (11) together with Theorem 2 implies that
\[ P_\alpha(\tilde{L}_{k,r}(s\alpha_k) > \delta/2 \mid \tilde{L}_{k,r}(s\alpha_k) \geq \delta) \]
\[ \geq 1 - P_\alpha(\tilde{L}_{r}(s\alpha_k + \tilde{L}_{k,r}(s\alpha_k)) > \delta/2 \mid \tilde{L}_{k,r}(s\alpha_k) \geq \delta) \]
\[ \sum_{j=2i\alpha_k/\alpha_\delta}^{\infty} P_\alpha(\tilde{L}_{r}(s\alpha_k + j\alpha_k) > j\alpha_k - i\alpha_k) P_\alpha(\tilde{L}_{k,r}(s\alpha_k) = j\alpha_k) \]
\[ = 1 - \frac{1}{P_\alpha(\tilde{L}_{k,r}(s\alpha_k) \geq 2i\alpha_k)} \]
\[ \sum_{j=2i\alpha_k/\alpha_\delta}^{\infty} \frac{2(m + j\alpha_k/\alpha_\delta)}{3(j\alpha_k - i\alpha_k)^2} \]
\[ \frac{P_\alpha(\tilde{L}_{k,r}(s\alpha_k) = j\alpha_k)}{P_\alpha(\tilde{L}_{k,r}(s\alpha_k) \geq 2i\alpha_k)} \]
\[ \geq 1 - \frac{2(m + 2i\alpha_k)}{3(i\alpha_k)^2} \]
\[ = \frac{1 - 8(m + \delta\alpha_k^{-1})\alpha_k^2}{3\delta^2}, \]

since the bound in Theorem 2 is monotone in \( k \), and the first term in the second summation is maximized when \( j = 2i\alpha_k/\alpha_\delta \). From (12) and Theorem 2 it follows that for \( \delta = 2i\alpha_k \)

\[ P_\alpha \left\{ \max_{1 \leq r \leq k, 0 \leq t \leq s\alpha_k} \left| L_{k,r}(t) \right| > \delta \right\} \leq \frac{8m\alpha_k^2}{3\delta^2} \left( 1 - \frac{8(m + \delta\alpha_k^{-1})\alpha_k^2}{3\delta^2} \right)^{-1}. \]

Letting \( s \) become large, the corollary is established.

We now complete the measure \( P_\alpha \) on \( P^\alpha \), and denote the resulting space by \( (\Omega^\alpha, P^\alpha, P_\alpha) \). Henceforward a.s. will refer to this space.

By induction on \( r > k \) and Definition 3 it is seen easily that

\[ w_k(m\alpha_k) = w_r(E_{r-1} \cdots E_k(m\alpha_k)), \]

which may be written

\[ w_k(m\alpha_k) = w_r(m\alpha_k + L_{k,r}(m\alpha_k)). \]

Also, from Corollary 1 and (11) it follows that \( \lim_{s \to \infty} L_{k,r}(m\alpha_k) \) exists a.s. Accordingly, we define a process \( R_\alpha(t) \) for certain (random) values of \( t \) by setting, whenever \( \lim_{s \to \infty} L_{k,s}(m\alpha_k) \) exists(3),

\[ R_\alpha \left( m\alpha_k + \lim_{s \to \infty} L_{k,s}(m\alpha_k) \right) = w_k(m\alpha_k). \]

Indeed, this definition is independent of \( k \) and \( m \). For if \( m_1\alpha_k + \lim_{s \to \infty} L_{k,s}(m_1\alpha_k) = m_2\alpha_k + \lim_{s \to \infty} L_{k,s}(m_2\alpha_k), \) \( k < r, \) then (11) implies that \( m_1\alpha_k + L_{k,r}(m_1\alpha_k) + \lim_{s \to \infty} L_{r,s}(m_1\alpha_k + L_{k,r}(m_1\alpha_k)) = m_2\alpha_k + \lim_{s \to \infty} L_{r,s}(m_2\alpha_k) \). By considering \( r < s < \infty \) and Definition 3 it follows that \( m_1\alpha_k + L_{k,r}(m_1\alpha_k) = m_2\alpha_k \). In more detail, Definition 3 has the consequence that \( \alpha_k^{-1}(m_2\alpha_k + L_{r,s}(m_2\alpha_k)) \) is the least

(3) This definition will be amended for a set of probability 0 by Theorem 3, and to this extent is temporary.
integer such that for $M_{r+i} \cdots M_{r+1} w_r = w_r$ the ordered set
\[
\{ w_r(\alpha), w_r(2\alpha), \ldots, w_r(\alpha^{-1}(m_2\alpha_r + L_{r,s}(m_2\alpha_r))) \}
\]
contains the ordered set $\{ w_r(\alpha), w_r(2\alpha), \ldots, w_r(m_2\alpha_r) \}$ as an ordered sub-
set. Thus for each $w \in \Omega_\omega$, $m_2\alpha_r + L_{r,s}(m_2\alpha_r)$ determines $m_2$ uniquely. From
$m_2\alpha_k + L_{k,s}(m_1\alpha_k) = m_2\alpha_r$ and (14) it follows that $w_k(m_1\alpha_k) = w_r(m_2\alpha_r)$, hence
Definition (15) is independent of $k$ and $m$.

It is an immediate consequence of Corollary 1 that the times of the form
$m_1\alpha_k + \lim_{s \to \omega} L_{k,s}(m_1\alpha_k)$, $0 \leq m < \infty$, $0 < k < \infty$, are a.s. dense in $(0, \infty)$. We
next prove that $R_\omega(t)$ is a.s. continuous on this dense set, uniformly in finite
time intervals. This will allow us to conclude that if $R_\omega(t)$ is defined for the
remaining times by continuity (except for $w$ in a set of probability 0) then
the result is Brownian motion.

**Lemma 2.** For $\epsilon$, $\delta$, and $T$ positive, $\epsilon < T$, and $k$ sufficiently large,
\[
P^\omega\left\{ \max \left| R_\omega \left( m_1\alpha_k + \lim_{s \to \omega} L_{k,s}(m_1\alpha_k) \right) - R_\omega \left( m_2\alpha_k + \lim_{s \to \omega} L_{k,s}(m_2\alpha_k) \right) \right| > \delta \right\}
\leq \frac{4T}{\delta(2\pi\epsilon)^{1/2}} \exp\left( -\frac{\delta^2}{8\epsilon} \right),
\]
where the maximum is over $0 \leq m_1 < m_2 \leq \alpha_1^{-1}T$, $m_2 - m_1 < \epsilon\alpha_1^{-1}$.

**Proof.** By (14) the lemma is reduced to an estimation of the analogous
max $| R_k(m_1\alpha_k) - R_k(m_2\alpha_k) |$. For large $k$ the De Moivre-Laplace limit theorem
implies that $R_k(t+\epsilon) - R_k(t)$ is asymptotically normally distributed with ex-
pectation 0 and variance $\epsilon$. Combining this with the representation of $R_k(m\alpha_k)$
as the sum of $m$ independent increments we have, by using the reflection prin-
ciple [2, p. 392] and the usual inequalities for the normal distribution [3,
p. 166]: for $k$ large,
\[
P^\omega\left\{ \max \left| R_k(m_1\alpha_k) - R_k(m_2\alpha_k) \right| > \delta \right\}
\leq \frac{2T}{\epsilon} \left( 1 - \phi \left( \frac{\delta}{2\epsilon^{1/2}} \right) \right) \leq \frac{4T}{\delta(2\pi\epsilon)^{1/2}} \exp\left( -\frac{\delta^2}{8\epsilon} \right),
\]
where the maximum is taken as above and $\phi(x)$ is the normal distribution.

The sets $\{ m_1\alpha_k + \lim_{s \to \omega} L_{k,s}(m_1\alpha_k), 0 \leq m_1 < \infty \}$ are nondecreasing in $k$, as
follows from (11), and they a.s. become dense in $(0, T)$ as $k$ approaches $\infty$,
by Corollary 1 with $m = T\alpha_1^{-1}$. Letting $k$ become large, and choosing $\epsilon$ small
in Lemma 2, we find that $R_\omega(t)$ is uniformly continuous a.s. on this dense set,
and letting $T$ become large the same continuity holds a.s. in all bounded
subintervals of $(0, \infty)$. This leads to the following definition and theorem:
Theorem 3. The paths of \( R_k(t) \) a.s. converge uniformly in finite time intervals to continuous functions. Let \( R_\omega(t) \) be defined by

\[
R_\omega(t) = \begin{cases} 
\lim_{k \to \infty} R_k(t) & \text{if this limit exists uniformly in finite intervals and is continuous for } 0 \leq t < \infty, \\
0 & \text{otherwise.}
\end{cases}
\]

Then \( R_\omega(t) \) is a separable Brownian motion process on \((\Omega_\omega, F^\omega, P^\omega)\). Moreover, for \( \epsilon, \delta > 0 \), and \( k \) large

\[
P^\omega \left\{ \max_{0 \leq t \leq T, \epsilon \geq 1} \left| R_\omega(t) - R_\epsilon(t) \right| > \delta \right\}
\]

\[
\leq \frac{16T\alpha_k}{3\epsilon^2} \left( 1 - \frac{8(T + \epsilon)\alpha_k}{3\epsilon^2} \right)^{-1} + \frac{16T}{\delta(6\pi\epsilon)^{1/2}} \exp \left( -\frac{\delta^2}{96\epsilon} \right).
\]

Proof. It is sufficient to consider only the set of probability 1 where \( R_\omega(t) \) can be defined for all \( t \) by (15) and its extension by continuity. It will first be shown that within every time interval in which \( R_\omega(t) \) has oscillation greater than \( 2\beta_k \) there exist times \( m_1\alpha_k + \lim_{s \to \infty} L_{\omega,s}(m_1\alpha_k) \) and \( m_2\alpha_k + \lim_{s \to \infty} L_{\omega,s}(m_2\alpha_k) \) at which the maximum and minimum of \( R_\omega(t) \) in the interval are approached to within \( \beta_k \). For, by choosing \( r \) large, the maximum and minimum can be approximated as closely as desired at times of the form \( m_r \alpha_k + \lim_{s \to \infty} L_{\omega,s}(m_r \alpha_k) \) and thus, according to (15), by values of the form \( R_\epsilon(m_r \alpha_k) \). Between any two values \( R_r(t_1) \) and \( R_r(t_2) \) differing by at least \( 2\beta_k \) there exist values \( n_1\beta_k \) and \( n_2\beta_k \) taken on at times between \( t_1 \) and \( t_2 \), and approximating \( R_r(t_1) \) and \( R_r(t_2) \), respectively, to within \( \beta_k \). It follows from Definition 2 of the mappings \( M_k \), moreover, that it can even be assumed that the values \( n_1\beta_k \) and \( n_2\beta_k \) are taken on by \( R_r(t) \) at times of the form

\[
E_{r-1}E_{r-2} \cdots E_k(m_i \alpha_r) + \lim_{s \to \infty} L_{\omega,s}(E_{r-1}E_{r-2} \cdots E_k(m_i \alpha_r))
\]

it follows that the times \( m_i \alpha_k + \lim_{s \to \infty} L_{\omega,s}(m_i \alpha_k) \), \( i = 1, 2 \), are in the original interval for \( R_\omega(t) \). Since \( R_\omega(m_i \alpha_k + \lim_{s \to \infty} L_{\omega,s}(m_i \alpha_k)) = R_r(E_{r-1}E_{r-2} \cdots E_k(m_i \alpha_r) = n_i \beta_k \), by (15), the remark is established. The argument has also established the following fact: for \( r > k \) there exist within every time interval in which \( R_r(t) \) has oscillation greater than \( 2\beta_k \) times of the form \( m_i \alpha_k + L_{\omega,s}(m_i \alpha_k) \) and \( m_i \alpha_k + L_{\omega,s}(m_i \alpha_k) \) at which the maximum and minimum of \( R_r(t) \) in the interval are approached to within \( \beta_k \).

\(^{(4)}\) It is apparent that the completing of \((\Omega_\omega, F^\omega, P^\omega)\) is theoretically unnecessary since all sets considered must actually be in \( F_\omega \). However, it permits a considerable simplification in the proofs.
Because of these facts, and writing $L_{k, \infty}(m\alpha_k)$ in place of $\lim_{n \to \infty} L_{k, n}(m\alpha_k)$, we have for $\delta > 4\beta_k$ and $\epsilon > 0$,

\[
P_{\infty}^\epsilon \left\{ \begin{array}{l}
\text{l.u.b.} \\
0 \leq s \leq \tau_k, \delta \leq r \leq \infty \\
| R_r(t) - R_k(t) | > \delta/2
\end{array} \right. \\
\leq P_{\infty}^\epsilon \left\{ \begin{array}{l}
\text{l.u.b.} \\
0 \leq s \leq \tau_k, \delta \leq r \leq \infty \\
| R_r(m\alpha_k + \Delta) - R_r(m\alpha_k + L_k, r(m\alpha_k)) | > \delta/2
\end{array} \right. \\
+ P_{\infty}^\epsilon \left\{ \begin{array}{l}
\text{l.u.b.} \\
0 \leq s \leq \tau_k, \delta \leq r \leq \infty \\
| L_{k, r}(m\alpha_k) | > \epsilon
\end{array} \right. \\
+ P_{\infty}^\epsilon \left\{ \begin{array}{l}
\text{l.u.b.} \\
0 \leq s \leq \tau_k, \delta \leq r \leq \infty \\
| R_k(m_1\alpha_k) - R_k(m_2\alpha_k) | > \delta/2 + 2\beta_k
\end{array} \right.,
\]

(18)

where the first inequality is derived by writing $t = m\alpha_k + \Delta$ for some $m$ and $0 \leq \Delta < \alpha_k$, and using (14) and (15). Applying Corollary 2 and (16), respectively, to the last two terms of (18), and letting $k$ become large, (17) follows immediately from the triangle inequality.

Finally, since the increments of $R_m(t)$ are a.s. limits of the corresponding increments of $R_k(t)$ as $k$ becomes large, and since these increments are obviously asymptotically stationary and independent with mean 0 and variance equal to the corresponding increments of time, it follows easily that the increments of $R_m(t)$ have the same properties. Hence $R_m(t)$ is a process with a.s. stationary independent increments and continuous path functions, and it follows from a known theorem [2, p. 420] that $R_m(t)$ is a separable Brownian motion process. This completes the proof of Theorem 3.

**Further remarks.** (a) A converse to Theorem 3 may also be proved. Namely, any complete separable Brownian motion space contains a subspace which is a replica of $(\Omega_m, F_m^\omega, P_m^\omega)$ minus a set of measure 0. To define the random walk $R_k$ one utilizes the fact that the time of first arrival by a Brownian motion in $(\Omega_m, F_m^\omega, P_m^\omega)$.

(b) By defining the projections of a Brownian motion in $n$ dimensions on the coordinate axes separately, using Theorem 3 and the fact that these projections are independent, one immediately obtains a definition of $n$-dimensional Brownian motion in a product space $\prod_{i=1}^n (\Omega_m, F_m^\omega, P_m^\omega)$ as a uniform limit in finite time intervals of random walks $R_{k,n}(t)$ on the $n$-dimensional lattice whose points are the integral (vector) multiples of $n^{-1/2}\beta_k$. The transition probabilities for these $R_{k,n},$ analogous to (1), are

$$
p((n_1n^{-1/2}\beta_k, \ldots, n_nn^{-1/2}\beta_k), ((n_1+1)n^{-1/2}\beta_k, \ldots, (n_n+1)n^{-1/2}\beta_k)) = 2^{-n},
$$
and the time intervals between successive steps are of length \( \alpha_k \). An analogue of (17) is evident.

(c) It is possible to extend the construction of \((\Omega_\alpha, F^\alpha_\omega, P^\alpha_\omega)\) to apply not only to Brownian motion but to any diffusion process on a finite or infinite interval of the real line provided the process has no singular points and involves no more than absorbing barrier boundary conditions. To do this, one can use a theorem of H. Trotter \([9]\) which states that almost all Brownian motion paths have sojourn times in intervals \((-\infty, \tau)\) before time \( t \) which have densities in \( x \) that are jointly continuous in \( x \) and \( t \). After a continuous change of scale in \( x \) (to the “natural scale” of Feller \([4]\)) the diffusion is governed by an infinitesimal generator of the form \( d/dm \, d/dx \) (the “intrinsic form” \([4]\)) where \( m(x) \) is a right continuous strictly increasing function on the domain of definition (and to any such \( m(x) \) there corresponds a diffusion process in the natural scale). When the domain is \((-\infty, \infty)\) it follows easily from Trotter’s theorem that if we define a family of processes by

\[
R_k,n,m(x)(t) = R_k(M(t)\alpha_k) \quad \text{where} \quad M(t) \quad \text{is the integer determined by}
\]

\[
\sum_{i=0}^{M(t)} 2\alpha_k n \left( m \left( \frac{j(i) + 1}{n} \right) - m \left( \frac{j(i)}{n} \right) \right) \leq t \leq \sum_{i=0}^{M(t)+1} 2\alpha_k n \left( m \left( \frac{j(i) + 1}{n} \right) - m \left( \frac{j(i)}{n} \right) \right),
\]

in which \( j(i) \) is the function on \( \Omega_\alpha \) satisfying

\[
\frac{j(i)}{n} \leq R_k(i\alpha_k) < \frac{j(i) + 1}{n},
\]

then for a suitable increasing sequence \( n(k) \) the processes \( R_k,n,m(x)(t) \) converge a.s. uniformly in finite time intervals to a limit process \( R_\infty,m(x)(t) \). It can then be proved directly that \( R_\infty,m(x)(t) \) is a diffusion process with infinitesimal generator \( d/dm \, d/dx \). It follows in this way that the space \((\Omega_\alpha, F^\alpha_\omega, P^\alpha_\omega)\) suffices to define the entire class of such diffusion processes.

From this definition it is evident what interpretation is to be given to the measure \( dm(x) \). Roughly speaking, it determines the local “rate,” or time scale, for the process \( R_\infty,m(x)(t) \), relative to the Brownian motion \( R_\infty(t) \), but in the sense that when \( 2dm(x) \) is greater at \( x \) than \( dx \), then \( R_\infty,m(x)(t) \) “runs more slowly” at \( x \) than does \( R_\infty(t) \). A closer examination, taking into account the fact that the measure \( m(dx) \) may be singular with respect to Lebesgue measure, is not much facilitated by the construction of \((\Omega_\alpha, F^\alpha_\omega, P^\alpha_\omega)\), and we shall be content to refer to better sources \([4; 6]\). As for the proofs of the last statements, the convergence proof is routine, while the identification of the measure \( dm(x) \) with the infinitesimal generator \( d/dm \, d/dx \) will follow directly from \([6]\), and is due, in fact, to these authors. The proofs are therefore
omitted along with the evident modifications necessary to apply the construction to processes on an interval with absorbing boundaries.

References


Princeton University,
Princeton, New Jersey
University of Minnesota,
Minneapolis, Minnesota