Z. Wahrscheinlichkeitstheorie verw. Geb. 22, 251–267 (1972)
© by Springer-Verlag 1972

Stochastic Abelian and Tauberian Theorems

Ward Whitt *

1. Introduction and Summary

Let $\{X_n, n \geqq 0\}$ be a sequence of random variables and consider the discounted sum

$$D(z) = \sum_{k=0}^{\infty} z^k X_k, \qquad 0 < z < 1. \tag{1.1}$$

If we let $\{X_n\}$ represent a sequence of random payments evenly spaced in time, then $D(z)$ represents the present value of all future payments when the discount factor is z. More generally, we can let a stochastic process $\{S(t), t \geqq 0\}$ represent a cumulative income process. Then the integral

$$D(s) = \int_0^{\infty} e^{-sv}\, dS(v), \qquad s > 0, \tag{1.2}$$

represents the present value of all future income with a rate of interest s. Recently Gerber [8] proved a central limit theorem for $D(z)$ when $\{X_n\}$ is a sequence of i.i.d. (independent and identically distributed) random variables with mean μ, variance σ^2, and finite third moment. As a byproduct of a discounted version of the Berry-Esséen theorem, Gerber showed that the normalized random variable

$$Y(z) = \sigma^{-1}(1-z^2)^{\frac{1}{2}}\,[D(z) - \mu(1-z)^{-1}] \tag{1.3}$$

is asymptotically $N(0, 1)$ as $z \to 1$, where $N(a, b)$ denotes the normal distribution with mean a and variance b. Gerber also obtained a similar result for $D(s)$ when $S(t)$ is a compound Poisson process.

It seems quite natural in this setting to look for some direct connection between ordinary stochastic limit theorems and associated discounted stochastic limit theorems. In the spirit of the classical Abelian and Tauberian theorems, cf. Widder [34], we would like to say one holds if the other does, thus eliminating the need to prove discounted stochastic limit theorems in each of the multitude of situations in which ordinary stochastic limit theorems are known. In order to obtain stochastic analogues of the classical Abelian and Tauberian theorems, we turn to the function spaces associated with weak convergence of probability measures, cf. Billingsley [1]. A cursory inspection of the random variable $D(s)$ in (1.2) shows that it depends on the entire process $\{S(t), t \geqq 0\}$ rather than any one random variable obtained by looking at the process $\{S(t), t \geqq 0\}$ at a single time point. However, if we form the discounted stochastic process

$$D(s, t) = \int_0^{t} e^{-sv}\, dS(v), \qquad t \geqq 0, \tag{1.4}$$

* Partially supported by National Science Foundation Grant GK-27866.

then a connection can be made between the stochastic process $\{D(s,t),\ t\geqq 0\}$ and the stochastic process $\{S(t),\ t\geqq 0\}$. The connection is a homeomorphism on $D[0,r]$ for any $r>0$, which implies for any $r>0$ that there is convergence in the function space $D[0,r]$ for a sequence of random functions induced by discounted stochastic processes if and only if there is convergence for the corresponding sequence of random functions induced by the undiscounted stochastic processes. We only need to apply celebrated continuous mapping theorems, e.g. Theorem 5.1 of [1]. We do this not only for weak convergence, but also for convergence almost everywhere and in probability, so that we obtain discounted functional laws of large numbers and discounted functional laws of the iterated logarithm as well as discounted functional central limit theorems. For these other modes of convergence, our work relates to that of Gapoškin [7], Hanson and Wright [9], and references there.

It is possible to obtain discounted limit theorems for all these modes of convergence because there are continuous mapping theorems for each mode. In fact, it is only necessary to apply the obvious continuous mapping theorem associated with almost sure convergence ($f(X_n)\to f(X)$ a.s. if $X_n\to X$ a.s. and f is continuous) because each mode of convergence can be expressed in terms of almost sure convergence. For convergence in probability of a sequence of random elements in a separable metric space, it is well known that $X_n\to X$ in prob. if and only if every subsequence of $\{X_n\}$ has a further subsequence converging almost surely to X. For weak convergence in a complete separable metric space, the reduction to almost sure convergence is due to Skorohod [25], with extensions and applications by Dudley [5], Pyke [24], and Wichura [33]. If $X_n\Rightarrow X$, where X_n and X are random elements with values in a separable metric space and $\Rightarrow$ denotes weak convergence, then there exists a probability space $(\Omega^*, \mathscr{B}^*, P^*)$ on which are defined random elements Y_n and Y such that $Y_n\to Y$ a.s., $Y_n\sim X_n$, and $Y\sim X$, where $\sim$ denotes equality in distribution. Continuous mapping theorems follow immediately from such representations. These almost sure representations are extremely important because they show that the arguments can be carried out without reference to the probability measures.

It turns out that the functional limit theorems above are not sufficient to obtain discounted limit theorems for the present value of the entire stochastic process $\{S(t),\ t\geqq 0\}$. For this purpose, we introduce a new topology on the function space $D[0,\infty)$ which is stronger than Stone's [28] topology on $D[0,\infty)$. Our new topology is a natural generalization of a topology put on $C[0,\infty)$ by Müller [18]. This topology also brings us closer to the classical Abelian-Tauberian situation because in this topology we can go from functional limit theorems for the undiscounted processes to functional limit theorems for the discounted processes, but not the other way. Under Gerber's hypotheses [8], Corollary 4.1 here or Theorem 1 of [18] and Theorem 3.1 here imply Gerber's central limit theorem for $Y(z)$ in (1.3).

One principal conclusion is that we can get discounted stochastic limit theorems, of both the functional and the ordinary kind, if we can prove functional limit theorems for the basic stochastic process $\{S(t),\ t\geqq 0\}$. Thus, we might ask how hard is this condition to verify. It turns out not to be as stringent as it might appear. For example, consider functional central limit theorems. Roughly speaking, functional central limit theorems hold whenever ordinary central limit theorems hold. In addition to functional central limit theorems for a sequence of partial sums of i.i.d. random variables with finite variance (Theorems 10.1 and 16.1 of [1]), similar theorems exist for dependent sequences, Chapter 4 of [1]; partial sums from triangular arrays, Theorem 3.1 of [21], p. 220 of [20]; conditional sums, [15, 16]; random sums, Section 17 of [1], Lemma 1 and Corollary 1 of [12]; processes with stationary and independent increments as well as partial sums of i.i.d. random variables in the domain of attraction of a stable law, Theorem 2.7 of [26]; renewal processes, other counting processes, and first passage time processes, Theorem 17.3 of [1], Theorem 1 of [13]; functionals of Markov chains, [6] and [27]; and random walks, birth-and-death processes, and diffusion processes, [29]. Undoubtedly, this list is incomplete, but it is at least representative. A warning is appropriate here, however. All of these functional central limit theorems have been proved in $D[0,1]$. Since they usually hold in $D[0,r]$ for any $r>0$, they usually hold in $D[0,\infty)$ with Stone's topology [28], but they have yet to be demonstrated in $D[0,\infty)$ with the stronger topology introduced in Section 3. However, Müller's success [18] in the case of partial sums of independent random variables satisfying the Lindeberg condition suggests that corresponding theorems hold in the other cases as well.

In addition to the simple financial setting described at the outset, there are significant applications of discounted stochastic limit theorems to dynamic programming over stochastic processes. New stochastic criteria and stochastic sensitivity analyses in dynamic programming can be developed by focusing on the total reward stochastic process instead of the total expected reward (deterministic) process. Stochastic criteria can be defined in terms of limit theorems for the total reward stochastic processes. Central limit theorems for the total reward processes associated with dynamic programming over Markov chains were apparently first proved by Hatori [10, 11], but existing limit theorems for functionals on Markov chains and Markov renewal processes also serve this purpose; see [1], [6], and [22]. The theorems in this paper are significant in this regard because they relate stochastic averaging criteria with stochastic discounting criteria when the interest rate is small. In this way, we can obtain stochastic complements to recent optimization work by Denardo [3] and Veinott [31]. These ideas were introduced in an earlier version of this paper and will be discussed in more detail in a subsequent paper.

We now chart the way ahead. In Section 2 we establish equivalence of convergence in $D[0,r]$. We introduce the new topology on $D[0,\infty)$ in Section 3 and provide sufficient conditions for convergence of the present value of the entire discounted stochastic process. Section 4 is devoted the special case in which $\{S(t), t\geqq 0\}$ is generated from partial sums of i.i.d. random variables. We treat this case in order to obtain the distribution of the limiting process in Sections 2 and 3, but the proof there is also of independent interest. In Section 5 we give an inequality for the probability that the discounted process remains between two bounds, which we obtain by applying results of Skorohod [27] and Müller [18] for partial sums of uniformly bounded i.i.d. random variables. We conclude in Section 6 with a brief discussion of stochastic analogues of other summation methods.

2. Convergence Equivalence in $D[0, r]$

Let $D[0,r]$ be the function space consisting of all right-continuous real-valued functions on $[0,r]$ with the limits from the left everywhere, endowed with Skorohod's J_1 topology [25], which is the topology induced by the metrics d and d_0 in Chapter 3 of [1]. Let $D[0,\infty)$ be the corresponding space of functions on $[0,\infty)$ with Stone's extension [28] of Skorohod's J_1 topology from $D[0,1]$ to $D[0,\infty)$.

Let $\{S(t), t\geqq 0\}$ be a stochastic process in $D[0,\infty)$ with $S(0)=0$. If the set of paths of bounded variation in every finite interval has probability one, then the Stieltjes integral

$$\int_a^b e^{-sv}\,dS(v) \tag{2.1}$$

is defined for all finite a and b with probability one, cf. p. 7 of Widder [34], and we discount in this way. If the stochastic process $\{S(t), t\geqq 0\}$ is like Brownian motion and many other stochastic processes in that it is not of bounded variation in every-finite interval with probability one, then we understand (2.1) to be defined by the formula for integration by parts, that is, we assume

$$D(s,t)=\int_0^t e^{-sv}\,dS(v)=e^{-t}S(t)+s\int_0^t e^{-sv}S(v)\,dv, \qquad t\geqq 0. \tag{2.2}$$

We remark that this is consistent with various definitions of stochastic integrals, for example see Doob [4] and Paley *et al.* [19], but we shall take (2.2) as our starting point. It is easy to verify that we can retrieve $\{S(t), t\geqq 0\}$ from $\{D(s,t), t\geqq 0\}$ by

$$S(t)=\int_0^t e^{su}\,d_u D(s,u)=e^{st}D(s,t)-s\int_0^t e^{su}D(s,u)\,du. \tag{2.3}$$

The key to our Abelian and Tauberian theorems is

Lemma 2.1. *If $f\colon D[0,r]\to D[0,r]$ is defined for each $x\in D[0,r]$ by*

$$f(x)(t)=e^{-t}x(t)+\int_0^t e^{-v}x(v)\,dv, \qquad 0\leqq t\leqq r,$$

then f is a homeomorphism on $D[0,r]$ with the J_1 topology (or the uniform topology) and f^{-1} is given by

$$f^{-1}(x)(t)=e^{t}x(t)-\int_0^t e^{v}x(v)\,dv, \qquad 0\leqq t\leqq r.$$

Proof. It is easy to see that $f(f^{-1}(x))=f^{-1}(f(x))=x$ for all $x\in D[0,r]$. We only discuss the continuity of f because f^{-1} is treated in essentially the same way. We also only discuss the J_1 topology because the argument is similar (easier) with the uniform topology. We shall use the metric d in Chapter 3 of [1]. For this purpose, let Λ be the usual set of time deformations, that is, let Λ consist of all strictly-increasing continuous functions λ of $[0,r]$ onto $[0,r]$. Then $d(x,y)$ is defined for any $x,y\in D[0,r]$ as the infimum of those positive $\varepsilon>0$ for which there is a $\lambda\in\Lambda$ such that

$$\sup_{0\leqq t\leqq r}|\lambda(t)-t|\leqq\varepsilon \tag{2.4}$$

and

$$\sup_{0\leqq t\leqq r}|x(t)-y(\lambda(t))|\leqq\varepsilon. \tag{2.5}$$

Let $x\in D[0,r]$ be given and let $M=\sup_{0\leqq t\leqq r}|x(t)|$. Suppose $d(x,y)<\delta$. Now $d[f(x),f(y)]$ is the infimum over Λ of those $\varepsilon>0$ for which both

$$\sup_{0\leqq t\leqq r}|\lambda(t)-t| \tag{2.6}$$

and

$$\sup_{0\leqq t\leqq r}\left|e^{-t}x(t)+\int_0^t e^{-v}x(v)\,dv-e^{-\lambda(t)}y(\lambda(t))+\int_0^{\lambda(t)} e^{-v}y(v)\,dv\right| \tag{2.7}$$

are less than or equal to ε. Since $d(x,y)<\delta$, we can find a $\lambda\in\Lambda$ such that (2.4) and (2.5) hold with $\varepsilon=\delta$. Use this λ in (2.6) and (2.7). Then (2.6) is less than or equal to δ and it suffices to consider (2.7).

Now observe that (2.7) is less than or equal to

$$\sup_{0\leqq t\leqq r}|e^{-t}x(t)-e^{-\lambda(t)}y(\lambda(t))|+\sup_{0\leqq t\leqq r}\left|\int_0^t e^{-v}x(v)\,dv-\int_0^{\lambda(t)} e^{-v}y(v)\,dv\right|; \tag{2.8}$$

where

$$\begin{aligned}
&\sup_{0\leqq t\leqq r}|e^{-t}x(t)-e^{-\lambda(t)}y(\lambda(t))|\\
&\quad\leqq\sup_{0\leqq t\leqq r}|e^{-t}x(t)-e^{-\lambda(t)}x(t)|+\sup_{0\leqq t\leqq r}|e^{-\lambda(t)}x(t)-e^{-\lambda(t)}y(\lambda(t))|\\
&\quad\leqq M\sup_{0\leqq t\leqq r}|e^{-t}-e^{-\lambda(t)}|+\sup_{0\leqq t\leqq r}|x(t)-y(\lambda(t))| \qquad (2.9)\\
&\quad\leqq M\sup_{0\leqq t\leqq r}|\lambda(t)-t|+d(x,y)\\
&\quad\leqq(M+1)\,\delta;
\end{aligned}$$

and

$$\begin{aligned}
&\sup_{0\leqq t\leqq r}\left|\int_0^t e^{-v}x(v)\,dv-\int_0^{\lambda(t)} e^{-v}y(v)\,dv\right|\\
&\leqq\sup_{0\leqq t\leqq r}\left|\int_0^t e^{-v}x(v)\,dv-\int_0^{\lambda(t)} e^{-v}x(v)\,dv\right|+\sup_{0\leqq t\leqq r}\left|\int_0^{\lambda(t)} e^{-v}x(v)\,dv-\int_0^{\lambda(t)} e^{-v}y(v)\,dv\right|\\
&\leqq M\sup_{0\leqq t\leqq r}|\lambda(t)-t|+\sup_{0\leqq t\leqq r}\left|\int_0^t e^{-v}[x(v)-y(v)]\,dv\right| \qquad (2.10)\\
&\leqq M\delta+\sup_{0\leqq t\leqq r}\int_0^r|x(v)-y(v)|\,dv\\
&\leqq M\delta+\alpha(\delta),
\end{aligned}$$

where $\alpha(\delta)\to 0$ as $\delta\to 0$ because the J_1 topology is finer than that given by the metric $\int_0^r|x(t)-y(t)|\,dt$. This is easy to see because $d(x_n,x)\to 0$ implies $x_n(t)\to x(t)$ for all continuity points t of x, cf. p. 112 of [1], which means convergence almost everywhere with respect to Lebesgue measure. Then apply the Lebesgue dominated convergence theorem to get $\int_0^r |x_n(t)-x(t)|\,dt \to 0$, cf. [1], Problem 2, p. 123. Hence, $d[f(x), f(y)] \leqq (2M+2)\,\delta + \alpha(\delta) \to 0$ as $\delta \to 0$, so that continuity is demonstrated.

To include more general limit theorems, we consider a sequence of stochastic processes $\{\{S^i(t),\ t\geqq 0\},\ i\geqq 1\}$. Let $\{U_n\}$ be the associated sequence of undiscounted random functions induced in $D[0,\infty)$ or $D[0,r]$ by this sequence of stochastic processes, that is, let

$$U_n(t) = S^n(n\,t)/\phi(n), \qquad t\geqq 0, \tag{2.11}$$

where ϕ: $(0,\infty)\to(0,\infty)$ is the appropriate normalization function. For the usual central limit theorems, $\phi(n)=n^{\frac{1}{2}}$, but $\phi(n)=n^{\alpha}$, $\alpha>0$, and $\phi(n)=(2n\log\log n)^{\frac{1}{2}}$ are possible. We explicitly assume that $\phi(n)\to\infty$ as $n\to\infty$. Now let $\{D_n\}$ be the corresponding sequence of discounted random functions induced in $D[0,\infty)$ or $D[0,r]$:

$$D_n(t) = \phi(n)^{-1} \int_0^{nt} e^{-v/n}\, dS^n(v), \qquad t\geqq 0. \tag{2.12}$$

We have set the discount rate in the n-th system at $1/n$. Our main result follows immediately from Lemma 2.1 because in $D[0,r]$, after an integration by parts, we get

$$D_n = f(U_n)$$

and

$$U_n = f^{-1}(D_n). \tag{2.13}$$

For the statement of the theorem, let $\Rightarrow$ denote weak convergence and let d be the metric inducing Skorohod's J_1 topology [25] on $D[0,r]$. Let the non-random functions X, Y, and C be defined by

$$\begin{aligned} X(t) &= 1-e^{-2t}, & t&\geqq 0,\\ Y(t) &= 1-e^{-t}, & t&\geqq 0,\\ C(t) &= c\,t, & t&\geqq 0. \end{aligned} \tag{2.14}$$

Let D_0 be the subset of $D[0,r]$ consisting of all non-decreasing functions $\psi(t)$ with $0\leqq\psi(t)\leqq r$. Let $\circ$ be the composition map, defined for any $(x,y)\in D[0,r]\times D_0$ by

$$(x\circ y)(t) = x(y(t)), \qquad t\geqq 0, \tag{2.15}$$

cf. p. 144 of [1].

Theorem 2.1. *For $n\geqq 1$, let U_n and D_n be the undiscounted and discounted random functions in $D[0,r]$ defined in* (2.11) *and* (2.12). *Let X, Y, and C be as in* (2.14). *Let $\circ$ be the composition map in* (2.15). *Let f be as in Lemma* 2.1. *And let W be the Wiener process in $D[0,r]$. Then*

(a) *$U_n\Rightarrow U$ if and only if $D_n\Rightarrow D$, where $D=f(U)$ and $U=f^{-1}(D)$;*

(b) *$U=W+C$ if and only if $D=(2^{-\frac{1}{2}})(W\circ X+cY)$;*

(c) *if $\{U_n\}$ and U are defined on a common probability space, then $d(U_n, U)\to 0$ with probability one (in probability) if and only if $d(D_n, f(U))\to 0$ with probability one (in probability);*

(d) *the sequence* $\{U_n\}$ *is relatively compact in* $D[0,r]$ *and its set of limit points is the compact set* $K(c)$ *if and only if the sequence* $\{D_n\}$ *is relatively compact in* $D[0,r]$ *with the sequence of limit points* $f(K(c))$, *cf. Strassen* [30].

Proof. (a) Integration by parts in (2.12) yields

$$D_n(t)=e^{-t}S^n(n\,t)/\phi(n)+(1/n\,\phi(n))\int_0^{nt} e^{-v/n}S^n(v)\,dv, \tag{2.16}$$

and after the change of variables $v=u/n$, we obtain

$$D_n(t)=e^{-t}S^n(n\,t)/\phi(n)+\int_0^t e^{-v}[S^n(n\,v)/\phi(n)]\,dv. \tag{2.17}$$

It only remains to apply Lemma 2.1 with the continuous mapping theorem, cf. Theorem 5.1 of [1]. It is easy to see that this result remains unchanged if $S^n(0)\neq 0$ as long as $S^n(0)/\phi(n)\Rightarrow 0$ as $n\to\infty$; we only need to apply Theorem 4.1 of [1].

(b) The invariance principle associated with weak convergence specifies a unique limit when U or D is determined. This limit is determined in Section 4 by considering a special case. Note that $e^{-k/n}\sim\left(1-\frac{1}{n}\right)^k$ so that we can work with either $D(z)$ or $D(s)$, cf. (1.1) and (1.2).

(c) The proof of part (a) applies here too but with the continuous mapping theorems associated with convergence a.s. and convergence in probability. As we remarked in the introduction, all these continuous mapping theorems can be obtained by using almost surely convergent representations. The functional strong laws should not be taken too seriously because it is easy to show that functional strong laws are equivalent to ordinary strong laws. This is not true for weak laws, however.

(d) Use the continuous mapping theorem associated with the functional law of the iterated logarithm, cf. Corollary to Theorem 3 of Strassen [30]. The setting in [30] is $C[0,1]$, but it is easily extended to $D[0,r]$. Again, this continuous mapping theorem is an immediate consequence of the almost sure convergence of subsequences.

3. A Stronger Topology on $D[0,\infty)$

It is easy to see that convergence in $D[0,\infty)$ with Stone's topology [28] is equivalent to convergence in $D[0,r_n]$ for all n, where $\{r_n,\ n\geqq 1\}$ is some sequence of positive numbers such that $r_n\to\infty$ as $n\to\infty$. Hence, Theorem 2.1 here applies to $D[0,\infty)$ as well as $D[0,r]$. We would now like to go from $D_n\Rightarrow D$ in $D[0,\infty)$ to $\lim_{t\to\infty} D_n(t)\Rightarrow\lim_{t\to\infty} D(t)$ in order to get limits for random variables such as $Y(z)$ in (1.3). First note, however, that $\lim_{t\to\infty} f(x)(t)$ is not even defined for all $x\in D[0,\infty)$ if f is the discounting map defined in Lemma 2.1: if $x(t)=e^{2t}$, $t\geqq 0$, then

$$f(x)(t)=e^{-t}x(t)+\int_0^t e^{-v}x(v)\,dv=2e^t-1,\qquad t\geqq 0. \tag{3.1}$$

Therefore, let us consider the subset of $D[0,\infty)$ containing only functions x for which

$$\lim_{t\to\infty}\frac{x(t)}{t^\alpha}=0, \tag{3.2}$$

where α is a fixed positive constant (usually $\alpha=1$). We refer to this space as $L\equiv L[0,\infty)$ and endow it with the relative topology. We have chosen this particular subset of $D[0,\infty)$ because in the presence of the strong law of large numbers all the random elements U_n in (2.11) can be regarded as elements of L with $\alpha=1$.

Now we would like to show that the map $g: L\to R$ defined by $g(x)=\lim_{t\to\infty} f(x)(t)$ or

$$g(x)=\int_0^\infty e^{-v}x(v)\,dv \tag{3.3}$$

is continuous, but this is not true. For example, let $x(t)=0$, $t\geqq 0$, and

$$x_n(t)=\begin{cases} e^n, & n\leqq t<n+1\\ 0, & t<n \text{ and } t\geqq n+1.\end{cases} \tag{3.4}$$

Then for each $r>0$, $\sup_{0\leqq t\leqq r}|x_n(t)-x(t)|=0$ for $n>r$, so that $d(x_n,x)\to 0$ in L, but $|g(x_n)-g(x)|=1-e^{-1}$ for all $n\geqq 1$.

The difficulty just encountered suggests that we should introduce a stronger topology on L. There are many precedents for using both stronger and weaker topologies on $D[0,r]$ and $D[0,\infty)$, cf. Chernoff [2], Lamperti [14], Pyke and Shorack [23], Skorohod [25], Whitt [32], and Woodroofe [35]. We shall use a new topology on L corresponding to Müller's topology [18] on $L[0,\infty)\cap C[0,\infty)$, which can be defined by the metric e:

$$e(x,y)=\sup_{t\geqq 0}\frac{|x(t)-y(t)|}{1+t^\alpha}, \tag{3.5}$$

where α is a fixed positive constant, also see (5.6) of [2] and (2.3) of [23]. Of course, we want to allow small time deformations. Let Λ' be the set of all strictly-increasing continuous maps of $[0,\infty)$ onto itself. Let $\alpha>0$ be fixed and let $m(x,y)$ be defined for any $x,y\in L$ as the infimum of those $\varepsilon>0$ for which there is a $\lambda\in\Lambda'$ such that

$$\sup_{t\geqq 0}\frac{|x(t)-y(\lambda(t))|}{1+t^\alpha}\leqq\varepsilon \tag{3.6}$$

and

$$\sup_{t\geqq 0}|\lambda(t)-t|\leqq\varepsilon. \tag{3.7}$$

Let $m_0(x,y)$ be defined in the same way with (3.8) instead of (3.7):

$$\sup_{s\neq t}\left|\log\frac{\lambda(t)-\lambda(s)}{t-s}\right|\leqq\varepsilon. \tag{3.8}$$

Lemma 3.1. *The spaces (L,m) and (L,m_0) are separable metric spaces with the same topology, but m_0 makes L complete, while m does not.*

Proof. The arguments on pp. 111–116 of [1] are easily extended to cover this case. Use the modulus of continuity

$$v'_x(\delta)=\inf_{\{t_i\}}\max_{0<i\leqq r} v_x[t_{i-1},t_i), \tag{3.9}$$

where v_x is defined in Lemma 3.2 below and the infimum extends over the finite sets $\{t_i\}$ of points satisfying

$$0=t_0<t_1<\cdots<t_r=\infty, \tag{3.10}$$

and $t_i-t_{i-1}>\delta$, $1\leqq i\leqq r$, cf. p. 110 of [1].

Before considering discounted limit theorems in (L, m), we shall examine the structure of (L, m) in more detail. From (3.2) and Lemma 1 on p. 110 of [1], we immediately get

Lemma 3.2. *For each $x\in L$ and each $\varepsilon>0$, there exist finitely many points $t_0, t_1, \ldots, t_r$ such that*

$$0=t_0<t_1<\cdots<t_r=\infty \tag{3.11}$$

and

$$v_x[t_{i-1},t_i)<\varepsilon, \quad 1\leqq i\leqq r, \tag{3.12}$$

where

$$v_x[t_{i-1},t_i)=\sup\left\{\frac{|x(t)-x(s)|}{1+t^{\alpha}}:\ t_{i-1}\leqq s,\ t<t_i\right\}. \tag{3.13}$$

Let c_r: $L[0,\infty)\to D[0,r]$ be the restriction of x in L to $[0,r]$, that is, let

$$c_r(x)(t)=x(t), \quad 0\leqq t\leqq r. \tag{3.14}$$

Lemma 3.3. *A set B has compact closure in (L, m) if and only if $c_r(B)$ has compact closure in $D[0, r]$ for each $r>0$ and*

$$\lim_{t\to\infty}\sup_{x\in B}\frac{|x(t)|}{t^{\alpha}}=0. \tag{3.15}$$

Proof. This is a minor modification of Theorem 14.3 of [1].

Lemma 3.4. *A sequence of probability measures $\{P_n, n\geqq 1\}$ on (L, m) is tight if and only if*

(i) *$\{P_n c_r^{-1}, n\geqq 1\}$ is tight for each $r>0$, and*

(ii) *for each positive ε and η, there exists a t_0 such that for all $n\geqq 1$*

$$P_n\left(\left\{x:\ \sup_{t>t_0}\left|\frac{x(t)}{t^{\alpha}}\right|>\varepsilon\right\}\right)\leqq\eta. \tag{3.16}$$

Proof. Lemma 3.4 follows from Lemma 3.3 in the same way that Theorem 15.2 follows from Theorem 14.3 and Theorem 8.2 follows from the Arzelà-Ascoli theorem in [1]. Also see p. 176 of [18].

Let π_t: $L\to R^k$ be the coordinate projection defined by $\pi_t(x)=x(t)$.

Lemma 3.5. *The Borel σ-field associated with (L, m) coincides with the σ-field on L generated by the coordinate projections.*

Proof. Apply the argument on pp. 121–122 of [1].

For a probability measure P on (L, m), let T_P consist of those t in $[0,\infty)$ for which the projection π_t is continuous almost everywhere with respect to P.

Lemma 3.6. *$P_n \Rightarrow P$ on (L, m) if and only if $\{P_n, n \geqq 1\}$ is tight and*

$$P_n \pi_{t_1,\ldots,t_k}^{-1} \Rightarrow P \pi_{t_1,\ldots,t_k}^{-1}$$

in R^k whenever $t_1, \ldots, t_k$ all lie in T_P.

Proof. Again the argument for $D[0,1]$ applies, cf. pp. 35, 123, and 241 of [1].

For an example of a weak convergence theorem on (L, m), see Theorem 1 of [18]. The way to check the additional condition (3.16) needed to get weak convergence in (L, m) from weak convergence in $D[0,1]$ is indicated on p. 177 of [18]. We now return to our discounted limit theorems. Instead of Lemma 2.1, we have

Lemma 3.7. *Let $f: (L, m) \to (L, m)$ be defined as in Lemma 2.1 and let $g: (L, m) \to R$ be as in (3.3). Then f and g are continuous, but f^{-1} is not. The same is true if the metric e in (3.5) is used instead of m throughout.*

Proof. The argument is similar to the proof of Lemma 2.1 so we omit most of it. The continuity of g can be obtained as a by-product of the proof for f. Directly, we have

$$|g(x_n) - g(x)| = \left| \int_0^\infty e^{-t} [x_n(t) - x(t)]\, dt \right| \leqq \int_0^\infty e^{-t} |x_n(t) - x(t)|\, dt \to 0 \qquad (3.17)$$

as $n \to \infty$ by virtue of the Lebesgue dominated convergence theorem. It is easy to see that $m(x_n, x) \to 0$ implies that $x_n(t) \to x(t)$ at all continuity points of x, cf. p. 112 of [1], which means convergence almost everywhere with respect to Lebesgue measure. For sufficiently large n, an integrable dominating function is

$$h(t) = e^{-t}(1 + t^\alpha)\, 3 \sup_{t \geqq 0} \frac{|x(t)|}{1 + t^\alpha}, \qquad t \geqq 0. \qquad (3.18)$$

To see that f^{-1} is not continuous, let $x(t) = 0$, $t \geqq 0$, and

$$x_n(t) = \begin{cases} 1 & n \leqq t < n+1 \\ 0 & t < n,\ t \geqq n+1. \end{cases} \qquad (3.19)$$

Then

$$\sup_{t \geqq 0} \frac{|x_n(t) - x(t)|}{1 + t^\alpha} = \sup_{t \geqq 0} \frac{|x_n(t)|}{1 + t^\alpha} = \frac{1}{1 + n^\alpha} \to 0, \qquad (3.20)$$

but

$$m[f^{-1}(x_n), f^{-1}(x)] = \sup_{t \geqq 0} \frac{|f^{-1}(x_n)(t)|}{1 + t^\alpha} \geqq \frac{f^{-1}(x_n)(n)}{1 + n^\alpha} = \frac{e^n}{1 + n^\alpha} \to \infty. \qquad (3.21)$$

In fact, if $x(t) = t\, e^{-t}$, $t \geqq 0$, then $f^{-1}(x) \notin (L, m)$.

Let U_n and D_n be the undiscounted and discounted random functions defined in (2.11) and (2.12); let X, Y, and C be as in (2.14); let $\circ$ be the composition map in (2.15); let f be as in Lemma 2.1 and g be as in (3.3); let W be the Wiener process in (L, m); and let $D_n(\infty)$ be defined by

$$\begin{aligned} D_n(\infty) &= \lim_{t \to \infty} D_n(t) = \phi(n)^{-1} \int_0^\infty e^{-v/n}\, dS^n(v) \\ &= \int_0^\infty e^{-v}\, U_n(v)\, dv. \end{aligned} \qquad (3.22)$$

Theorem 3.1. *In the setting above,*

(a) *if* $U_n \Rightarrow U$ *in* (L, m), *then* $D_n \Rightarrow f(U)$ *in* (L, m) *and* $D_n(\infty) \Rightarrow g(U)$ *in* R;

(b) *if* $U_n \Rightarrow W + C$ *in* (L, m), *then* $D_n \Rightarrow (2^{-\frac{1}{2}})(W \circ X + cY)$ *in* (L, m) *and* $D_n(\infty) \Rightarrow N(c/2, \frac{1}{2})$ *in* R;

(c) *if* $m(U_n, U) \to 0$ *in probability (with probability one) in* (L, m), *then* $m(D_n, f(U)) \to 0$ *in* (L, m) *and* $|D_n(\infty) - g(U)| \to 0$ *in* R *in probability (with probability one)*;

(d) *if* $m(U_n, C) \to 0$ *in probability (with probability one) in* (L, m), *then* $m(D_n, cY) \to 0$ *in* (L, m) *and* $|D_n(\infty) - c| \to 0$ *in* R *in probability (with probability one)*;

(e) *if* $\{U_n\}$ *is relatively compact in* (L, m) *and its set of limit points is the compact set* $K(c)$, *then* $\{D_n\}$ *is relatively compact in* (L, m) *with the set of limit points* $f(K(c))$ *and* $\{D_n(\infty)\}$ *is relatively compact in* R *with the set of limit points* $g(K(c))$;

(f) *if* $|U_n(1) \to U(1)| \to 0$ *in* R *with probability one with* $\phi(n) = n^\alpha$, *then* $|D_n(\infty) - \Gamma(\alpha+1)\, U(1)| \to 0$ *in* R *with probability one.*

Proof. Parts (a)–(e) are a consequence of Lemma 3.7 and the continuous mapping theorems. The distribution in (b) is obtained in Section 4. Part (f) is the standard Abelian theorem, which can be found on p. 181 of Widder [34]. The standard counterexamples show that the converse in (f) is not true, cf. p. 186 of [34]. For (e) a slightly different space is actually more appropriate, cf. Section 2 of [18].

4. Discounted Sums of i.i.d. Random Variables

In order to determine the form of the limiting process in Theorems 2.1(b) and 3.1(b), we consider the special case in which the stochastic processes $\{S^i(t), t \geq 0\}$ are constructed from partial sums of i.i.d. random variables. The invariance principle associated with weak convergence implies that this process is the limiting process in the more general setting of Theorems 2.1(b) and 3.1(b). At the same time, we are providing a direct proof of Gerber's discounted central limit theorem [8] for $Y(z)$ in (1.3) and its weak convergence generalization. We use Prohorov's function space generalization of the Lindeberg-Feller central limit theorem ([21], Theorem 3.1) and a time transformation in the manner of [1], Section 17. The direct application of Prohorov's theorem involves an artificial transformation of the time scale. The original time scale is then restored by an inverse time transformation.

Assume that $\{X_{ni}, i \geq 1\}$ is a sequence of i.i.d. random variables for each n; $EX_{ni} = \mu_n \to \mu$; $\sigma^2(X_{ni}) = \sigma_n^2 \to \sigma^2$; $0 < \sigma^2$, $\sigma_n^2 < \infty$; and $E|X_{ni}|^{2+\delta}$ is uniformly bounded in n for some $\delta > 0$. (The bound on $E|X_{ni}|^{2+\delta}$ is unnecessary with a single sequence; for double sequences the uniform bound is an easily verifiable substitute for the more general Lindeberg condition.) For each $n \geq 1$ and z, $0 < z < 1$, there is a discounted sum of i.i.d. random variables $D_n(z)$ defined by

$$D_n(z) = \sum_{i=0}^{\infty} z^i X_{ni}. \tag{4.1}$$

We first verify that $D_n(z)$ is well defined.

Lemma 4.1. *For each $n \geqq 1$ and z, $0<z<1$, the discounted sum in* (4.1) *is an a.e. convergent series.*

Proof. Let $U_{ni} = z^i (X_{ni} - \mu_n)$. Then $\sigma^2(U_{ni}) = \sigma_n^2 z^{2i}$ and

$$\sum_{i=0}^{\infty} \sigma^2(U_{ni}) = \sigma_n^2 (1-z^2)^{-1} < \infty.$$

Hence, $\sum_{i=0}^{\infty} U_{ni}$ is a convergent series a.e., cf. p. 236 of [17]. Since

$$\sum_{i=0}^{\infty} \mu_n z^i = \mu_n (1-z)^{-1} < \infty,$$

(4.1) is convergent a.e. as well. We could also apply a martingale convergence theorem here, cf. p. 393 of [17].

Now we apply Prohorov's theorem, cf. p. 220 of [20]. For this purpose, we let $z = (1-1/n)$ and define two new double sequences of normalized random variables $\{Y_{ni}\}$ and $\{Z_{ni}\}$ by setting

$$Y_{ni} = (n\,\sigma_n^2)^{-\frac{1}{2}} (2-1/n)^{\frac{1}{2}} (1-1/n)^i (X_{ni} - \mu_n)$$

and

$$Z_{ni} = (n\,\sigma_n^2)^{-\frac{1}{2}} (2-1/n)^{\frac{1}{2}} (1-1/n)^i X_{ni}. \tag{4.2}$$

It is easy to check that the conditions of Prohorov's theorem are satisfied for $\{Y_{ni}\}$. For each $n \geqq 1$, $\{Y_{ni}\}$ is i.i.d., $EY_{ni} = 0$, and $\sum_{i=0}^{\infty} \sigma^2(Y_{ni}) = 1$. Furthermore, the Lindeberg condition is satisfied. Note, however, that we are adding an infinite number of random variables in each row instead of a finite number. The theorem applies in this case too as long as the series $\sum_{i=0}^{\infty} \sigma^2(X_{ni})$ converges. This is easily verified by truncating in each row at a point increasing very rapidly with n (for example, set $k_n = n^2$) and then applying Theorem 4.1 of [1].

We now define the two sequences of random functions $\{A_n\}$ and $\{B_n\}$ induced in $D[0,1]$ by the row sums from $\{Y_{ni}\}$ and $\{Z_{ni}\}$. For $t \in [0,1]$, let

$$A_n(t) = \begin{cases} \sum_{i=0}^{\infty} Y_{ni}, & t=1 \\ \sum_{i=0}^{k} Y_{ni}, & \sum_{i=0}^{k} \sigma^2(Y_{ni}) \leqq t < \sum_{i=0}^{k+1} \sigma^2(Y_{ni}) \\ 0, & 0 \leqq t < \sigma^2(Y_{n0}), \end{cases}$$

$$B_n(t) = \begin{cases} \sum_{i=0}^{\infty} Z_{ni}, & t=1 \\ \sum_{i=0}^{k} Z_{ni}, & \sum_{i=0}^{k} \sigma^2(Y_{ni}) \leqq t < \sum_{i=0}^{k+1} \sigma^2(Y_{ni}) \\ 0, & 0 \leqq t < \sigma^2(Y_{n0}). \end{cases} \tag{4.3}$$

Prohorov's theorem was originally proved for linearly-interpolated versions of $\{A_n\}$ in $C[0,1]$, but it is now well known that it applies equally well in $D[0,1]$ to $\{A_n\}$ in (4.3). Hence, we have

Theorem 4.1. *Let A_n and B_n be defined by* (4.3) *with Y_{ni} and Z_{ni} in* (4.2). *Let W be the Wiener process in $D[0,1]$. Then*

$$\text{(a)}\quad A_n \Rightarrow W \text{ in } D[0,1];$$

and

$$\text{(b)}\quad \text{if } (2n)^{\frac{1}{2}}\mu_n\sigma_n^{-1}\to c, \qquad -\infty<c<+\infty,$$

then

$$B_n \Rightarrow W+E \text{ in } D[0,1],$$

where

$$E(t)=c[1-(1-t)^{\frac{1}{2}}], \qquad 0\leqq t\leqq 1.$$

Proof. We have remarked that (a) is an immediate consequence of Prohorov's theorem, cf. [20], p. 220. Part (b) is obtained by adding the (non-random) translation functions to both sides of the result in (a), using Theorems 4.4 and 5.1 of [1]. Although addition is not continuous in $D[0,1]$, it is measurable in $D[0,1]$ and continuous in $C[0,1]$. Note that the n-th translation term has the value $(n\sigma_n^2)^{-\frac{1}{2}}(2-1/n)^{\frac{1}{2}}\mu_n\sum_{i=1}^{k}(1-1/n)^i$ at time t, where

$$n^{-1}(2-1/n)\sum_{i=0}^{i=k}(1-1/n)^{2i}\leqq t<n^{-1}(2-1/n)\sum_{i=0}^{i=k+1}(1-1/n)^{2i}.$$

Letting $k=[nt]$, we see that asymptotically the translation term has value $c(1-e^{-t})$ at $1-e^{-2t}$, where convergence is uniform on $[0,\infty)$. The desired result is then obtained by taking logarithms.

We can immediately apply the continuous mapping theorem to obtain the corresponding ordinary discounted central limit theorems. Part (a) contains Gerber's discounted limit theorem [8] for $Y(z)$ in (1.3).

Corollary 4.1. *Let $\{Y_{ni}\}$ and $\{Z_{ni}\}$ still be defined by* (4.2). *Then*

$$\text{(a)}\quad \sum_{i=0}^{\infty} Y_{ni}\Rightarrow N(0,1) \text{ in } R^1;$$

and

$$\text{(b)}\quad \text{if } (2n)^{\frac{1}{2}}\mu_n\sigma_n^{-1}\to c, \qquad -\infty<c<\infty,$$

then

$$\sum_{i=0}^{\infty} Z_{ni}\Rightarrow N(c,1) \text{ in } R^1.$$

Proof. Since W and $W+E$ are contained in $C[0,1]$, all projections are continuous functions almost everywhere on $D[0,1]$. Hence, (a) and (b) here are obtained from (a) and (b) in Theorem 4.1 by applying the projection at $t=1$ with Theorem 5.1 of [1].

We now obtain the promised functional central limit theorem by restoring the original timing. Define random functions T_n and V_n in $D[0,\infty)$ by setting

$$\begin{aligned} T_n(t) &= \sum_{i=0}^{[nt]} Y_{ni}, \quad t \geqq 0, \\ V_n(t) &= \sum_{i=0}^{[nt]} Z_{ni}, \quad t \geqq 0, \end{aligned} \tag{4.4}$$

where Y_{ni} and Z_{ni} are defined in (4.2). Of course, since $T_n(t)$ and $V_n(t)$ have finite limits as $t \to \infty$ with probability one (Lemma 4.1), T_n and V_n can be regarded as elements of $L[0,\infty)$ with $\alpha \geqq 1$.

Theorem 4.2. *Let T_n and V_n be defined by* (4.4); *let W be the Wiener process on $D[0,1]$; and let X, Y, and C be defined in* (2.14). *Then*

$$\text{(a)} \quad T_n \Rightarrow W \circ X \ \text{ in } D[0,\infty);$$

and

$$\text{(b)} \quad \text{if } (2n)^{\frac{1}{2}} \mu_n \sigma_n^{-1} \to c, \qquad -\infty < c < \infty,$$

then

$$V_n \Rightarrow W \circ X + c\,Y \ \text{ in } D[0,\infty).$$

Proof. Note that $T_n = A_n \circ X_n$ and $V_n = B_n \circ X_n$, where X_n is a nondecreasing deterministic transformation of $[0,\infty)$ onto $[0,1)$. Since $\lim_{t\to\infty} X_n(t) = 1$, we can use $[0,1]$ instead of $[0,1)$. Furthermore,

$$\lim_{n\to\infty} X_n(t) = 1 - e^{-2t}, \qquad t \geqq 0,$$

uniformly in t. Hence, the argument of [1], p. 145, together with Theorem 4.1 implies that $A_n \circ X_n \Rightarrow W \circ X$ and $B_n \circ X_n \Rightarrow (W+E) \circ X$ in $D[0,n]$ for each n, and thus in $D[0,\infty)$. Finally, $(W+E) \circ X = W \circ X + c\,Y$.

The discrepancy between Theorems 4.2 and 2.1(b) is due to the $(2-1/n)^{\frac{1}{2}}$ term in Y_{ni} and Z_{ni} in (4.2). The connection with Theorems 2.1(b) and 3.1(b) is easily made. The usual random function induced by the undiscounted partial sums of the X_{ni} is U_n and $V_n = f(U_n)$, where f is defined in Lemma 2.1 as before. For Theorem 2.1(b), we apply Prohorov's theorem (p. 220 of [20]) to get $U_n \Rightarrow W + C$ in $D[0,r]$. For Theorem 3.1(b), we apply the double sequence version of Theorem 1 of [18] which states that $U_n \Rightarrow W + C$ in (L, m). The strong law of large numbers holds in each row of the double sequences. Hence, the negligible set for the double sequence is just the union of the negligible sets associated with the rows. Theorem 3.1(a) then implies that $V_n \Rightarrow f(U)$ in (L, m). Theorem 4.2 identifies the distribution of $f(U)$. Section 3 implies that the limits in Theorem 4.2 also hold in the space (L, m), but it is not necessary to verify this directly in Theorem 4.2 in order to identify the limiting distribution, cf. [1], p. 35.

5. The Discounted Process between Two Bounds

In the setting of Section 4, assume that $P\{|X_{ni}| > M\} = 0$ for all n. Skorohod ([27], p. 169) has evaluated the rate of convergence in this situation of the probability

$$\Pr\left\{g_1(t) < \frac{S^n_{[nt]} - \mu_n [n t]}{(n\sigma^2)^{\frac{1}{2}}} < g_2(t),\ 0 \leqq t \leqq 1\right\} \tag{5.1}$$

to the probability

$$\Pr\{g_1(t) < W(t) < g_2(t),\ 0 \leqq t \leqq 1\} \tag{5.2}$$

for a large class of continuous functions g_1 and g_2. These results have been extended by Müller ([18], Section 4) to the semi-infinite time interval $[0, \infty)$. We briefly note that these results can be directly applied to get one-sided rate of convergence results of the same kind for the associated discounted processes.

We say that $u \leqq v$ in $D[a, b]$ if $u(t) \leqq v(t)$ for all t in $[a, b]$. Now note that f: $D[a, b] \to D[a, b]$ is monotonic: if $u \leqq v$, then $f(u) \leqq f(v)$. Hence,

$$\begin{aligned}&\{x \mid g_1(t) \leqq x(t) \leqq g_2(t),\ a \leqq t \leqq b\}\\ &\quad \subseteq \{x \mid f(g_1)(t) \leqq f(x)(t) \leqq f(g_2)(t),\ a \leqq t \leqq b\},\end{aligned} \tag{5.3}$$

so that, with T_n in (4.4) and appropriate functions g_1 and g_2, there exists a constant A such that

$$\begin{aligned}&\Pr\{f(g_1)(t) \leqq T_n(t) \leqq f(g_2)(t),\ a \leqq t \leqq b\}\\ &\quad \geqq \Pr\left\{g_1(t) \leqq \frac{S^n_{[nt]} - \mu_n^{[nt]}}{(n\sigma^2)^{\frac{1}{2}}} \leqq g_2(t),\ a \leqq t \leqq b\right\}\\ &\quad \geqq \Pr\{g_1(t) \leqq W(t) \leqq g_2(t),\ a \leqq t \leqq b\} - A n^{-\frac{1}{4}} \log n.\end{aligned} \tag{5.4}$$

The inequality the other way does not follow by this argument because $f^{-1}: D \to D$ is not monotonic.

6. Other Summation Methods

The previous sections can be generalized to yield stochastic analogues of other kinds of summation. Let U_n be defined in (2.11), but instead of (2.12), let D_n be defined by

$$D_n(t) = \phi(n)^{-1} \int_0^{nt} h(v/n)\, dS^n(v), \qquad 0 \leqq t \leqq r, \tag{6.1}$$

where h: $[0, r] \to R$ has a continuous derivative h' on $[0, r]$. Integration by parts yields

$$D_n(t) = h(t)\, U_n(t) - \int_0^t h'(v)\, U_n(v)\, dv, \qquad 0 \leqq t \leqq r. \tag{6.2}$$

Continuity for f in Lemma 2.1 extends to this case if f is now defined by

$$f(x)(t) = h(t)\, x(t) - \int_0^t h'(v)\, x(v)\, dv, \qquad 0 \leqq t \leqq r. \tag{6.3}$$

To get a continuous f^{-1}, we also require that $h(t) \neq 0$ on $[0, r]$. Then f^{-1} is also given by (6.3), but with $h(t)$ replaced by $1/h(t)$ everywhere, which means $h'(v)$ is replaced by $-h'(v)/h(t)^2$.

For example, $h(t) = 1 - t$, $0 \leqq t \leqq 1$, corresponds to a stochastic analogue of Cesàro $(C, 1)$ convergence:

$$D_n(t) = (1 - t)\, U_n(t) - \int_0^t U_n(v)\, dv, \qquad 0 \leqq t \leqq 1, \tag{6.4}$$

and

$$D_n(1) = \int_0^1 U_n(v)\,dv = n^{-1} \int_0^n [S^n(v)/\phi(n)]\,dv. \tag{6.5}$$

If $U_n \Rightarrow W$ in $D[0,1]$, then $D_n(1) \Rightarrow N(0, \frac{1}{3})$, where the distribution of the limit is easily obtained from the case of a single sequence of i.i.d. variables $\{X_n, n \geqq 1\}$ with mean 0 and variance 1; then

$$D_n(1) = n^{-\frac{3}{2}} \sum_{k=1}^{n} S_k = n^{-\frac{3}{2}} \sum_{k=1}^{n} (n-k) X_k. \tag{6.6}$$

Since $h(1)=0$, we cannot construct the inverse map f^{-1} on $D[0,1]$. For example, if $x(t)=t$, $0 \leqq t \leqq 1$, then $f^{-1}(x)(t) = 1 - \log|1-t|$, $0 \leqq t \leqq 1$, which blows up at $t=1$. The inverse exists of course on $D[0,s]$ for $s<1$. The situation is just as before. For $s<1$, there is equivalence on $D[0,s]$ for functional limit theorems of the ordinary $(C,0)$ kind and the Cesàro $(C,1)$ kind, but on the space $D[0,1]$, which is necessary in order to apply the projection, we can only go one way, just as in the deterministic case.

Another possible stochastic analogue of Cesàro $(C,1)$ convergence is obtained by setting $D_n = g(U_n)$, where

$$g(x)(t) = \int_0^t s^{-\frac{1}{2}} x(s)\,ds, \qquad t \geqq 0. \tag{6.7}$$

The function g in (6.7) is continuous from (D,d) to (D,d) but not from (L,m) to (L,m). With a single sequence of random variables, instead of (6.6), we have

$$D_n(1) = n^{-1} \sum_{k=1}^{n} k^{-\frac{1}{2}} S_k = n^{-1} \sum_{k=1}^{n} X_k \sum_{k=1}^{n} k^{-\frac{1}{2}}, \tag{6.8}$$

so that $D_n(1) \Rightarrow N(0, \frac{2}{3})$ if $U_n \Rightarrow W$ in (D,d). As before, the general theme is the possibility of obtaining stochastic analogues of deterministic summation methods by exploiting continuous mapping theorems in the function space setting.

Acknowledgments. The author is grateful to J. Sethuraman, J. G. Wendel, W. Vervaat, and the referees for pointing out references and errors and otherwise helping to make this a better paper.

References

1. Billingsley, P.: Convergence of Probability Measures. New York: John Wiley and Sons 1968.
2. Chernoff, H.: Estimation of the mode. Ann. Inst. Statist. Math. **16**, 31–41 (1964).
3. Denardo, E. V.: Markov renewal programs with small interest rate. Ann. Math. Statistics **42**, 477–496 (1971).
4. Doob, J.: Stochastic Processes. New York: Wiley 1953.
5. Dudley, R. M.: Distances of probability measures and random variables. Ann. Math. Statistics **39**, 477–496 (1968).
6. Freedman, D.: Some invariance principles for functionals of a Markov chain. Ann. Math. Statistics **38**, 1–7 (1967).
7. Gapoškin, V. F.: The law of the iterated logarithm for Cesàro's and Abel's method of summation. Theor. Probab. Appl. **10**, 411–420 (1965).
8. Gerber, H. U.: The discounted central limit theorem and its Berry-Esséen analogue. Ann. Math. Statistics **42**, 389–392 (1971).

9. Hanson, D.L. and F.T. Wright: Some convergence results for weighted sums of independent random variables. Z. Wahrscheinlichkeitstheorie verw. Geb. **19**, 81–89 (1971).
10. Hatori, H.: On Markov chains with rewards. Kōdai Math. Sem. Reports **18**, 184–192 (1966).
11. Hatori, H.: On continuous time Markov processes with rewards. Kōdai Math. Sem. Reports **18**, 212–218 (1966).
12. Iglehart, D.L. and D.P. Kennedy: Weak convergence of the average of flag processes. J. Appl. Probab. **7**, 747–753 (1970).
13. Iglehart, D. and W. Whitt: The equivalence of functional central limit theorems for counting processes and associated partial sums. Ann. Math. Statistics **42**, 1372–1378 (1971).
14. Lamperti, J.: On the convergence of stochastic processes. Trans. Amer. Math. Soc. **104**, 430–435 (1962).
15. Liggett, T.: An invariance principle for conditional sums of independent random variables. J. Math. Mech. **18**, 559–570 (1968).
16. Liggett, T.: Weak convergence of conditioned sums of independent random vectors. Trans. Amer. Math. Soc. **152**, 195–213 (1970).
17. Loève, M.: Probability Theory. 3rd. ed. Princeton, New Jersey: Van Nostrand 1963.
18. Müller, D.W.: Verteilungs-Invarianzprinzipien für das starke Gesetz der großen Zahl. Z. Wahrscheinlichkeitstheorie und verw. Geb. **10**, 173–192 (1968).
19. Paley, R.E.A.C., N. Wiener, and A. Zygmund: Note on random functions. Math. Z. **37**, 647–668 (1933).
20. Parthasarathy, K.R.: Probability Measures on Metric Spaces. New York: Academic Press 1967.
21. Prohorov, Yu.: Convergence of random processes and limit theorems in probability theory. Theor. Prob. Appl. **1**, 157–214 (1956).
22. Pyke, R. and R. Schaufele: Limit theorems for Markov renewal processes. Ann. Math. Statistics **35**, 1746–1764 (1964).
23. Pyke, R. and G.R. Shorack: Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage Theorems. Ann. Math. Statistics **39**, 755–771 (1968).
24. Pyke, R.: Applications of almost surely convergent constructions of weakly convergent processes. Proc. Int. Symp. Prob. and Inf. Theory, pp. 187–200. Berlin-Heidelberg-New York: Springer 1969.
25. Skorohod, A.V.: Limit theorems for stochastic processes. Theor. Probab. Appl. **1**, 261–290 (1956).
26. Skorohod, A.V.: Limit theorems for stochastic processes with independent increments. Theor. Prob. Appl. **2**, 138–171 (1957).
27. Skorohod, A.V.: Studies in the Theory of Random Processes. Reading, Mass.: Addison Wesley 1965.
28. Stone, C.: Weak convergence of stochastic processes defined on semi-infinite time intervals. Proc. Amer. Math. Soc. **14**, 694–696 (1963).
29. Stone, C.: Limit theorems for random walks, birth and death processes, and diffusion processes. Illinois J. Math. **7**, 638–660 (1963).
30. Strassen, V.: An invariance principle for the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie und verw. Geb. **3**, 211–226 (1964).
31. Veinott, A.F., Jr.: Discrete dynamic programming with sensitive discount optimality criteria. Ann. Math. Statistics **40**, 1635–1660 (1969).
32. Whitt, W.: Weak convergence of first passage time processes. J. Appl. Probab. **8**, 417–422 (1971).
33. Wichura, M.J.: On the construction of almost uniformly convergent random variables with given weakly convergent image laws. Ann. Math. Statistics **41**, 284–291 (1970).
34. Widder, D.F.: The Laplace Transform. Princeton, New Jersey: Princeton University Press, 1946.
35. Woodroofe, M.: On the weak convergence of stochastic processes without discontinuities of the second kind. Z. Wahrscheinlichkeitstheorie und verw. Geb. **11**, 18–25 (1968).

Ward Whitt
Department of Administrative Sciences
Yale University
2 Hillhouse Avenue
New Haven, Connecticut 06520
USA

(Received June 17, 1971)