SOME LIMIT THEOREMS AND INEQUALITIES FOR WEIGHTED
AND NON-IDENTICALLY DISTRIBUTED EMPIRICAL PROCESSES

by

Kenneth Sidney Alexander
B.S., University of Washington
(1979)

SUBMITTED TO THE DEPARTMENT OF MATHEMATICS
IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS OF THE DEGREE OF
DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1982

© Kenneth Sidney Alexander

The author hereby grants to M.I.T. permission to reproduce
and to distribute copies of this thesis document in whole
or in part.

Signature of Author............................................................... Department of Mathematics, August 6, 1982

Certified by................................................................. Richard M. Dudley
Thesis Supervisor

Accepted by................................................................. Nesmith C. Ankeny
Chairman, Departmental Graduate Committee
2.

SOME LIMIT THEOREMS AND INEQUALITIES FOR WEIGHTED AND NON-IDENTICALLY DISTRIBUTED EMPIRICAL PROCESSES

by

KENNETH SIDNEY ALEXANDER

Submitted to the Department of Mathematics on August 6, 1982 in partial fulfillment of the requirements for the degree of Doctor of Philosophy

ABSTRACT

Sharp bounds in probability are derived for the supremum of the normalized empirical process $v_n$ over classes of sets satisfying a combinatorial condition, including the non-identically distributed case. From this is derived a bounded law of the iterated logarithm for this supremum, and a central limit theorem for truncated weighted empirical processes. In the identically distributed case with law $P$, conditions are found on $q$ so that an invariance principle holds for the weighted process $v_n/q^*P$, where $q$ is a function on $[0,1]$. As a corollary, a functional law of the iterated logarithm for $v_n/q^*P$, and necessary and sufficient conditions on $q$ for the central limit theorem to hold, are obtained.

Thesis Supervisor: Richard M. Dudley

Title: Professor of Mathematics
TABLE OF CONTENTS

1. Introduction 5
2. Background and Preliminaries 12
3. Exponential Bounds for the Suprema of Normalized Empirical Processes 32
4. A Bounded LIL for the Normalized Empirical Process 68
5. Weighted Empirical Processes with Truncation 81
6. Weighted Empirical Processes without Truncation 98
Index of Notation 133
References 135
ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Professor Richard M. Dudley, for his devotion of much time and effort to guidance of my work and careful checking for my errors. I am grateful to everyone else with whom I consulted in the course of this project, especially Professors Walter Philipp, Ronald Pyke, and Galen Shorack.

My fiancée Chris Czarnecki and my family provided moral support and encouragement throughout this endeavor, heightening my successes and cushioning my failures, making my efforts a pleasure.

I am indebted to Professor Bruce Erickson for fostering my original interest in probability, and to the National Science Foundation which supported me financially as I pursued that interest.

I wish to thank Phyllis Ruby for her excellent work typing this thesis.
I. Introduction

Given a sequence $X_1, X_2, \ldots$ of independent random variables taking values in a measurable space $(X, \mathcal{A})$ with laws $\mathcal{L}(X_i) = P(i)$, we consider for each $A \in \mathcal{A}$ the fraction $P_n(A)$ of the first $n$ r.v.'s (random variables) which fall in $A$. Our particular interest is in the deviation of $P_n(A)$ from its expected value

$$\bar{P}_n(A) := \frac{1}{n} \sum_{i=1}^{n} P(i)(A)$$

as $A$ varies over some subclass $C$ of $\mathcal{A}$, especially for large $n$. (We use "=" to mean "equals by definition.") We call $\nu_n := n^{1/2}(P_n - \bar{P}(n))$ the $n$th normalized empirical process for $\{P(i), i \geq 1\}$. We can view $\nu_n$ as a stochastic process indexed by $C$, or equivalently as a random element of a space of functions on $C$. We seek two main types of results: first, bounds on the tail of $\sup_{C \in \mathcal{C}} |\nu_n(C)|$, and second, limit theorems for $\nu_n$, including CLT's (central limit theorems) and LIL's (laws of the iterated logarithm). The first type of result is usually the key to proving the second.

If $C$ is a finite set and all $P(i) = \text{some } P$, then by the finite dimensional CLT we have

$$(1.1) \quad \mathcal{L}(\nu_n(C_1), \ldots, \nu_n(C_k)) \Rightarrow N(0, \Sigma)$$

in $\mathbb{R}^k$ as $n \to \infty$, where $\Sigma_{ij} := P(C_i \cap C_j) - P(C_i)P(C_j)$. 
6.

Hence, letting \( \alpha := \max_{1 \leq i \leq n} \sum_{i} = \sup_{C} \text{var} \left( \nu_{n}(C) \right) \), we have for all \( M > 0 \):

\[
\lim_{n \to \infty} \Pr \left[ \sup_{C} |\nu_{n}(C)| > M \right] \leq \sum_{i=1}^{k} \lim_{n \to \infty} \Pr \left[ |\nu_{n}(C_{i})| > M \right] = K \exp \left( -\frac{M^2}{2\alpha} \right)
\]

where \( K \) is a constant depending on \( k \) and \( \phi \) is the distribution function of the normal law \( N(0,1) \).

When we allow \( C \) to be infinite, however, the situation is more complex. If, for example, \( X \) is \([0,1]\), \( C \) is the Borel sets in \( X \), and all \( P^{(i)} \) are the uniform law \( P \) on \([0,1]\), then the (random) set \( C := \{X_{1}, \ldots, X_{n}\} \) satisfies \( P_{n}(C) = 1, P(C) = 0 \), so \( |\nu_{n}(C)| = n^{1/2} \). Hence \( \sup_{C} |\nu_{n}(C)| = n^{1/2} \) a.s. Vapnik and Červonenkis (1971), though, showed that if \( C \) satisfies a certain combinatorial condition, stated by saying \( C \) is a Vapnik-Červonenkis (or VC) class and satisfied by many classes of interest in applications, then

\[
(1.3) \quad \Pr \left[ \sup_{C} |\nu_{n}(C)| > M \right] \leq 4(n^{V} + 1) \exp \left( -\frac{M^2}{8} \right)
\]

for \( M \geq 2^{1/2} \), where \( v \) is a constant depending on \( C \).
7.

Devroye (1982) improved (1.3) to

\[
\Pr[\sup_{C} |\nu_{n}(C)| > M] \leq K(n^{V} + 1)\exp(-2M^{2})
\]

where \( K \) is a universal constant. But (1.3) and (1.4) cannot be improved to an inequality as good as (1.2), for Kiefer and Wolfowitz (1958) show that for the VC class \( \mathcal{C} := \{(-\infty, x]: x \in \mathbb{R}^{2}\} \), there is a law \( P \) which gives

\[
\lim_{n \to \infty} \Pr[\sup_{C} |\nu_{n}(C)| > M] \sim 8M^{2}\exp(-M^{2}/2\alpha)
\]

as \( M \to \infty \), where again \( \alpha := \sup_{C} \text{var}(\nu_{n}(C)) \). Kiefer (1961) did prove the next best thing to (1.2) in a special case: he showed that when \( \mathcal{C} := \{(-\infty, x]: x \in \mathbb{R}^{d}\} \) and all \( P_{(i)} \) are equal to some \( P \), for each \( \varepsilon > 0 \) there is a constant \( K \) depending on \( d \) and \( \varepsilon \) such that

\[
\Pr[\sup_{C} |\nu_{n}(C)| > M] \leq K \exp(-(2-\varepsilon)M^{2})
\]

for all \( M > 0 \). (1.5) is nearly as good as (1.2) if \( \alpha \) is 1/4, its largest possible value since

\[
\sup_{C} \text{var}(\nu_{n}(C)) = \sup_{C} P(C)(1 - P(C)) \leq 1/4.
\]

The inequalities (1.3) and (1.4) actually require some measurability assumptions since the events they refer to
may not be measurable in general. We omit these here and in the remainder of this introduction.

All of this motivates us to seek generalizations of (1.5) of the form

$$\Pr[\sup_{C} |v_n(C)| > M] \leq K \exp\left(-(1-\varepsilon)M^2/2\alpha\right)$$

for all $\varepsilon > 0$. We further seek to include the non-identically distributed case, where the $P(i)$ may be distinct. These goals will be accomplished in Chapter 3.

Returning to (1.1), we are led to ask whether the processes $v_n$ on a possible infinite class $C$ converge in law to a Gaussian process, i.e. whether a CLT holds. This information would enable us to find the limiting distribution of many functions of $v_n$. Whether or not the CLT holds depends of course on the $\sigma$-algebra on which we wish the weak convergence of laws to occur; a reasonable choice turns out to be the $\sigma$-algebra generated by all balls in the space of functions on $C$ in which we view $v_n$ as taking values. Dudley (1978) showed this convergence does occur in the identically distributed case when $C$ is a VC class. In Chapter 5 we generalize this result to the non-identically distributed case.
Kuelbs and Dudley (1980) showed that \( \{ v_n \} \) satisfies a compact LIL in the identically distributed case when \( \mathcal{C} \) is a VC class. In the non-identically distributed case there may be no compact LIL, but we can look for a bounded LIL as follows. If \( \mathcal{C} \) is a finite class \( \{ C_1, \ldots, C_k \} \), and \( \sum_{i=1}^{\infty} P_i(C_j) = \infty \) for each \( j \leq k \), then by a classical LIL for real-valued r.v.'s (see Stout (1974)),

\[
\limsup_{n \to \infty} \frac{|v_n(C_j)|}{n (2\sigma_{nj}^2 \log \log (n\sigma_{nj}^2))^{1/2}} = 1 \quad \text{a.s.}
\]

where \( \sigma_{nj}^2 = \text{var} (v_n(C_j)) \), so

\[
(1.6) \quad \limsup_{n \to \infty} \sup_{C \in \mathcal{C}} \frac{|v_n(C)|}{n (2\alpha_n \log \log (n\alpha_n))^{1/2}} = 1 \quad \text{a.s.}
\]

where \( \alpha_n := \max_j \sigma_{nj}^2 = \sup_{C \in \mathcal{C}} \text{var} (v_n(C)) \). In Chapter 4 we generalize (1.6) to possibly infinite VC classes \( \mathcal{C} \).

Even more detailed information about \( v_n \) may be obtained by weighting it at each set \( C \in \mathcal{C} \) according to the size of \( C \), that is, by studying the weighted empirical process \( v_n(C)/q(\bar{P}_n(C)), C \in \mathcal{C} \), where \( q \) is a non-negative function on \([0,1]\). If \( q \) is taken to be small when \( \bar{P}_n(C) \) is small, then the process
\( \nu_n / q \circ \overline{P}(n) \) in effect magnifies the normalized deviation 
\( \nu_n(C) \) of \( P_n(C) \) from \( \overline{P}(n)(C) \) for small sets \( C \), where the true deviation is likely to be small. A natural question to ask is: given \( \{P_i\} \) and a VC class \( \mathcal{C} \), for which \( q \) do the processes \( \nu_n / q \circ \overline{P}(n) \) satisfy a CLT or compact LIL? In the identically distributed case, if \( q \) is continuous and positive this CLT follows easily from Dudley's (1978) CLT for the unweighted processes \( \nu_n \). On the other hand, if \( q(t) \to 0 \) too fast as \( t \to 0 \) or 1, or if \( q(t) = 0 \) for some \( 0 < t < 1 \), then \( \nu_n / q \circ \overline{P}(n) \) may be very large on some sets with high probability, precluding any CLT or LIL. Thus the cases of interest are those \( q \) for which \( q(t) > 0 \) for \( t \in (0,1) \) and \( q(t) \to 0 \) as \( t \to 0 \) or as \( t \to 1 \), and the question is how fast \( q(t) \) can approach 0 as \( t \to 0 \) or 1 if we wish the CLT or LIL to hold. O'Reilly (1974) showed that for \( \mathcal{C} := \{[0,x]: x \in [0,1]\} \), all \( P_i \) uniform on \([0,1]\), and \( q \) satisfying certain regularity assumptions, the processes \( \nu_n / q \circ \overline{P}(n) \) satisfy a CLT if and only if

\[
\int_0^1 \exp \left( -\varepsilon q^2(t)/t \right) \frac{dt}{t} < \infty \text{ for all } \varepsilon > 0
\]

and
11.

\[ \int_0^1 \exp \left( - \frac{q^2(1-t)}{t} \right) \frac{dt}{t} < \infty \quad \text{for all} \quad \varepsilon > 0. \]

James (1975) showed that in the same setting, \( \frac{v_n}{q_0 \bar{F}(n)} \) satisfies a compact LIL if and only if

\[ \int_0^1 \frac{dt}{q^2(t) \log \log \left( \frac{1}{(t(1-t))} \right)} < \infty. \]

In Chapters 5 and 6 we extend these results to more general classes \( \mathbb{C} \) and laws \( P(i) \).
II. Background and Preliminaries

Throughout this thesis we shall work in and assume the following context: \((X, \mathcal{A})\) is a measurable space, \(\mathcal{C} \subset \mathcal{A}\), and \(\{P_i\}, i \geq 1\) is a sequence of laws (i.e. probability measures) on \((X, \mathcal{A})\). In the identically distributed case, i.e. when all \(P_i\) are equal, we call this law \(P\).

Let

\[ P := \prod_{i=1}^{\infty} P_i \]

be the product measure on \((X^\infty, \mathcal{A}^\infty)\). The \(n\)th empirical measure \(P_n\) is defined by

\[ P_n(A) := \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}(A), \ A \in \mathcal{A} \]

where \(X_i\) is the \(i\)th coordinate function on \(X^\infty\). Thus \(P_n\) is a function from \((X^\infty, \mathcal{A}^\infty)\) to the space \(\ell^\infty(\mathcal{A})\) of all bounded functions on \(\mathcal{A}\) with the sup norm, and \(P_n(A)\) is the fraction of the first \(n\) coordinates \(X_i\) which fall in \(A\). If we put a law on \((X^\infty, \mathcal{A}^\infty)\), then \(P_n\) and each \(X_i\) become r.v.'s. For example, on \((X^\infty, \mathcal{A}^\infty, P)\) we can view \(P_n\) as a measure obtained from
independent random points $X_1, \ldots, X_n$ sampled according to laws $P(1), \ldots, P(n)$ respectively. In general if $H(1), H(2), \ldots$ are any laws on $(X, \mathcal{A})$, $\mathbb{H} := \prod_{i=1}^{\infty} H(i)$, and $B \subset \lambda^\infty(\mathcal{A})$, then $\mathbb{H}[P_n \in B]$ means the probability that $P_n$ is in the set $B$ when each $X_i$ is an r.v. with law $H(i)$.

Note that when confusion is possible we always use subscripts in parentheses for non-random objects.

Next define, as in the introduction,

$$\overline{P}(n) := \frac{1}{n} \sum_{i=1}^{n} P(i)$$

and

$$\nu_n := n^{1/2}(P_n - \overline{P}(n)),$$

so $\overline{P}(n)(A) = E \overline{P}(n)(A)$ for $A \in \mathcal{A}$. We wish to study the restriction of $\nu_n$ to classes $\mathcal{C} \subset \mathcal{A}$. Given a finite subset $F$ of $X$ and a subset $E$ of $F$, we say $E$ is cut out by $\mathcal{C}$ from $F$ if $E = F \cap C$ for some $C \in \mathcal{C}$. Let $\Delta^\mathcal{C}(F)$ be the number of subsets of $F$ cut out by $\mathcal{C}$, and

$$m^\mathcal{C}(n) := \max \{\Delta^\mathcal{C}(F): F \subset X, \text{card}(F) = n\}, \ n \geq 1.$$
14.

Then $m^C(n) \leq 2^n$ for all $n$. We say $C$ is a Vapnik-Červonenkis (or VC) class if

\begin{equation}
(2.1) \quad m^C(n) < 2^n \quad \text{for some } \ n \geq 1
\end{equation}

and define the Vapnik-Červonenkis index $V(C)$ of $C$ to be the least $n \geq 1$ such that the inequality in (2.1) holds. Thus a VC class is one which does not contain enough sets to cut out all subsets of any $n$-element subset of $X$ when $n$ is large.

The concept of a VC class is quite a general one, as the following lemmas from Dudley (1978) show.

**Lemma 2.1.** Let $\mathcal{F}$ be a $d$-dimensional ($d < \infty$) vector space of real-valued functions on $X$, and $C := \{ \{ x : f(x) > 0 \} : f \in \mathcal{F} \}$. Then $C$ is a VC class with $V(C) = d + 1$.

**Lemma 2.2.** Let $C$ be a VC class and $k < \infty$. Let $\mathcal{G} := \cup \{ \sigma(C_1, \ldots, C_k) : C_i \in C \text{ for all } i \leq k \}$, where $\sigma(C_1, \ldots, C_k)$ is the algebra generated by $C_1, \ldots, C_k$. Then $\mathcal{G}$ is a VC class.

Note that $\mathcal{G}$ consists of those sets which can be formed from at most $k$ sets in $C$ by the operations of union, intersection, and complementation.
15.

Perhaps the most significant fact about VC classes is the following, due to Vapnik and Červonenkis (1971).

**Lemma 2.3.** If $\mathcal{C}$ is a VC class then $m^{\mathcal{C}}(n) \leq n^{V(\mathcal{C})} + 1$ for all $n \geq 1$.

Thus for any class $\mathcal{C} \subset \mathcal{A}$, either $m^{\mathcal{C}}(n) = 2^n$ for all $n$ (if $\mathcal{C}$ is not a VC class) or $m^{\mathcal{C}}(n)$ grows only like a power of $n$ (if $\mathcal{C}$ is a VC class.)

We use $A \setminus B$ to denote the set difference $A \cap B^c$.

A special case of Lemma 2.2 is the following:

**Lemma 2.4.** Let $\mathcal{C}$ be a VC class, and $v \geq V(\mathcal{C})$. Let $\mathcal{C}_1$ be a subset of $\{A \setminus B: A, B \in \mathcal{C}\}$. Let $v_1$ be the least positive integer such that $(v_1^{V(\mathcal{C})} + 1)^2 < 2^{v_1}$. Then $\mathcal{C}_1$ is a VC class with $V(\mathcal{C}_1) \leq v_1$.

**Proof.** This follows from Lemma 2.3 and the fact that

$\Delta_1^\mathcal{C}(F) \leq (\Delta^\mathcal{C}(F))^2$ for all finite $F \subset X$.

Using Lemmas 2.1 and 2.3 we can show the following are each VC classes:

1. Any subset of a VC class
2. A finite union of VC classes
3. All rectangles in $\mathbb{R}^d$
16.

(4) All ellipsoids in \( \mathbb{R}^d \)
(5) All half spaces in \( \mathbb{R}^d \)
(6) The class \( \{(-\infty,x]: x \in \mathbb{R}^d\} \) of all lower rectangles in \( \mathbb{R}^d \), where \( (-\infty,x] = \prod_{i=1}^{d} (-\infty, x_i] \).

Given a law \( H \) on \( (X, \mathcal{A}) \), define the pseudometric \( d_H \) on \( \mathcal{A} \) by \( d_H(A,B) := H(A \Delta B) \). Then for \( \epsilon > 0 \) and \( \mathcal{C} \subset \mathcal{A} \) define

\[
N(\epsilon, \mathcal{C}, H) := \min \{k: \text{There exist } C_1, \ldots, C_k \in \mathcal{C} \text{ such that } \min_{i<k} d_H(C,C_i) < \epsilon \text{ for all } C \in \mathcal{C} \}.
\]

The most important property of VC classes for our purposes is the following modified version of Lemma 7.13 of Dudley (1978).

**Lemma 2.5.** There exists an increasing sequence of constants \( \{A(v), v \geq 1\} \) such that if \( \mathcal{C} \) is a VC class and \( v \geq V(\mathcal{C}) \) then

\[
N(\epsilon, \mathcal{C}, H) \leq A(v) \epsilon^{-2v}
\]

for all laws \( H \) and \( 0 < \epsilon \leq 1 \). Hence in particular \( \mathcal{C} \) is totally bounded for \( d_H \).

Note that the bound in Lemma 2.5 does not depend on the law \( H \).
17.

We turn our attention next to the convergence in law of \( \{v_n, n \geq 1\} \) and related processes. Let \( \mathcal{C} \) be a VC class. For \( P \) a law on \((X, \mathcal{A})\) let \( W_P \) be a mean-zero Gaussian process indexed by \( \mathcal{C} \) with covariance \( \sigma(A, B) := P(A \cap B) \), and let \( G_P(C) := W_P(C) - P(C)W_P(X) \), \( C \in \mathcal{C} \). Then \( G_P \) has covariance \( \rho(A, B) := P(A \cap B) - P(A)P(B) \). It follows from Theorem 1.1 of Dudley (1973) and Lemma 2.5 that we may take \( G_P \) to have bounded uniformly \( d_P \)-continuous sample paths on \( \mathcal{C} \). As discussed in the introduction, \( G_P \) is the proper limit in law to seek for \( \{v_n\} \) in the identically distributed case. \( G_P \) and \( v_n \) are random elements of \( \ell^\infty(\mathcal{C}) \), so we look for convergence in law of \( v_n \) to \( G_P \) in a subspace \( D \) of \( \ell^\infty(\mathcal{C}) \) large enough to contain \( v_n \) and \( G_P \) almost surely. Such a \( D \) must in general be non-separable if \( X \) is uncountable, for then the set \( \{\delta_x - P: x \in X\} \) of possible values for \( v_1 \) may form an uncountable discrete set in \( D \). It is consistent with standard set theory to assume that any law \( \mu \) on the Borel sets of a non-separable metric space is concentrated on some separable subspace, i.e. \( \mu(E) = 1 \) for some separable \( E \) (Marczewski and Sikorski (1948)), so the law \( \mathcal{L}(v_n) \) is not in general defined on all Borel sets in \( D \). Thus we must consider weak convergence on smaller \( \sigma \)-algebras. But the limit
law $\mathcal{L}(G_p)$ is concentrated on the space $C_b,u(C,d_p)$ of bounded uniformly $d_p$-continuous functions on $C$, which is separable since (by Lemma 2.5) $C$ is totally bounded. Hence we examine weak convergence to laws concentrated on separable subspaces.

Given a metric space $(S,d)$, a $\sigma$-algebra $\mathcal{A}$ of subsets of $S$, and a law $\mu$ and sequence of laws $\{\mu_n, n \geq 1\}$ on $(S,\mathcal{A})$, we say $\{\mu_n\}$ converges weakly to $\mu$ on $\mathcal{A}$ (or "in $(S,\mathcal{A})$") if $\int f d\mu_n \to \int f d\mu$ for all bounded continuous real-valued functions $f$ on $S$ which are measurable with respect to $\mathcal{A}$. If $Y$ and $\{Y_n, n \geq 1\}$ are $S$-valued r.v.'s measurable with respect to $\mathcal{A}$, we say $\{Y_n\}$ converges in law to $Y$ on $\mathcal{A}$ (or "in $(S,\mathcal{A})$") if the laws $\{\mathcal{L}(Y_n)\}$ converge weakly to $\mathcal{L}(Y)$ on $\mathcal{A}$. Let $\mathcal{B}_B(S,d)$ denote the Borel $\sigma$-algebra of $(S,d)$ and $\mathcal{B}_b(S,d)$ the $\sigma$-algebra generated by all balls $\{x: d(x,x_0) < r\}$ ($x_0 \in X, r > 0$) in $S$. (We write only $\mathcal{B}_B$ or $\mathcal{B}_b$ when no confusion is possible.)

If $E$ is a separable subspace of $S$ then $\mathcal{B}_b$ and $\mathcal{B}_B$ have the same relativization to $E$, that is,

$$(2.2) \quad \{A \cap E: A \in \mathcal{B}_b\} = \{A \cap E: A \in \mathcal{B}_B\}.$$  

We have $E \in \mathcal{B}_b$ if $E$ is also closed; in fact,
letting $T$ be a countable dense subset of $E$ and $B(x,r)$ the open ball of radius $r$ about $x \in S$, we have

$$E = \bigcap_{m=1}^{\infty} \bigcup_{t \in T} B(t, l/m) \in \mathcal{B}_b.$$  

Hence a law $\mu$ on $(S, \mathcal{B}_b)$ which is concentrated on a separable subspace of $S$ can be viewed as a law on $(S, \mathcal{B}_B)$, by setting $\mu(A) := \mu(A \cap E)$ for all $A \in \mathcal{B}_B$.

The following theorem is due to Wichura (1970). Another proof is in Fernandez (1974).

**Theorem 2.6.** Let $(S, d)$ be a metric space and let $\mathcal{B}_n$, $n \geq 1$, be $\sigma$-algebras with $\mathcal{B}_b \subset \mathcal{B}_n \subset \mathcal{B}_B$ for all $n$. Let $\mu_n$ be a law defined on $\mathcal{B}_n$ for each $n$, and let $\mu$ be a law on $\mathcal{B}_b$ which is concentrated on some separable subspace of $S$. Suppose $\mu_n$ converges weakly to $\mu$ on $\mathcal{B}_b$. Then there exist r.v.'s $Y$ and $Y_n$, $n \geq 1$, all defined on the same probability space, such that $Y_n$ is measurable into $\mathcal{B}_n$, $\mathcal{L}(Y_n) = \mu_n$, $Y$ is measurable into $\mathcal{B}_B$, $\mathcal{L}(Y) = \mu$, and $Y_n \rightarrow Y$ a.s. \[\square\]

**Corollary 2.7.** Let $(S, d)$ be a metric space and $\mathcal{B}$ a $\sigma$-algebra with $\mathcal{B}_b \subset \mathcal{B} \subset \mathcal{B}_B$. Let $\mu$ and $\mu_n$, $n \geq 1$, be laws on $(S, \mathcal{B})$. If $\mu_n \rightarrow \mu$ weakly on $\mathcal{B}_b$, then $\mu_n \rightarrow \mu$ weakly on $\mathcal{B}$. 
20.

Proof. Let \( Y \) and \( Y_n, n \geq 1 \), be as in Theorem 2.6, with \( B_n = B \) for all \( n \). If \( f \) is a bounded continuous function on \( S \) measurable with respect to \( B \), then by bounded convergence, \( \int fd\mu_n = Ef(Y_n) \to Ef(Y) = \int fd\mu \). □

This corollary justifies restricting our attention to weak convergence on \( B_b \).

Given a set \( A \) in a metric space \((S,d)\) and \( \epsilon > 0 \), let \( A^\epsilon \) denote the set \( \{x \in S: d(x,A) < \epsilon\} \).

If \( \mu \) is a law on \((S,B_S)\), a set \( A \in B_B \) is called a **continuity set** for \( \mu \) if \( \mu(\partial A) = 0 \), where \( \partial A \) is the boundary of \( A \).

The next theorem is a characterization of weak convergence.

**Theorem 2.8.** Let \((S,d)\) be a metric space, \( E \) a closed separable subspace of \( S \), and \( \mu \) and \( \mu_n, n \geq 1 \), laws on \((S,B_B)\), with \( \mu \) concentrated on \( E \). Then the following are equivalent:

(i) \( \mu_n \to \mu \) weakly on \( B_b \);

(ii) \( \mu_n(A) \to \mu(A) \) for all \( A \in B_B \) which are continuity sets for \( \mu \);

(iii) \( \mu_n(A) \to \mu(A) \) for every \( A \) which is a finite union of open balls which are continuity sets for \( \mu \).
Proof. Suppose first that (i) holds. As discussed above, we may consider $\mu$ to be a law on $(S, \mathcal{B}_B)$. Let $A \in \mathcal{B}_b$ be a continuity set for $\mu$, and let $\varepsilon > 0$.

Then we can choose $m \geq 1$ such that $\mu((\overline{A} \cap E)^{1/m} \setminus (A \cap E)) < \varepsilon$. Define $f$ on $S$ by $f(x) := \max (1 - m d(x, \overline{A} \cap E), 0)$.

Since $\overline{A} \cap E$ is separable we have $(\overline{A} \cap E)^{c} \in \mathcal{B}_b$ for all $\varepsilon > 0$, so $f$ is $\mathcal{B}_b$-measurable, and

$$
(2.3) \quad \mu(A) \leq \int_S f d\mu \leq \mu(A) + \varepsilon.
$$

Let $F := \overline{A} \cap \{x: f(x) < 1 - \varepsilon\}$. Then $E \cap F = \emptyset$, so for each $x \in E$ there is a ball $B(x, r(x))$ contained in the complement $F^c$ of $F$. Let $\delta > 0$. Taking a countable subcover we have $E \subseteq \bigcup_{n=1}^{\infty} B(x_n, \delta r(x_n))$ for some sequence $\{x_n\} \subseteq E$. Choose $N$ such that

$$
\mu(\bigcup_{n=1}^{N} B(x_n, \delta r(x_n))) > 1 - \delta,
$$

and define $U_\delta := \bigcup_{n=1}^{N} B(x_n, \delta r(x_n))$. Define $g$ on $S$ by

$$
g(x) := \min (1, \min_{1 \leq n \leq N} \frac{d(x, x_n)}{r(x_n)}).
$$

Then $g$ is continuous, bounded and $\mathcal{B}_b$-measurable, $g \leq \delta$ on $U_\delta$, and $g = 1$ on $F$. Hence
22.

\[
\limsup_n \mu_n(A \cap \{x: f(x) \leq 1 - \varepsilon\}) \\
\leq \limsup_n \int_S gd\mu_n \\
= \int_S gd\mu \\
\leq \delta + \mu(U_\delta^c) \\
< 2\delta.
\]

It follows that \( \mu_n(A \cap \{x: f(x) \leq 1 - \varepsilon\}) \to 0 \) as \( n \to \infty \) so using (2.3) we see that

\[
\limsup_n \mu_n(A) = \limsup_n \mu_n(A \cap \{x: f(x) > 1 - \varepsilon\}) \\
\leq \frac{1}{1 - \varepsilon} \limsup_n \int_S fd\mu_n \\
= \frac{1}{1 - \varepsilon} \int_S fd\mu \\
\leq \frac{1}{1 - \varepsilon} (\mu(A) + \varepsilon).
\]

Since the same holds for \( A^c \) in place of \( A \), and \( \varepsilon \) is arbitrary, (ii) follows.

Clearly (ii) implies (iii), so we assume (iii) and prove (i). Let \( f \) be continuous, bounded, and \( \mathcal{B}_b \)-measurable on \( S \). Let \( T := \{t \in \mathbb{R}: \mu(\{x: f(x) = t\}) > 0\} \).

Then \( T \) is at most countable. Fix \( t \not\in T \), and let
23.

A:= \{x: f(x) > t\}. Then for each x ∈ A ∩ E there is an r(x) > 0 such that B(x,r(x)) ⊂ A and
μ(∪B(x,r(x))) = 0. Taking a countable subcover we have

that A ∩ E ⊂ ∪\limits_{n=1}^{∞} B(x_n,r(x_n)) for some sequence \{x_n\} ⊂ A ∩ E.

Fix ε > 0 and choose k < ∞ such that

\[ \mu(\bigcup_{i=1}^{k} B(x_i,r(x_i))) > \mu(A ∩ E) - \varepsilon. \]

By (iii) we have

\[ \liminf_{n} \mu_n(A) \geq \liminf_{n} \mu_n(\bigcup_{i=1}^{k} B(x_i,r(x_i))) \]

\[ = \mu(\bigcup_{i=1}^{k} B(x_i,r(x_i))) \]

\[ > \mu(A) - \varepsilon. \]

Since the same holds for A = \{x: f(x) < t\}, we have

(2.4) \ μ_n(\{x: f(x) > t\}) → μ(\{x: f(x) > t\}) for all t \not\in T.

From (2.4) it is easy to prove that \( \int f d\mu_n \to \int f d\mu, \)
and (i) follows.

The functions \( \Pi_C(\psi) := \psi(C), C \in \mathcal{C}, \) are called the coordinate functions. If \( \nu \) is a stochastic process indexed by \( \mathcal{C}, \) then the laws on \( \mathbb{R}^d \) of the r.v.'s
(v(C_1),...,v(C_d)), where d \geq 1 and C_1,...,C_d \in \mathcal{C}, are called the finite dimensional distributions of v.

A stochastic process v indexed by a topological space is called sample-continuous if some version of v has continuous sample paths, i.e. if there is a stochastic process with continuous sample paths and the same finite-dimensional distributions as v. For \delta, \varepsilon > 0 and H a law on (X, \mathcal{A}), let

\[ B_{\delta, \varepsilon}(H) := \{ f \in L^\infty(\mathcal{C}) : \text{There exist } A, B \in \mathcal{C} \text{ with } \]
\[ H(A \Delta B) < \delta, |f(A) - f(B)| > \varepsilon \} \text{.} \]

For any law \Pr let \Pr^* and \Pr_* denote the corresponding outer and inner measures respectively.

The following is a variant of Theorem 1.2 of Dudley (1978). The proof is essentially the same.

**Theorem 2.9.** Let \mathcal{C} \subset \mathcal{A} and let \psi_n, n \geq 1, be \ell^\infty(\mathcal{C})-valued r.v.'s on a probability space (\Omega, \mathcal{B}, \Pr) with \psi_n(C) measurable for each C \in \mathcal{C}. Let P be a law on (X, \mathcal{A}). Let D be a subspace of \ell^\infty(\mathcal{C}) containing \mathcal{C}_{b,u}(\mathcal{C}, d_p) and suppose each \psi_n is measurable into (D, \mathcal{B}_b). Then \psi_n converges in law on \mathcal{B}_b to a Gaussian process G with sample paths in \mathcal{C}_{b,u}(\mathcal{C}, d_p).
if

(2.5) for every $\varepsilon > 0$ there exist $\delta > 0$ and $n_0$ such
that $n > n_0$ implies $\Pr^*[\psi_n \in B_{\delta, \varepsilon}(P)] < \varepsilon$,

(2.6) $C$ is totally bounded for $d_p$,

and

(2.7) the finite dimensional distributions of $\psi_n$ converge

to those of $G$,

and only if (2.5) and (2.7) hold.

Using his similar theorem, Dudley (1978) showed that
for a particular choice of the space $D$, $\nu_n$ converges
in law to $G_p$ on $\mathcal{B}_b(D, d_p)$ in the identically dis-
tributed case, under certain measurability assumptions.
We refer the reader to Dudley's paper for details.

Measurability assumptions are in general needed
because, as discussed above, $\nu_n$ may not be measurable
into $(\ell^\infty(C), \mathcal{B}_B)$, and because such r.v.'s as $\sup_C |\nu_n(C)|$
may not be measurable. Hence we define here a measura-
bility condition to suit our purposes. Let

$$P_n := \frac{1}{n} \sum_{i=n+1}^{2n} \delta_{X_i},$$

$$\nu_n^0 := n^{1/2}(P_n - P'_n),$$
and let
\[ \mathbb{P}(n) := \prod_{i=1}^{n} \mathbb{P}(i) \]
be the product measure on \( (X^n, \mathcal{A}^n) \). We say a class \( \mathcal{D} \subseteq \mathcal{A} \) is \( n \)-deviation-measurable for \( \{ \mathbb{P}(i), i \geq 1 \} \) (or just "for \( \mathbb{P} \)" if all \( \mathbb{P}(i) = \) some \( \mathbb{P} \)) if

\[
\begin{align*}
(2.8) & \sup_{\mathcal{D}} |n^{1/2}(\mathbb{P}_n(D) - b\bar{\mathbb{P}}_n(D))|, \\
(2.9) & \sup_{\mathcal{D}} |\nu_0(D)|, \\
(2.10) & \sup_{\mathcal{D}} |\mathbb{P}_{2n}(D) - b\bar{\mathbb{P}}(n)(D)|
\end{align*}
\]

are each \( \mathbb{P}^2_n \)-completion measurable on \( (X^{2n}, \mathcal{A}^{2n}) \) when \( b = 1 \). We say \( \mathcal{D} \) is \( n \)-deviation measurable with constants for \( \{ \mathbb{P}(i), i \geq 1 \} \) if this measurability holds for all \( b > 0 \). The proof of the following lemma is obvious.

**Lemma 2.10.** Let \( \mathcal{D} \subseteq \mathcal{A} \) and let \( f, g \) be random elements of \( \ell^\infty(\mathcal{D}) \) such that \( f(D) \) and \( g(D) \) are measurable for each \( D \in \mathcal{D} \). Suppose that with probability 1 there exists a countable subset \( \mathcal{D}_0 \) of \( \mathcal{D} \) such that for each \( D \in \mathcal{D} \) there exists a sequence \( \{D_n\} \) in \( \mathcal{D}_0 \) with
27.

\[ f(D_n) + f(D) \text{ and } g(D_n) + g(D). \text{ Then } \sup_D |f(D) - g(D)| \]
is a measurable r.v. In particular, if these hypotheses are satisfied for \((f,g) = (P_n, b\overline{F}(n)), (P_n', P_n'), \text{ and } (P_{2n}, b\overline{F}(n))\) for all \(b > 0\), then \(D\) is \(n\)-deviation measurable with constants for \(\{P(i), i \geq 1\}\). \[\] 

Given a sequence \(\{y_n\}\) in a topological space, we let \(C(\{y_n\})\) denote the cluster set of \(\{y_n\}\), i.e. the set of all limits of subsequences of \(\{y_n\}\). Let \(B\) be a Banach space with norm \(\|\cdot\|_B\), \(Y\) a \(B\)-valued r.v., \(Y_1, Y_2, \ldots\) independent copies of \(Y\), \(S_n := \sum_{i=1}^{n} Y_i\), and \(w_n := (2n \log \log n)^{1/2}\). \(Y\) is said to satisfy the compact LIL in \(B\) if there is a compact set \(K\) in \(B\) such that

\[\lim_{n} \|S_n/w_n - K\|_B = 0 \text{ a.s.}\]

and

\[C(\{S_n/w_n\}) = K \text{ a.s.}\]

A discussion of the nature of \(K\) can be found in Goodman, Kuelbs, and Zinn (1980). Let \(C_B[0,1]\) denote the space of continuous functions from \([0,1]\) to \(B\) with norm \(\|f\|_C := \sup_{0 \leq t \leq 1} \|f(t)\|_B\). Define random elements \(\zeta_n\) of \(C_B[0,1]\) by
28.

$$\zeta_n(t) = S_{nt} \text{ if } t = k/n, \ k = 0, 1, \ldots, n$$

with $\zeta_n$ linear in between points $k/n$. $Y$ is said to satisfy the functional LIL in $B$ if there is a compact set $K$ in $C_B[0,1]$ such that $\lim_{n} \|\zeta_n/w_n - K\|_C = 0$ a.s. and $C(\{\zeta_n/w_n\}) = K$ a.s. By considering only $\zeta_n(1)$ we see that the functional LIL implies the compact LIL.

$Y$ is said to satisfy the CLT in $(B, \mathcal{B})$ if $L(S_n/n^{1/2})$ converges weakly on $\mathcal{B}$ to a Gaussian law.

When no confusion is possible we may abuse notation and say $\{S_n/n^{1/2}\}$ satisfies the compact or functional LIL or the CLT when $Y$ does.

A bounded LIL is any result which says a sequence $\{Y_n: n \geq 1\}$ of $B$-valued r.v.'s with partial sums $S_n$ satisfies $\limsup_{n} \|S_n\|/w_n < \infty$.

The compact LIL implies a bounded LIL, since the compact limit set $K$ is bounded. In fact, it may be shown (see Pisier (1975)) that if $B$ is separable and $Y$ satisfies the compact LIL in $B$, then

$$\sup \{\|x\|_B: x \in K\}$$

(2.11)

$$= \sup \{\text{var } (f(Y))^{1/2}: f \in B^*, f(x) \leq 1 \text{ for all } \|x\|_B \leq 1\},$$
29.

where $B^*$ is the dual of $B$. In particular, we have the following.

**Lemma 2.11.** If $C \subset Q$ is totally bounded and an r.v. $Y$ satisfies the compact LIL in $C_b, u(C, d_p)$, then

$$\lim \sup_{n} \frac{\|S_n\|/w_n}{\sup \{\text{var} \ (Y(C))^{1/2} : C \in C\}} \ a.s.$$  

The next lemma is from Kuelbs and Le Page (1973).

**Lemma 2.12.** A Gaussian r.v. taking values in a separable Banach space satisfies the CLT and the compact and functional LIL's.

This lemma makes the following invariance principle of Dudley and Philipp (1982) quite useful.

**Theorem 2.13.** Let $\mathcal{F} \subset L^2(X, \Omega, \mathbb{P})$ be a class of functions such that

1. (2.12) $\mathcal{F}$ is totally bounded in $L^2$, and

2. (2.13) for every $\varepsilon > 0$ there exists $\delta > 0$ and $n_0$ such that for $n \geq n_0$,

$$\mathbb{P}^* \left[ \sup \{|(f-g)dB_n| : f, g \in \mathcal{F}, \int (f-g)^2 d\mathbb{P} < \delta^2 \} > \varepsilon \right] < \varepsilon.$$  

Then there exists a sequence $\{Y_j, j \geq 1\}$ of independent
identically distributed Gaussian processes indexed by $\mathcal{F}$, with sample functions a.s. uniformly continuous on $\mathcal{F}$ with respect to the $L^2$ norm, such that

\[(2.14) \quad \mathbb{E} Y_1(f) = 0 \quad \text{for all } f \in \mathcal{F},\]

\[(2.15) \quad \mathbb{E} Y_1(f)Y_1(g) = \int fgdP - \int fdP \cdot \int ggdP \quad \text{for all } f,g \in \mathcal{F},\]

and

\[(2.16) \quad n^{-1/2} \max_{k \leq n} \sup_{f \in \mathcal{F}} \left| \sum_{j=1}^{k} f(X_j) - \int fdP - Y_j(f) \right| \to 0 \quad \text{as } n \to \infty\]

in probability, and in $L^p$ for all $p < 2$.

If in addition there exists a measurable function $F$ with $|f| \leq F$ for all $f \in \mathcal{F}$ and

\[(2.17) \quad \int F^2 dP < \infty,\]

then the convergence in (2.16) can be obtained in $L^2$ also. If

\[(2.18) \quad \int F^2 / \log \log F \, dP < \infty,\]

then the $Y_j$ can be chosen so that, instead of (2.16),

\[(2.19) \quad \sup_{f \in \mathcal{F}} \left| \sum_{j=1}^{n} f(X_j) - \int fdP - Y_j(f) \right| = o((n \log \log n)^{1/2}) \quad \text{a.s.}\]
The r.v. in (2.16) may not be measurable in general. Therefore we take convergence in probability of r.v.'s $Z_n$ to $Z$ to mean $\Pr^*([|Z_n-Z| > \varepsilon]) \to 0$ for all $\varepsilon > 0$, and convergence in $L^p$ to mean $\mathbb{E}W_n^p \to 0$ for some measurable r.v. $W_n \geq |Z_n-Z|.$

The Gaussian processes $Y_i$ in Theorem 2.13 are defined on the product of $(X^\omega, \mathcal{G}^\omega, \mathbb{P})$ and $[0,1]$ with Lebesgue measure.
III. Exponential Bounds for the Suprema of Normalized Empirical Processes

For $C \subseteq A$ and $n \geq 1$ define

$$
\alpha(C, n) := \sup_{C} \text{var}_P(v_n(C))
= \sup_{C} \frac{1}{n} \sum_{i=1}^{n} P(i)(C) (1 - P(i)(C)).
$$

For $x = (x_1, x_2, \ldots) \in X^\infty$ define

$$
x^{(n)} := (x_1, \ldots, x_n).
$$

Our main theorem of this chapter will be the following.

Theorem 3.1. Let $n \geq 1$, and let $C \subseteq A$ be a VC class such that

(3.1) $C$ is $k$-deviation measurable for $P(i)$ for all $i \leq n$ and $k \geq 1$,

and

(3.2) $\sup_{C} |v_n(C)|$ is measurable.

Let $v \geq V(C)$, $0 < \varepsilon \leq 1$, and $\alpha > 0$. Then there exist constants $M_0 = M_0(\varepsilon, \alpha, v)$, $n_0 = n_0(\varepsilon, \alpha, v)$, $K = K(v)$ such that if
(3.3) \( n \geq n_0, \alpha \geq \alpha(\mathcal{C},n), \) and \( M_0 \leq M < \begin{cases} 3n^{1/2}a\varepsilon/8 & \text{if } \alpha < \frac{1}{4} \\ \infty & \text{if } \alpha > \frac{1}{4} \end{cases} \)

then

\[ \mathbb{P}\left[ \sup_\mathcal{C} |v_n(C)| > M \right] \leq K \exp \left( -(1-\varepsilon)M^2/2a \right). \]

Given \( \delta > 0 \) there exists \( \alpha_0 = \alpha_0(\varepsilon,\nu,\delta) \) such that

(3.4) \( \alpha \leq \alpha_0 \) implies \( M_0(\varepsilon,\alpha,\nu) < \delta. \)

Since \( \alpha(\mathcal{C},n) \leq 1/4 \) always, this theorem is never vacuous, in that (3.3) is always satisfied for \( \alpha = 1/4 \) and \( M,n \) large.

The property (3.4) enables us to apply the theorem to classes of "small" sets, which is useful in proving statements like (2.5).

The following lemma uses ideas from Pollard (1981). In its proof, and several that follow, we use iterated integrals in place of conditional expectations to avoid measurability difficulties.

**Lemma 3.2.** Let \( H(1), \ldots, H(m) \) be laws on \((X,\mathcal{C}),\) \( H := \prod_{i=1}^m H(i), \) and \( H := H^2 \) on \((X^{2m},\mathcal{C}^{2m}). \) Let
$$H_m := \frac{1}{m} \sum_{i=1}^{m} \delta x_i, \quad H'_m := \frac{1}{m} \sum_{i=m+1}^{2m} \delta x_i, \quad \bar{H}(m) := \frac{1}{m} \sum_{i=1}^{m} H(i),$$

$$\mu_m := m^{1/2} (H_m - \bar{H}(m)), \quad \mu'_m := m^{1/2} (H'_m - \bar{H}(m)).$$

Let

$$D \subset \Omega, \quad y > 0, \quad \text{and} \quad \theta \geq \sup_{\mathcal{D}} \frac{1}{m} \sum_{i=1}^{m} H(i)(D)(1 - H(i)(D)).$$

Then

$$(3.5) \quad H[\sup_{\mathcal{D}} |\mu_m(D)| > y] \leq 2H[\sup_{\mathcal{D}} |\mu'_m(D)| > y - (2\theta)^{1/2}]$$

whenever both events in (3.5) are $H$-completion measurable.

**Proof.** Let $\mu'_m := m^{1/2} (H'_m - \bar{H}(m))$ and let $|| \cdot ||$ denote the sup norm for functions on $\mathcal{D}$. We have

$$\sup_{\mathcal{D}} E(\mu'_m(D)^2) = \sup_{\mathcal{D}} \frac{1}{m} \sum_{i=1}^{m} H(i)(D)(1 - H(i)(D)) \leq \theta$$

so by Chebyshev's inequality, if $x^{(m)} \in \{x^{(m)}: |\mu_m(D_0)| > y\}$ for some $D_0 \in \mathcal{D}$, then

$$\int_{X^m} 1[||\mu_m|| > y - (2\theta)^{1/2}] dH(x_{m+1}, \ldots, x_{2m})$$

$$\geq \int_{X^m} 1[||\mu'_m|| \leq (2\theta)^{1/2}] dH(x_{m+1}, \ldots, x_{2m})$$

$$\geq 1 - \frac{1}{2\theta} E(\mu'_m(D_0)^2)$$

$$\geq 1/2.$$
Hence

\[ H[|u_m^0| > y - (2\theta)^{1/2}] \]

\[ = \int \int \frac{1}{x^m x^m} \left[ |\|u_m^0\| > y - (2\theta)^{1/2} \right] dH(x_{m+1}, \ldots, x_{2m}) dH(x_1, \ldots, x_m) \]

\[ \geq \int [ |u_m| > y] 1/2 \ dH(x_1, \ldots, x_m) \]

\[ = \frac{1}{2} H[|u_m| > y]. \]

**Lemma 3.3.** For integers \( r \geq 0 \)

\[ \sum_{j=r+1}^{\infty} j^{1/2} 2^{-j} \leq 3 \cdot 2^{-r/2}. \]

**Proof.** Let \( f(x) := x^{1/2} 2^{-x} \). Then \( f'(x) < 0 \) for \( x \geq 1 \), so if \( r \geq 1 \) then

\[ \sum_{j=r+1}^{\infty} j^{1/2} 2^{-j} \leq \int_{r}^{\infty} f(x) dx \]

\[ = \int_{r}^{\infty} 2u^{2} 2^{-u} du \]

\[ = \frac{r^{1/2} 2^{-r}}{\log 2} + \int_{r^{1/2}}^{\infty} \frac{1}{\log 2} 2^{-u} du \]

\[ \leq \frac{r^{1/2} 2^{-r}}{\log 2} + \int_{r^{1/2}}^{\infty} 2u \log 2 \frac{1}{2r^{1/2} (\log 2)^2} 2^{-u} du \]
36.

\[ = 2^{-r} \left( \frac{r^{1/2}}{\log 2} + \frac{1}{2r^{1/2} (\log 2)^2} \right) \]

\[ \leq 3r^{1/2} e^{-r} \]

\[ \leq 3 \cdot 2^{-r/2} \]

The case \( r = 0 \) follows easily from the case \( r = 1 \). 

The next lemma also uses ideas from Pollard (1981).

We define \( K_1(v) := 2A(v) + 3A(v)^2, \ v \geq 1 \).

**Lemma 3.4.** Let \( H(1), \ldots, H(m), \ H, \) and \( \mu_0 \) be as in Lemma 3.2, and let \( \mathcal{C} \subset \mathcal{A} \) be a VC class. Let \( b > 0, \beta \geq \sup \ H(m)(D), \ u > 0, \)

\[ (3.6) \quad r := \left[ \frac{u^2}{128(b+\beta)V(\mathcal{C}) \log 2} \right] \]

where \([\cdot]\) denotes the integer part. Suppose

\[ (3.7) \quad (2^{13}V(\mathcal{S}))^{1/2} 2^{-r/2} \leq u/2. \]

Then

\[ H[\sup_{\mathcal{A}} |\mu_0(D)| > u] \leq H[\sup_{\mathcal{A}} |H_2m(D) - \ H(m)(D)| > b] \]

\[ + K_1(V(\mathcal{S})) \exp \left( -\frac{u^2}{32(b+\beta)} \right) \]
provided the events in (3.8) are $\mathbb{H}$-completion measurable.

Proof. Let $||\cdot||$ denote the sup norm over $\mathcal{G}$. Let $Z := \{-1,1\}$, let $\mathcal{P}_Z$ be the power set of $Z$, and let $F\{-1\} := F\{1\} := 1/2$. Let

$$\varepsilon_i(z) := z_i, \ z = (z_1, z_2, \ldots) \in Z^\infty,$$

and let $G := \mathbb{H} \times F_m$ on $(X^{2m} \times \mathbb{Z}^m, \mathcal{Q}^{2m} \times \mathcal{B}_{\mathbb{Z}}^m)$. Define

$$\mu^0_m := m^{1/2} \sum_{i=1}^{m} \varepsilon_i(\delta_{X_i} - \delta_{X_{n+i}}), \text{ on } X^{2m} \times \mathbb{Z}^m,$$

$$W := \{x(2m) : ||H_{2m} - \overline{H}(m)|| > b \} \subset X^{2m}.$$

Then

(3.9) $\mathbb{H}[||\mu^0_m|| > u] = G[||\mu^0_m|| > u].$

(This is clear conditional on any fixed values of the $\varepsilon_i$.)

We have

$$G[||\mu^0_m|| > u] = \int \int 1_{\{||\mu^0_m|| > u\}} dF^m d\mathbb{H}$$

(3.10)

$$< \mathbb{H}(W) + \sup_{x(2m) \notin W} F^m[||\mu^0_m|| > u].$$
38.

Fix $x^{(2m)} \not\in W$. Then from the definition of $W$,

\[(3.11) \quad \sup_{\mathcal{F}} H_{2m}(D) \leq b + \beta.\]

Let $m_j := N(2^{-2j}, \mathcal{F}, H_{2m})$. Then for each $j \geq 0$ there exist sets $D_{j1}, \ldots, D_{jm_j} \in \mathcal{F}$ such that $\min_i H_{2m}(D \Delta D_{j_i}) < 2^{-2j}$ for every $D \in \mathcal{F}$.

Define $\eta_j, j \geq 0$, by

\[\eta_j^2 := 480V(\mathcal{F})(\log 2)j2^{-2j}.\]

By (3.7) and Lemma 3.3 we have

\[(3.12) \quad \sum_{j=r+1}^{\infty} \eta_j \leq u/2.\]

For $D \in \mathcal{F}$ and $j \geq 0$ let $D_j(D)$ be one of the $D_{ji}, i \leq m_j$, such that $H_{2m}(D \Delta D_j(D)) < 2^{-2j}$. If $j$ is large enough so $2^{-2j} < 1/2m$, then $H_{2m}(D \Delta D_j(D)) = 0$ for all $D \in \mathcal{F}$, so $\tilde{\mu}_m^0(D) = \tilde{\mu}_m^0(D_j(D))$. It follows that

\[\mu_m(D) = \mu_m(D_r(D)) + \sum_{j=r+1}^{\infty} [\mu_m^0(D_j(D)) - \mu_m^0(D_{j-1}(D))].\]

Also
39.

\[ H_{2m}(D_j(D) \Delta D_{j-1}(D)) \leq H_{2m}(D \Delta D_j(D)) + H_{2m}(D \Delta D_{j-1}(D)) \leq 2^{-2j} + 2^{-2(j-1)} = 5 \cdot 2^{-2j}. \]

Letting \( \gamma_j := 5 \cdot 2^{-2j} \), it follows using (3.12) that

\[ F_m^m \left[ \sup_{\mathcal{G}} |\tilde{\mu}_m^0(D(D))| > u \right] \]

\[ \leq F_m^m \left[ \sup_{\mathcal{G}} |\tilde{\mu}_m^0(D_\mathcal{F}(D))| > u/2 \right] \]

\[ + \sum_{j=r+1}^{\infty} F_m^m \left[ \sup_{\mathcal{G}} |\tilde{\mu}_m^0(D_j(D)) - \tilde{\mu}_m^0(D_{j-1}(D))| > \eta_j \right] \]

(3.13)

\[ \leq m_{r+1} \max \{ F_m^m[|\tilde{\mu}_m^0(D_{\mathcal{F}i})| > u/2] : i \leq m_r \}
\]

\[ + \sum_{j=r+1}^{\infty} m_j m_{j-1} \max \{ F_m^m[|\tilde{\mu}_m^0(D_{ji}) - \tilde{\mu}_m^0(D_{(j-1)k})| > \eta_j] : i \leq m_j, k \leq m_{j-1}, H_{2m}(D_{ji} \Delta D_{(j-1)k}) \leq \gamma_j \}. \]

Now \( \tilde{\mu}_m^0(D_{ji}) - \tilde{\mu}_m^0(D_{(j-1)k}) = m^{-1/2} \sum_{\lambda=1}^{m} \varepsilon_{n} h_{\lambda}, \)

where \( h_{\lambda} := (1_{D_{ji}} - 1_{D_{(j-1)k}})(X_{\lambda}) - (1_{D_{ji}} - 1_{D_{(j-1)k}})(X_{m+\lambda}), \)

so
40.

\[ \sum_{\ell=1}^{m} h_\ell^2 \leq 2 \sum_{\ell=1}^{2m} (1_{D_{ji}} - 1_{D(j-1)k})^2(x_\ell) \]

\[ \leq 4m2^m (D_{ji}, D(j-1)k) \]

\[ \leq 20m2^{-2j} \]

if \( H_{2m}(D_{ji}, D(j-1)k) \leq Y_j \). Hence by Theorem 2 of Hoeffding (1963), the terms on the right side of (3.13) satisfy

\[ P^m[|\tilde{\mu}_m(D_{ji}) - \tilde{\mu}_m(D(j-1)k)| > \eta_j] \]

\[ \leq 2 \exp \left( -2m\eta_j^2/4 \sum_{\ell=1}^{m} h_\ell^2 \right) \]

(3.14)

\[ \leq 2 \exp \left( -\eta_j^2/(40 \cdot 2^{-2j}) \right) \]

\[ = 2 \exp \left( -12jV(\xi)\log 2 \right). \]

Similarly, \( \tilde{\mu}_m(D_{ri}) = m^{-1/2} \sum_{\ell=1}^{m} \epsilon_\ell g_\ell \), where \( g_\ell := 1_{D_{ri}}(X_\ell) - 1_{D_{ri}}(X_{m+\ell}) \), so by (3.11),

\[ \sum_{\ell=1}^{m} g_\ell^2 \leq 2mH_{2m}(D_{ri}) \leq 2m(b+b). \]

Hence applying Theorem 2 of Hoeffding (1963) again,
\[ F^m[|\mu_m^0(D, r)| > u/2] \leq 2 \exp \left(-2\mu u^2/16 \sum_{\lambda=1}^{m} g_{\lambda}^2 \right) \]  
(3.15)

Now by Lemma 2.5 we have \( m_j \leq A(V(\mathcal{S})2^4jV(\mathcal{S})) \), so by (3.13), (3.14), and (3.15) and the definition of \( r \),

\[ F^m[\sup_{D} |\mu_m^0(D)| > u] \]

\[ \leq 2A(V(\mathcal{S})) e^{4rV(\mathcal{S})\log 2} e^{-u^2/16(b+\beta)} \]
\[ + \sum_{j=r+1}^{\infty} 2A(V(\mathcal{S})) e^{8jV(\mathcal{S})\log 2} e^{-12jV(\mathcal{S})\log 2} \]
\[ \leq 2A(V(\mathcal{S})) e^{-u^2/32(b+\beta)} \]
\[ + 2A(V(\mathcal{S})) e^{-4(r+1)V(\mathcal{S})\log 2}/(1 - e^{-4V(\mathcal{S})\log 2}) \]
\[ \leq [2A(V(\mathcal{S})) + 3A(V(\mathcal{S}))^2] e^{-u^2/32(b+\beta)}. \]

The lemma now follows from (3.9), (3.10), and (3.16). \( \square \)

Let \( N \) be a positive integer, to be specified later, and define
Y := \{1, \ldots, N\}

B_Y := \text{the power set of } Y

Q(\{i\}) := 1/N \text{ for all } i \in Y

Q := Q^\infty, \text{ a measure on } (Y^\infty, B_Y^\infty)

\sigma(i)(y) := (i-1)N + y_1, \text{ for } y = (y_1, y_2, \ldots) \in Y^\infty

\tau(i)(y) := (i-1)N + y_{N+i}

For each \((x,y) \in X^\infty \times Y^\infty\) we get probability measures defined as follows:

\begin{align*}
\nu_n^{(1)} &:= \frac{1}{n} \sum_{i=1}^{n} \delta_{X_{\sigma(i)}} \\
\nu_n^{(2)} &:= \frac{1}{n} \sum_{i=1}^{n} \delta_{X_{\tau(i)}} \\
\nu_n &:= \frac{1}{n} \sum_{i=1}^{n} \frac{jN}{i=(j-1)N+1} \delta_{X_i}.
\end{align*}

Note that \(\nu_{nN} = \frac{1}{n} \sum_{j=1}^{n} \nu_n^{(j)}\). Define

\begin{align*}
\nu_n^{(j)} &:= n^{1/2}(\nu_n^{(j)} - \nu_{nN}) , \quad j = 1, 2 \\
\nu_n^{0} &:= n^{1/2}(\nu_n^{(1)} - \nu_n^{(2)}) \\
\mathbf{P}(N) &:= \prod_{i=1}^{\infty} P_{nN}^{(i)}
\end{align*}
Pr(N) := P(N) × Q, a measure on \((X^n × Y^n, \mathcal{A}_X × \mathcal{A}_Y)\)

Pr := P × Q.

Under the law Pr(N) on \(X^n × Y^n\), we can think of \(P_n^{(1)}\) and \(P_n^{(2)}\) as each being obtained by sampling \(n\) groups of \(N\) points each from \(X\), with the \(i\)th group sampled according to \(P(i)\), then choosing one point randomly from each of the \(n\) groups.

For fixed \(x^{(nN)}\), \(P_n^{(1)}\) is a version of the empirical measure which has \(n\) sample points chosen according to \(P_{1,n}, ..., P_{2,n}\) respectively. The latter measures are of course non-random for fixed \(x^{(nN)}\). The corresponding normalized empirical process is \(v_n^{(1)}\). \(P_n^{(2)}\) is an independent copy (for fixed \(x^{(nN)}\)) of \(P_n^{(1)}\).

The next lemma will in a sense make it sufficient to prove Theorem 3.1 with \(P(i)\) replaced by \(P_{i,N}\). After such replacement we can take advantage of the fact that \(P_{i,N}\) is bounded below in the sense that \(P_{i,N}(C) ≥ 1/N\) if \(P_{i,N}(C) > 0\).

Lemma 3.5. Let \(n ≥ 1\) and let \(C ⊂ \mathcal{A}\) be a VC class satisfying the measurability assumptions (3.1) and (3.2). Let \(0 < \varepsilon < 1, α > 0, v ≥ V(C),\) and \(M < n^{1/2}\). Suppose
(3.17) \[ N \geq \max \left( \frac{16n}{\varepsilon^2 \alpha}, \frac{17n \log n}{\varepsilon^2} \right), \]

(3.18) \[ n \geq 2^{18v} \]

and let

(3.19) \[ W := \cup_{j=1}^{n} \{ x(\mathbb{N}) : \sup_{C} |P_{j,N}(C) - P(j)(C)| > \varepsilon M/16n^{1/2} \}. \]

Then

\[ \mathbb{P}[\sup_{C} |\nu_n(C)| > M] \leq 2K_1(v) \exp \left( -M^2/2\alpha \right) \]

(3.20)

\[ + \sup_{x(\mathbb{N}) \not\in W} Q[\sup_{C} |\nu_n^{(1)}(C)| > (1 - \varepsilon/16)M] \]

**Proof.** Let \( || \cdot || \) denote the sup norm over \( C \). If \( x(\mathbb{N}) \not\in W \) then

\[ ||\nu_n^{(1)} - n^{1/2}(P_n^{(1)} - \overline{P}(n))|| = ||n^{1/2}(P_{nN} - \overline{P}(n))|| \]

\[ = ||n^{-1/2} \sum_{j=1}^{n} (P_{j,N} - P(j))|| \]

\[ \leq \varepsilon M/16. \]

It follows that
\[ \mathbb{P}[|\nu_n| > M] = \text{Pr}^N[|n^{1/2}(p_{n1} - \overline{p}_n)| > M] \]

(3.21) \[ = \int \int 1_{x,y} \mathbb{P}^N[|n^{1/2}(p_{n1} - \overline{p}_n)| > M] \]

\[ \leq \mathbb{P}^N(W) + \sup_{x(nN) \notin W} \mathcal{Q}[|\nu_n| > (1 - \frac{\epsilon}{16})M]. \]

Now

(3.22) \[ \mathbb{P}^N(W) \leq \sum_{j=1}^{n} P^N(j)[|N^{1/2}(p_{1N} - p_{jN})| > \epsilon M^{1/2}/16n^{1/2}]; \]

we wish to bound each of these \( n \) terms by applying

Lemmas 3.2 and 3.4 to \( \mathcal{C} := \mathcal{C}, m := N, \theta := 1/2, \)
\( y := \epsilon M^{1/2}/16n^{1/2}, u := \epsilon M^{1/2}/32n^{1/2}, b := 1, \beta := 1, \) and
\( H(i) := p_{jN} \) for all \( i \leq N. \) From Lemma 3.2 and (3.17) we get

\[ P^N(j)[|N^{1/2}(p_{1N} - p_{jN})| > \epsilon M^{1/2}/16n^{1/2}] \]

(3.23) \[ \leq 2P^{2N}[|N^{1/2}(p_{1N} - p_{2N})| > \epsilon M^{1/2}/16n^{1/2} - 1] \]

\[ \leq 2P^{2N}(j)[|N^{1/2}(p_{1N} - p_{2N})| > \epsilon M^{1/2}/32n^{1/2}]. \]

Let \( r \) be as in (3.6); then by (3.17) and (3.18),
\[ r > \left[ \frac{u^2}{256v \log 2} \right] \]

\[ = \left[ \frac{\varepsilon^2 M^2 N}{2^{18} vn \log 2} \right] \]

\[ \geq \frac{\log n}{2v \log 2} - 1 \]

\[ \geq 8. \]

Combining (3.24) with (3.17) and (3.18) we get

\[ (2^{13 V(\mathcal{D})})^{1/2} - r/2 \leq (2^5 v)^{1/2} \leq u/2, \]

so (3.7) holds and we may apply Lemma 3.4. Noting that the event on the right side of (3.8) is empty, we get

\[ P_j^{2N}[|N^{1/2}(P_{1,N} - P_{2,N})| > \varepsilon MN^{1/2}/32n^{1/2}] \]

\[ \leq K_1(v) \exp (-\varepsilon^2 M^2 N/2^{16} n). \]

By (3.17),

\[ \log n - \frac{\varepsilon^2 M^2 N}{2^{16} n} \leq - \frac{\varepsilon^2 M^2 N}{2^{17} n} \]

\[ \leq - M^2/2a. \]

Combining (3.22), (3.23), (3.25), and (3.26) gives
$\mathbb{P}^{(N)}(W) \leq 2nK_1(v) \exp \left( -\varepsilon^2 M^2 N / 2^{16} n \right)$

$\leq 2K_1(v) \exp \left( -M^2 / 2\alpha \right)$.

This and (3.21) prove the lemma. \[\square\]

For $C \subseteq \mathcal{A}$, $\gamma > 0$, and $H$ a law on $(X, \mathcal{A})$, define $\mathcal{C}_1(\gamma, C, H) := \{ A \mid B : A, B \in C, H(A \setminus B) < \gamma \}$.

**Lemma 3.6.** Let $C \subseteq \mathcal{A}$ be a VC class, $n \geq 1$, $0 < \varepsilon < 1$, $\alpha > \alpha(C, n)$, $V \geq V(C)$, and $M < n^{1/2}$. If $\alpha < 1/4$ we suppose that

$$M < 3n^{1/2} \alpha \varepsilon / 8.$$  

Let $H(1), \ldots, H(n), H, \overline{H}(n)$, and $\mu_n$ be as in Lemma 3.3, and suppose that

$$\sup_{C} |H(i)(C) - P(i)(C)| \leq \varepsilon M / 16 n^{1/2} \quad \text{for all } i \leq n.$$  

Let $\gamma := \exp \left( -\varepsilon M^2 / 16 \alpha v \right)$. Then

$$H^* \left[ \sup_{C} |\mu_n(C)| > (1 - \frac{\varepsilon}{16}) M \right]$$

$\leq 2A(v) \exp \left( -(1-\varepsilon) M^2 / 2\alpha \right) + H^* \left[ \sup_{\mathcal{C}_1(\gamma, C, H(n))} |\mu_n(C)| > 3\varepsilon M / 32 \right]$.  

\[\square\]
Proof. Let \( m := N(\gamma, \mathcal{C}, H(n)) \); then there exist \( C_1, \ldots, C_m \in \mathcal{C} \) such that \( \min_{i \leq m} H(n)(C \Delta C_i) < \gamma \) for all \( C \in \mathcal{C} \). If \( |\mu_n(C)| > (1 - \frac{\varepsilon}{16})M \) for some \( C \in \mathcal{C} \), then there exists \( i \leq m \) such that \( H(n)(C \Delta C_i) < \gamma \), so \( C \setminus C_i \) and \( C_i \setminus C \) are in \( \mathcal{C}_1(\gamma, \mathcal{C}, H(n)) \), and hence either \( |\mu_n(C_i)| > (1 - \frac{\varepsilon}{4})M \) or \( |\mu_n(C \setminus C_i)| > 3\varepsilon M/32 \) or \( |\mu_n(C_i \setminus C)| > 3\varepsilon M/32 \). It follows that

\[
H^* \left[ \sup_{C} |\mu_n(C)| > (1 - \frac{\varepsilon}{16})M \right] \\
\leq m \max_{i \leq m} H[|\mu_n(C_i)| > (1 - \frac{\varepsilon}{4})M] \\
+ H^* \left[ \sup_{\mathcal{C}_1(\gamma, \mathcal{C}, H(n))} |\mu_n(C)| > 3\varepsilon M/32 \right].
\]

Fix \( C \in \mathcal{C} \), and let \( \sigma^2 := \text{var}_H(\mu_n(C)) \). Define \( f(t) := t(1-t), t \in [0,1] \); then \( f(t+h) \leq f(t) + |h| \), so using (3.28),

\[
\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} f(H(i)(C)) \\
\leq \frac{1}{n} \sum_{i=1}^{n} f(P(i)(C)) + \varepsilon M/16n^{1/2} \\
\leq \alpha(C, n) + \varepsilon M/16n^{1/2}.
\]
49.

If \( \alpha < 1/4 \) then by (3.27),

\[
\frac{\epsilon M}{16n^{1/2}} \leq \frac{3\alpha \epsilon^2}{128} \leq \frac{\alpha \epsilon}{4}.
\]

If \( \alpha \geq 1/4 \) then since \( M < n^{1/2} \),

\[
\frac{\epsilon M}{16n^{1/2}} < \frac{\epsilon}{16} \leq \frac{\alpha \epsilon}{4}.
\]

Hence by (3.31),

\[
(3.32) \quad \alpha^2 \leq (1 + \frac{\epsilon}{4}) \alpha.
\]

Let \( \tilde{M} := (1 - \frac{\epsilon}{4}) M \), \( \lambda_1 := \tilde{M}/n^{1/2} \alpha^2 \), and \( \lambda_2 := \tilde{M}/n^{1/2} \alpha (1 + \frac{\epsilon}{4}) \).

By Bernstein's inequality (see Hoeffding (1963)) and (3.32),

\[
(3.33) \quad \text{H}[|\mu_n(C)| > \tilde{M}] \leq 2 \exp \left( -n^{1/2} \tilde{M} \lambda_2/2 \right) \leq 2 \exp \left( -n^{1/2} \tilde{M} \lambda_1/2 \right) \leq 2 \exp \left( -n^{1/2} \tilde{M} \lambda_1/2 \right) (1 + \frac{1}{3} \lambda_1).
\]

If \( \alpha < 1/4 \), then \( \lambda_2 \leq M/n^{1/2} \alpha \leq 3\epsilon/8 \) so (3.33) gives

\[
\text{H}[|\mu_n(C)| > \tilde{M}] \leq 2 \exp \left( -n^{1/2} \tilde{M} \lambda_2/2 \right) (1 + \frac{\epsilon}{8})
\]

\[
(3.34) \quad \leq 2 \exp \left( -(1 - \frac{\epsilon}{4})^2 M^2/2\alpha (1 + \frac{\epsilon}{8}) (1 + \frac{\epsilon}{4}) \right)
\]

\[
\leq 2 \exp \left( -(1 - \frac{7}{8} \epsilon) M^2/2\alpha \right).
\]
If $\alpha > 1/4$ then Theorem 1 of Hoeffding (1963) gives

$$H[|\mu_n(C)| > M] \leq 2 \exp (-2M^2)$$

(3.35)

$$\leq 2 \exp (-\left(1 - \frac{7}{8}\epsilon\right)M^2/2\alpha)$$

By Lemma 2.5,

(3.36) $m \leq A(v)\gamma^{-2v} = A(v)\exp (\varepsilon M^2/8\alpha)$

Combining (3.30), (3.34), (3.35), and (3.36) gives (3.29).

For $v > 1$ let $v_1(v)$ be the least integer $k > 1$ such that $(k^v + 1)^2 < 2^k$. Then $v_1(v) \geq v_1(1) = 6$ for all $v$. Define also $K_2(v) := 2K_1(v_1(v)) + 8A(v_1(v))$.

Lemma 3.7. Let $C$ be a VC class, $v > V(C)$, $0 < \varepsilon < 1$, $M > 0$, and $\alpha > 0$. Let $\gamma := \exp (-\varepsilon M^2/16\alpha v)$ and suppose

(3.37) $N-1 \leq \max \left(\frac{2^{16}n}{\varepsilon^2\alpha}, \frac{2^{17}n \log n}{\varepsilon^2 M^2}\right)$

Then there exist constants

$$M_0 = M_0(\varepsilon, \alpha, v), n_1 = n_1(\varepsilon, \alpha, v)$$

such that if $n \geq n_1$ and $M \geq M_0$, then
Proof. Let \( v_1 := v_1(\gamma) \) and \( C_1 := C_1(\gamma, \mathcal{C}, P_{nN}) \), and let \( || \cdot || \) denote the sup norm over \( C_1 \). We wish to apply Lemmas 3.2 and 3.4 to \( \mathcal{C} := C_1 \), \( \gamma := 3\epsilon M/32 \), \( \theta := \gamma \), \( m := n \), \( H(i) := P_{i,N} \) for all \( i \), \( b := \beta := \alpha \epsilon^2/2^{13} \), and \( u := \epsilon M/16 \). Note that measurability is not a problem because all functions of \( v_n^{(1)} \) or \( \bar{v}_n^0 \) are \( \mathcal{B}_Y \)-measurable.

Let \( r \) be as in (3.6); then using Lemma 2.4,

\[
r = \left[ \frac{u^2}{128(b+\beta) V(\mathcal{C}) \log 2} \right] \\
\geq \left[ \frac{M^2}{32\alpha v_1 \log 2} \right].
\]

Let \( M_0(\epsilon, \alpha, v) \) be the least real number such that the following all hold for all \( M \geq M_0 \):

(3.38) \( 2^{1/2} \exp (-\epsilon M^2/32\alpha v) \leq \epsilon M/32 \)

(3.39) \( (2^{13} v_1)^{1/2} 2^{-r/2} \leq \epsilon M/32 \)

(3.40) \( \exp (-\epsilon M^2/16\alpha v) \leq \alpha \epsilon^2/2^{13} \)

(3.41) \( 4\alpha v_1 \leq M^2 \).

Let \( n_1(\epsilon, \alpha, v) := \max (2^{49} \alpha, 2^{46} v_1^{-13/3} \epsilon^{-32/3}, 2^{33} v_1^{-2} \epsilon^{-5}) \).
Fix $M \geq M_0$ and $n \geq n_1$. Now

\[
\sup_{C_1} \left( \sum_{i=1}^{n} P_i, N(C) (1 - P_i, N(C)) \right) \leq \sup_{C_1} P_{nN}(C) \leq \gamma
\]

so the hypotheses of Lemma 3.2 are satisfied; using (3.38), (3.5) becomes

\[
(3.42) \quad Q\left[ |y_n^{(1)}| > 3\varepsilon M/32 \right] \leq 2Q\left[ |\tilde{y}_n^0| > \varepsilon M/16 \right].
\]

By (3.40) we have \( \sup_{C_1} P_{nN}(C) \leq \gamma \leq \beta \), while (3.7) follows from (3.39). Thus the hypotheses of Lemma 3.4 are satisfied, and (3.8) becomes

\[
Q\left[ |\tilde{y}_n^0| > \varepsilon M/16 \right] \leq Q\left[ \left| \frac{1}{2} P_n^{(1)} + P_n^{(2)} - P_{nN} \right| > b \right]
+ K_1(v_{\tilde{C}}) \exp \left( -\frac{u^2}{32(b+\beta)} \right)
\]

(3.43)

\[
\leq 2Q\left[ \left| P_n^{(1)} - P_{nN} \right| > b \right]
+ K_1(v_{\tilde{C}}) \exp \left( -\frac{M^2}{2a} \right).
\]

Let \( k := N(1/nN, C_1, P_{nN}) \); then there exist \( C_1, \ldots, C_k \in C_1 \) such that \( C \in C_1 \) implies \( P_{nN}(C \Delta C_i) < 1/nN \) for some \( i \leq k \). But then \( P_{nN}(C \Delta C_i) = 0 \), so \( P_n^{(1)}(C \Delta C_i) = 0 \) also. Thus
53.

\[ Q[|P_n^{(1)} - P_{nN}| > b] \]

(3.44) \[ = Q[\max_{i \leq k} |P_n^{(1)}(C_i) - P_{nN}(C_i)| > b] \]

\[ \leq k \max_{i \leq k} Q[|P_n^{(1)}(C_i) - P_{nN}(C_i)| > b]. \]

Define

\[ g(\mu) := \begin{cases} 
\log \left( \frac{(1-\mu)/\mu}{(1-2\mu)} \right), & 0 < \mu < \frac{1}{2} \\
1/2\mu(1-\mu), & \frac{1}{2} \leq \mu < 1.
\end{cases} \]

Then \( g(\mu) \geq 2 \) (see Hoeffding (1963)) so

\[ g(\mu) \geq 2 \geq \log \frac{1}{\mu} > \frac{1}{2} \log \frac{1}{\mu}, \quad \mu > e^{-2} \]

(3.45)

\[ g(\mu) \geq \log \left( \frac{1}{\mu} - 1 \right) > (\log \frac{1}{\mu}) - 1 \geq \frac{1}{2} \log \frac{1}{\mu}, \quad \mu \leq e^{-2}, \]

and then also

\[ g(1-\mu) \geq \frac{1}{2} \log \frac{1}{1-\mu} > \frac{1}{2} \log \frac{1}{\mu}, \quad \mu > \frac{1}{2} \]

(3.46)

\[ g(1-\mu) = \frac{1}{2\mu(1-\mu)} > \frac{1}{2\mu} \geq \frac{1}{2} \log \frac{1}{\mu}, \quad \mu < \frac{1}{2}. \]

By Theorem 1 of Hoeffding (1963), (3.45), and (3.46),
\[ Q\left[ |P_n^{(1)}(C_i) - P_{nN}(C_i)| > b \right] \]

\[ \leq \exp(-nb^2 g(P_{nN}(C_i))) + \exp(-nb^2 g(1 - P_{nN}(C_i))) \]

(3.47)

\[ \leq 2 \exp(-\frac{1}{2}nb^2 \log \frac{1}{Y}) \]

\[ = 2 \exp(-na\varepsilon^5 M^2/2^{33}v). \]

**Case 1:** \( M^2/\alpha \leq 2 \log n \), so the second term on the right side of (3.37) is the larger. Now

\((\log n)/n \leq (\log 2^{49})/2^{49} \) since \( n \geq 2^{49} \), so using Lemma 2.5, (3.37), and (3.41) we get

\[ \log (k/A(v_1)) \leq 2v_1 \log (nN) \]

\[ \leq 2v_1 \log (2^{18}n^2(\log n)/\varepsilon^2 M^2) \]

\[ \leq 2v_1 \log (n^{3/2}30^2\varepsilon^2 \alpha). \]

**Case 2:** \( M^2/\alpha > 2 \log n \), so the first term on the right side of (3.37) is larger. Again since \( n \geq 2^{47} \), we have, using (3.37):

\[ \log (k/A(v_1)) \leq 2v_1 \log (nN) \]

\[ \leq 2v_1 \log (2^{17}n^2/\varepsilon^2 \alpha) \]

\[ \leq 2v_1 \log (n^{3/2}30^2\varepsilon^2 \alpha). \]
Now (in both cases) $n^{3/2}e^2 > 2^{68}$ since $n > n_1$, and $t^{-1/6}\log t$ is a decreasing function for $t > e^6$, so using the definition of $n_1$ and (3.41) we get

$$\log (k/A(v_1)) \leq 2v_1 \log \left(\frac{n^{3/2}e^2}{e^2}\right)$$

$$\leq 2v_1(2^{-68/68}\log 2)n^{1/2}/2^5\left(e^2\right)^{1/6}$$

$$\leq n^{1/2}v_12^{-9}\left(e^2\right)^{-1/6}(n^{1/2}a^{13/6}\epsilon^{16/3}e^{-23}\epsilon^{-1})$$

$$\leq nae^{5}v_1/2^{32}v$$

$$\leq nae^{5}M^2/2^{34}v.$$ (3.48)

Combining (3.42), (3.43), (3.44), (3.47), and (3.48) now gives

$$Q[||v_n^{(1)}|| > 3\epsilon M/32] \leq 8A(v_1)\exp(-nae^{5}M^2/2^{34}v)$$

$$+ 2K_1(v_1)\exp(-M^2/2a)$$

$$\leq K_2(v)\exp(-M^2/2a),$$

the last inequality following from the definition of $n_1$. [−]

We are now ready to assemble Lemmas 3.5-3.7 into a proof of our main theorem.
Proof of Theorem 3.1. We may assume \( M < n^{1/2} \), since \(|\nu_n(C)| \leq n^{1/2}\). Let \( N \) be the least integer satisfying (3.17); then \( N \) satisfies (3.37). Let \( M_0 \) and \( n_1 \) be as in Lemma 3.7, and let \( n_0 := \max(n_1, 2^{18v}) \). It is easy to see that (3.4) holds. Let \( W \) be as in (3.19), and \( \gamma \) as in Lemmas 3.6 and 3.7. We first apply Lemma 3.5, then apply Lemma 3.6 to the second term on the right side of (3.20), with \( H(i) := P_{i,N} \) for all \( i \); (3.28) follows from the definition of \( W \), and for fixed \( x(nN) \) there is an obvious correspondence between the \( H \) of Lemma 3.6 and \( Q \). Finally we apply Lemma 3.7 to the second term on the right side of (3.29). All of this gives

\[
\mathbb{P}\left[ \sup_C |\nu_n(C)| > M \right] \leq 2K_1(v) \exp(-M^2/2\alpha)
\]

\[
+ \sup_{x(nN) \notin W} \mathbb{Q}[\sup_C |\nu_n^{(1)}(C)| > (1 - \frac{\epsilon}{16})M]
\]

\[
\leq (2K_1(v) + 2A(v)) \exp(-(1-\epsilon)M^2/2\alpha)
\]

\[
+ \sup_{x(nN) \notin W} \mathbb{Q}\left[ \sup_{C_1(\gamma, C, P_{nN})} |\nu_n^{(1)}(C)| > 3\epsilon M/32 \right]
\]

\[
\leq (2K_1(v) + 2A(v) + K_2(v)) \exp(-(1-\epsilon)M^2/2\alpha)
\]

From the proof above we see that if \( \alpha > \alpha(C, n) \), in place of (3.3) it is sufficient to verify that
(3.38) - (3.41) all hold and that

(3.49) \( n \geq \max(2^{18 \alpha}, 2^{49 \alpha}, 2^{46 \alpha} \varepsilon^{-13/3}, 2^{33 \alpha} \varepsilon^{-5}). \)

To prove limit theorems for \( \nu_n \) as \( n \to \infty \) we may wish to apply Theorem 3.1 with \( \alpha := \alpha_n := \alpha(C, n) \) and \( n \to \infty \).

To satisfy (3.49) in this situation requires that \( n^{-3/13} = 0(\alpha_n) \). The next two theorems, however, are variants of Theorem 3.1 which are valid under conditions which may be weaker than (3.49); more about this later.

The proofs are also simpler.

Given \( \mathcal{F} \subseteq \mathcal{C} \) and \( n \geq 1 \) define

\[ \beta(\mathcal{F}, n) := \sup_{D} \bar{P}(n)(D). \]

**Theorem 3.8** Let \( \mathcal{F} \subseteq \mathcal{C} \) be a VC class which is \( n \)-deviation-measurable for \( \{P(i), i \geq 1\} \), \( \beta \geq \beta(\mathcal{F}, n) \), \( \nu \geq \nu(\mathcal{F}) \), and \( u > 0 \). Let

\[ r(1) := \left| \frac{u^2}{512 v \beta \log 2} \right| \]

(3.50)

\[ r(2) := \left| \frac{u^{1/2}}{2048 \nu \log 2} \right| \]

and suppose

(3.51) \( (2^{13 \nu})^{1/2} \leq r(i)/2 \leq u/2, \ i = 1, 2, \)
(3.52) \[ u^2/\beta \geq 36 \log 2, \]

and

(3.53) \[ n^{1/2}u \geq 512 \log 2. \]

Then

\[
\mathbb{P} \left[ \sup \{v_n(D) \mid > u \} \leq 4K_1(v)\exp(-u^2/256\beta) \right.
\]

\[
\left. + 4K_1(v)\exp(-n^{1/2}u/1024) \right. \]

\[
\mathbb{P}(n,2) := \left( \prod_{i=1}^n P(i) \right)^2, \text{ and} \]

\[
y_j := 4^j u \\
u_j := y_j/2^{1/2} = 4^ju/2^{1/2} \\
b_j := y_{j+1}/n^{1/2} = 4^{j+1}u/n^{1/2} \\
r_j := [u_j^2/128v(b_j+\beta)\log 2] \]

for \( j \geq 0 \). We first apply Lemma 3.2 with \( H(i) := P(i) \)

and \( \theta := \beta \) and observe that

\[
y_j - (2\beta)^{1/2} \geq (1 - (18 \log 2)^{-1/2})y_j \geq y_j/2^{1/2} = u_j \]
by (3.52), since \( y_j \geq u \). We then apply Lemma 3.4, noting that (3.7) with \( r = r_j \) follows from (3.51), since \( r_j \geq r_0 \geq \min(r^{(1)}, r^{(2)}) \). We get

\[
\mathbb{P}[||v_n|| > y_j] \leq 2\mathbb{P}(n, 2)[||v^0_n|| > u_j]
\]

\[
\leq 2\mathbb{P}(n, 2)[||P_{2n} - \bar{P}(n)|| > b_j]
+ 2K_1(v)\exp(-u^2_j/32(b_j + \beta))
\]

(3.55)

\[
\leq 4\mathbb{P}[||P_n - \bar{P}(n)|| > b_j]
+ 2K_1(v)\exp(-u^2_j/32(b_j + \beta))
\]

\[
= 4\mathbb{P}[||v_n|| > y_{j+1}]
+ 2K_1(v)\exp(-u^2_j/32(b_j + \beta)).
\]

By obvious induction we get from (3.55) that for \( J \geq 1 \),

\[
\mathbb{P}[||v_n|| > y_0] \leq 4^J\mathbb{P}[||v_n|| > y_J]
+ \sum_{j=0}^{J-1} 2K_1(v)4^j\exp(-u^2_j/32(b_j + \beta)).
\]

But if \( J \) is large enough so \( y_J \geq n^{1/2} \), it follows that
$60.$

$$
P[\| \mathbf{v}_n \| > u] \leq \sum_{j=0}^{\infty} 2K_1(v)4^j \exp(-u_j^2/32(b_j+\beta))
$$

(3.56) \quad \leq \sum_{j=0}^{\infty} 2K_1(v)4^j \exp(-u_j^2/64b_j)

$$
+ \sum_{j=0}^{\infty} 2K_1(v)4^j \exp(-u_j^2/64\beta).
$$

Now for $\tau \geq 2$ the function $f(j) = j\tau^{-j}$ on the nonnegative integers reaches its maximum value $\tau^{-1}$ at $j = 1$. It follows that

(3.57) \quad j \leq \tau^{j-1}$ for $\tau \geq 2$ and integers $j \geq 0$.

Also

(3.58) \quad \tau^j \geq 1 + (\tau-1)j$ for $\tau > 0$ and integers $j \geq 0$

since $f(x) = \tau^x$ is convex and equality holds at $j = 0$ and 1. Applying (3.57) with $\tau = 16$ to (3.52) we get

$$
u^2/\beta \geq 288(\log 4)j4^{-2j}, j \geq 0,$$

so

(3.59) \quad j \log 4 \leq 4^2j u^2/288\beta \leq u_j^2/128\beta.$$

Using (3.58) with $\tau = 16$ we get
(3.60) \[ u_j^2/128β = 4^j u^2/256β \geq (1+8j) u^2/256β. \]

Combining (3.59), (3.60), and (3.52),

\[
\sum_{j=0}^{∞} 2K_1(v)4^j \exp(-u_j^2/64β) \leq \sum_{j=0}^{∞} 2K_1(v) \exp(-u_j^2/128β) \\
\leq \sum_{j=0}^{∞} 2K_1(v) \exp(-(1+8j) u^2/256β) \\
= 2K_1(v) \exp(-u^2/256β)/(1-\exp(-u^2/32β)) \\
\leq 4K_1(v) \exp(-u^2/256β).
\]

Next we apply (3.57) with \( \tau = 4 \) to (3.53) and get

\[ u/n^2 \geq 256j(\log 4)/4^{j-1}, \quad j \geq 0 \]

so

(3.62) \[ j \log 4 \leq 4^{j-1} n^{1/2}/256 = u_j^2/128b_j. \]

Using (3.58) with \( \tau = 4 \) we get

(3.63) \[ u_j^2/128b_j = 4^j n^{1/2}/1024 \geq (1+2j)n^{1/2}/1024 \]

Combining (3.62), (3.63), and (3.53),
\[
\sum_{j=0}^{\infty} 2K_1(v)4^j \exp(-u_j^2/64b) \\
\leq \sum_{j=0}^{\infty} 2K_1(v) \exp(-u_j^2/128b_j) \\
(3.64) \leq \sum_{j=0}^{\infty} 2K_1(v) \exp(-(1+2j)u_n^{1/2}/1024) \\
= 2K_1(v) \exp(-u_n^{1/2}/1024)/(1 - \exp(-2u_n^{1/2}/1024)) \\
\leq 4K_1(v) \exp(-u_n^{1/2}/1024).
\]

Combining (3.56), (3.61), and (3.64) finishes the proof.

For \( v \geq 1 \) define \( K_3(v) := 2A(v) + 2K_1(v) + 8K_1(v_1(v)) \).

**Theorem 3.9.** Let \( \mathcal{C} \subset \mathcal{A} \) be a VC class satisfying the measurability assumptions (3.1) and (3.2). Let \( 0 < \varepsilon \leq 1, \alpha \geq \alpha(\mathcal{C}, n), v \geq V(\mathcal{C}) \), and \( M > 0 \). Let \( v_1 := v_1(v), \gamma := \exp(-\varepsilon M^2/16av), \)

\[
r^{(1)} := \left| \frac{9\varepsilon^2 M^2}{2^{19} v_1 \gamma \log 2} \right|, \\
r^{(2)} := \left| \frac{3\varepsilon M n^{1/2}}{2^{16} v_1 \log 2} \right|.
\]
If

\begin{align}
(3.65) \quad & \gamma \leq 9\epsilon^2/2^{17}, \\
(3.66) \quad & M \leq 3n^{1/2} \epsilon^2/2^{14}, \\
(3.67) \quad & \gamma \leq \epsilon^2 M^2/2^{12} \log 2, \\
(3.68) \quad & M^2/\alpha \geq \log 2, \\
(3.69) \quad & n \geq 2^{18} \nu 
\end{align}

and

\begin{align}
(3.70) \quad & (2^{13} \nu_1)^{1/2} - r(i)/2 \leq 3\epsilon M/64, \quad i = 1,2,
\end{align}

then

\begin{align}
(3.71) \quad & \mathbb{P} \left[ \sup_{C} \left| \nu_n(C) \right| > M \right] \leq K_3(\nu) \exp\left( -(1-\epsilon)M^2/2\alpha \right).
\end{align}

**Proof.** Let $N$ and $W$ be as in Lemma 3.5. Applying Lemma 3.5, then Lemma 3.6 with $H(i) := P_i, N'$ we get
64.

\[ \mathbb{P} [ \sup_{C} |v_n(C)| > M ] \leq 2K_1(v) \exp(-M^2/2\alpha) \]

\[ + \sup_{x(nN) \notin W} \mathbb{Q} [ \sup_{C} |v^{(1)}_n(C)| > (1 - \frac{\varepsilon}{16})M] \]

\[ \leq 2K_1(v) \exp(-M^2/2\alpha) + 2A(v) \exp(-(1-\varepsilon)M^2/2\alpha) \]

\[ + \sup_{x(nN) \notin W} \mathbb{Q} [ \sup_{C} |v^{(1)}_n(C)| > 3\varepsilon M/32]. \]

We now apply Theorem 3.8 to the last term on the right of (3.72), with \( u := 3\varepsilon M/32, \beta := \gamma, \) and \( \mathcal{C} := C_1(\gamma, C, P_{nN}). \) Then \( V(C) \leq v_1, \) so (3.51), with \( v_1 \) in place of \( v, \) follows from (3.70), (3.52) from (3.67), and (3.53) from (3.66) and (3.68). From (3.65) we have \( u^2/256\beta \geq M^2/2\alpha, \) and from (3.66) we have \( n^{1/2}u/1024 \geq M^2/2\alpha. \) Thus

\[ \mathbb{Q} [ \sup_{C_1(\gamma, C, P_{nN})} |v^{(1)}_n(C)| > 3\varepsilon M/32] \]

\[ \leq 4K_1(v_1) (\exp(-u^2/256\beta) + \exp(-n^{1/2}u/1024)) \]

\[ \leq 8K_1(v_1) \exp(-M^2/2\alpha), \]

and the theorem follows. \( \square \)
From the conditions (3.65) - (3.70) the following is immediate.

**Corollary 3.10.** For \( n \geq 1 \) let \( \mathcal{C}_n \subset \mathcal{A} \) be VC classes satisfying the measurability assumptions (3.1) and (3.2). Suppose \( V(\mathcal{C}_n) \leq v \) for all \( n \). Let \( 0 < \varepsilon \leq 1 \), \( \alpha(\mathcal{C}_n, n) \leq \alpha_n \leq 1/4 \), and \( M_n > 0 \). If

\[
(3.73) \quad (\alpha_n \log \frac{1}{\alpha_n})^{1/2} = o(M_n)
\]

and

\[
(3.74) \quad M_n = o(n^{1/2} \alpha_n),
\]

then

\[
(3.75) \quad \mathbb{P} [\sup_{\mathcal{C}_n} |v_n(\mathcal{C})| > M_n] \leq K_3(v) \exp(- (1-\varepsilon) M_n^2 / 2 \alpha_n)
\]

for all sufficiently large \( n \).

Note that (3.73) and (3.74) imply that \( n^{-1} \log n = o(\alpha_n) \).

In the proof of Theorem 3.9, the use of Lemma 3.5, and of the discrete approximations \( P_{i,N} \) for \( P(i) \), could be avoided by applying Lemma 3.6 directly with \( H(i) = P(i) \) instead of \( H(i) = P_{i,N} \). But we would then need to assume \( \mathcal{C}_1(\gamma, \mathcal{C}, P_{(n)}) \) to be n-deviation-measurable,
a hypothesis which may be a nuisance to verify in specific cases.

In comparing Theorem 3.1 to Theorem 3.9, it should be noted that the condition (3.66) is stronger than the last part of (3.3), but (3.66), (3.69), and (3.70) together may put weaker conditions on $n$ and $\alpha$ asymptotically than (3.49) does. To satisfy (3.49) for large $n$ and $\alpha = \alpha_n$ requires that $n^{-3/13} = o(\alpha_n)$, but (3.66), (3.69), and (3.70) only require that $n^{-1}\log n = o(\alpha_n)$. Corollary 3.10 could not be derived from Theorem 3.1, because (3.73) and (3.74) may hold without (3.49) becoming true for large $n$ and $\alpha = \alpha_n$.

In Theorem 3.1 the requirements that $n \geq n_0$ and $M \geq M_0$ may be dispensed with if we allow $K$ to depend also on $\varepsilon$ and $\alpha$. In particular, letting

\[ M_2(\varepsilon, \alpha, v) := \max(M_0(\varepsilon, \alpha, v), n_0(\varepsilon, \alpha, v)^{1/2}) \]

\[ K_0(\varepsilon, \alpha, v) := \max(K(v), \exp(M_2^2(\varepsilon, \alpha, v)/2\alpha)) \]

we get the following from that theorem.

**Corollary 3.11.** Let $\mathcal{C} \subseteq \mathcal{X}$ be a VC class satisfying the measurability assumptions (3.1) and (3.2).
Let \( v \geq V(C) \), \( \alpha > 0 \), and \( 0 < \varepsilon \leq 1 \). Then there exists a constant \( K_0 = K_0(\varepsilon, \alpha, v) \) such that if

\[
n \geq 1, \ \alpha \geq \alpha(C, n), \text{ and } 0 < M \leq \begin{cases} 
3n^{1/2} \alpha \varepsilon / 8 & \text{if } \alpha < 1/4 \\
\infty & \text{if } \alpha \geq 1/4 
\end{cases}
\]

then

\[
P[\sup_C |v_n(C)| > M] \leq K_0 \exp(-(1-\varepsilon)M^2/2\alpha).
\]

Theorem 3.1 is not true if the quantity \( 3n^{1/2} \alpha \varepsilon / 8 \) in (3.3) is replaced by \( \infty \). To see this, let all \( P(i) \) be the same law \( P \), let \( C \) consist of a single set \( C \) of probability \( \beta < 1/4 \), let \( \varepsilon > 0 \) be small enough so \( \beta > 4^{-1} (1-\varepsilon) (1-\beta) / 3\beta \), and let \( n \geq 1, \ \alpha := \beta (1-\beta), \ M := (2n (\log 4) / 3)^{1/2} (1-\beta), \ \text{and } \ K > 0. \) Then if \( n \) is sufficiently large we have

\[
P[|v_n(C)| > M] \geq P[P_n(C) = 1]
\]

\[
= \beta^n
\]

\[
> K \cdot 4^{-n} (1-\beta) (1-\varepsilon) / 3\beta
\]

\[
= K \exp(-(1-\varepsilon)M^2/2\alpha).
\]
IV. A Bounded LIL for the Normalized Empirical Process

In this chapter we will use Corollary 3.10 to establish a bounded LIL for \( \nu_n \), following a standard technique for such results. Let \( L_x := \log(\max(x, e)) \) and \( LL_x := L(L_x) \).

**Theorem 4.1.** If \( \mathcal{C} \subset \mathcal{Q} \) is a VC class satisfying the measurability conditions (3.1) and (3.2), 
\( \alpha_n := \alpha(\mathcal{C}, n) \), and

\[
(4.1) \quad \log \frac{1}{\alpha_n} = o(\LL_n),
\]

then

\[
(4.2) \quad \limsup_{n} \frac{\| \nu_n \|}{(2\alpha_n \LL_n)^{1/2}} = 1 \quad \text{a.s.}
\]

where \( \| \cdot \| \) denotes the sup norm for functions on \( \mathcal{C} \).

Conversely if \( \beta_n, n \geq 1 \), are positive real numbers such that

\[
(4.3) \quad \LL_n = o(\beta_n),
\]

then there exists a sequence \( P(i), i \geq 1 \), of p.m.'s on the integers \( \mathbb{Z} \) and a VC class \( \mathcal{C} \) of subsets of \( \mathbb{Z} \) such that
(4.4) \[ \log \frac{1}{\alpha_n} = o(\beta_n) \]

but

(4.5) \[ \limsup_n \frac{|\nu_n|}{(2\alpha_n L L n)^{1/2}} = \infty \text{ a.s.} \]

The following lemma is from Stout (1974) p. 262.

**Lemma 4.2.** For each \( \delta > 0 \) there are constants \( \rho(\delta) \) and \( \pi(\delta) \) satisfying the following: Let \( n \geq 1 \) and let \( Z_i, 1 \leq i \leq n \), be independent mean - 0 r.v.'s. Let \( S := \sum_{i=1}^{n} Z_i \) and \( s^2 := \sum_{i=1}^{n} EZ_i^2 \). If \( a > 0 \), \( \rho \geq \rho(\delta) \), \( ap \leq \pi(\delta) \), and \( |z_i| \leq \text{as a.s. for all } i \leq n \), then

\[ \Pr[S/s > \rho] \geq \exp(-(1+\delta)\rho^2/2). \]

**Proof of Theorem 4.1.** To prove the first half of the theorem we first show that

(4.6) \[ \limsup_n \frac{|\nu_n|}{(2\alpha_n L L n)^{1/2}} \leq 1 + 9\delta \text{ a.s. for all } \delta > 0. \]

Fix \( 0 < \delta < 1/36 \) and define

\[ u_n := (2\alpha_n L L n)^{1/2} \]
\[ M_n := (1+2\delta)u_n. \]

By (4.1) we have

\[(4.7) \quad n\alpha_n \uparrow \infty \text{ as } n \to \infty \]

so we can define

\[ n(k) := \min\{n \geq 1 : n\alpha_n \geq (1+4\delta)^k\}, \quad k \geq 1, \]

\[ n(0) := 0. \]

Define the events

\[ A_k := [||\nu_j|| > (1+9\delta)u_j \text{ for some } n(k-1) < j \leq n(k)] \]

\[ B_k := [||\nu_{n(k)}|| > M_n(k)]. \]

We will show that

\[(4.8) \quad \mathbb{P}(A_k) \leq 2\mathbb{P}(B_k), \quad k \text{ large}, \]

and

\[(4.9) \quad \sum_{k=1}^{\infty} \mathbb{P}(B_k) < \infty, \]

so \( \mathbb{P}[A_k \text{ i.o.}] = 0 \) and (4.6) will follow.

All further inequalities in this proof should be interpreted as being valid for sufficiently large (but
nonrandom) \( k \).

Fix \( k \geq 1 \) and define

\[
\tau := \min \{ j > n(k-1) : \|v_j\| > (1+9\delta)u_j \} \leq \infty
\]

\[
\mathbb{P}(L, m) := \prod_{i=l}^{m} \mathbb{P}(i), \quad l \leq m.
\]

Then

\[
\mathbb{P}(B_k) \geq \sum_{j=n(k-1)+1}^{n(k)} \mathbb{P}(B_k \cap [\tau = j])
\]

\[
= \sum_{j=n(k-1)+1}^{n(k)} \int_{\mathbb{R}^j} 1_{[\tau = j]} \prod_{i=n(k)-j+1}^{1} dB_{k,i}
\]

Fix \( j, n(k-1)+1 \leq j \leq n(k) \), and \( x(j) \in [\tau = j] \). Then there exists \( C \in \mathbb{C} \) with \( |v_j(C)| > (1+9\delta)u_j \). Define

\[
Y_1 := (\delta_{x_j} - \mathbb{P}(i))(C)
\]

for \( j+1 \leq i \leq n(k) \), and

\[
U := \sum_{i=j+1}^{n(k)} Y_i.
\]

Then

\[
n(k)^{1/2}v_{n(k)}(C) = j^{1/2}v_j(C) + U
\]

so \( |U| \leq j^{1/2}(1+9\delta)u_j - n(k)^{1/2}M_{n(k)} \) imposes

\[
|v_{n(k)}(C)| > M_{n(k)}.
\]

It follows that
72.

\[
\int_{x^{n(k)-j} \mathcal{B}_k} 1 \, d\mathcal{P}(j+1,n(k))
\]

(4.11)

\[
\geq \int_{x^{n(k)-j} \mathcal{B}_k} \frac{1}{|u|^j} (1+\delta)u_j^{1/2} (1+9\delta)u_j - n(k)1/2 M_n(k) \, d\mathcal{P}(j+1,n(k)).
\]

Now by (4.1),

(4.12) \quad \text{LLn}(k) \sim \text{LL}(n(k) \mathcal{a}_n(k)) \sim \text{LL}(1+4\delta)^k \sim \log k

as \ k \to \infty, so

(4.13) \quad n(k)u_{n(k)}^2 \sim (1+4\delta)^k \log k

and then

(4.14) \quad n(k-1)u_{n(k-1)}^2 \sim (1+4\delta)^{-1} n(k)u_{n(k)}^2.

It follows, using (4.7), that

\[
(j(1+9\delta)^2 u_j^2 \geq (1+9\delta)^2 n(k-1) u_{n(k-1)}^2
\]

\[
\geq (1+9\delta)^2 (1-4\delta) n(k) u_{n(k)}^2
\]

\[
\geq (1+\delta)^2 n(k) M_n^2(k)
\]

and then that
73.

\[ \frac{1}{2} \left( 1 + \theta \right) u_j - n(k) \frac{1}{2} M_k \geq \theta n(k) \frac{1}{2} M_n(k) \]

\[ \geq \left( 2n(k) \alpha_n(k) \right)^{1/2} \]

\[ \geq (2\theta \nu^2)^{1/2}. \]

Hence by Chebyshev's inequality, the right side of (4.11) is at least 1/2. Since \( x^{(j)} \in [\tau = j] \) is arbitrary, (4.10) then gives

\[ \mathbb{P}(B_k) \geq \frac{1}{2} \sum_{j=n(k-1)+1}^{n(k)} \mathbb{P} \left[ \tau = j \right] \int_{x^j}^{x^{j+1}} \mathbb{P}(1,j) \]

\[ = \frac{1}{2} \sum_{j=n(k-1)+1}^{n(k)} \mathbb{P} \left[ \tau = j \right] \]

\[ = \mathbb{P}(A_k)/2 \]

and (4.8) is proved.

To prove (4.9) we apply Corollary 3.10. (3.73) follows from (4.1), and (3.74) is clear. Thus, using (4.12),

\[ \mathbb{P}(B_k) \leq K_3(V(C)) \exp \left( - \frac{(1-\theta) M^2_n(k)}{2 \alpha_n(k)} \right) \]

\[ \leq K_3(V(C)) \exp \left( - \frac{(1+2\theta) \ln(k)}{} \right) \]

\[ \leq K_3(V(C)) k^{-(1+\theta)} \]
and (4.9), and then (4.6), follow.

To prove the reverse inequality in (4.2) we will use Lemma 4.2. Again fix $0 < \delta < 1/36$. For $C \in \mathcal{C}$ and $n \geq 1$ define $\gamma_n(C) := n \var{\nu_n(C)}$ so

$$\sup_C \gamma_n(C) = n\alpha_n.$$  

(4.15)

Fix $T > 2/\delta$ and this time define $n(k) := \min\{n \geq 1: n\alpha_n \geq T^k\}$, $n(0) := 0$. By (4.15) for each $k \geq 1$ there exists $C_k \in \mathcal{C}$ with

$$\gamma_n(k)(C_k) \geq (1-\delta) n(k) \alpha_n(k).$$  

(4.16)

For $k \geq 1$ define

$$Z_{k,i} := \left( \delta X_{n(k-1)+i} - P(n(k-1)+i))(C_k) \right), \quad i \geq 1 \; \text{with}$$

$$m(k) := n(k) - n(k-1)$$

$$S_k := \sum_{i=1}^{m(k)} Z_{k,i}$$

$$s_k^2 := \sum_{i=1}^{m(k)} E Z_{k,i}^2$$

$$a_k := 1/s_k$$
75.
\[ \rho_k := \frac{1-\delta}{1-\delta} (2LLn(k))^{1/2} \]

\[ W_k := n(k)^{-1/2} s_k \]

\[ r_k := (1-\delta)u_n(k). \]

Similarly to (4.14) we have

\[ (4.17) \quad n(k-1)u_n^2(k-1) \prec T_n^{-1} n(k)u_n^2(k) \quad \text{as } k \to \infty, \]

while \( T_n^{-1} \prec \delta/2 < \delta(1-\delta), \) so

\[ s_k^2 = \gamma_n(k)(C_k) - \gamma_n(k-1)(C_k) \]

\[ (4.18) \quad \geq (1-\delta)n(k)\alpha_n(k) - n(k-1)\alpha_n(k-1) \]

\[ \geq (1-\delta)^2 n(k)\alpha_n(k). \]

Let \( \pi(\delta) \) and \( \rho(\delta) \) be as in Lemma 4.2; from (4.18) and (4.1) we have \( a_k^2 \rho_k^2 \leq \pi(\delta), \) while clearly \( \rho_k^2 \geq \rho(\delta). \) Similarly to (4.12) we have \( LLn(k) \sim \log k, \) which along with Lemma 4.2 and (4.18) gives
\[ \mathbb{P}[W_k > r_k] = \mathbb{P}[S_k / s_k > n(k)^{1/2}r_k / s_k] \]
\[ \geq \mathbb{P}[S_k / s_k > r_k / (1-\delta)^{1/2}n(k)] \]
\[ = \mathbb{P}[S_k / s_k > \rho_k] \]
\[ \geq \exp(-(1+\delta)(1-7\delta)^2(1-\delta)^{-2}Ln(k)) \]
\[ \geq k^{-(1-\delta)}. \]

Since the \( W_k \) are independent it follows that

\[ \limsup_k W_k / u_{n(k)} \geq 1 - 7\delta \quad \text{a.s.} \] (4.19) Using (4.6) we have

\[ \limsup_k \frac{|\nu_{n(k)}(C_k) - W_k|}{u_{n(k)}} \]
\[ = \limsup_k \frac{n(k)^{1/2}|\nu_{n(k-1)}(C_k)|}{u_{n(k)}} \]
\[ \leq \limsup_k \frac{n(k-1)^{1/2}u_{n(k-1)}}{n(k)^{1/2}u_{n(k)}} \]
\[ \leq \limsup_k \frac{n(k-1)^{1/2}u_{n(k-1)}}{n(k)^{1/2}u_{n(k)}} \]
\[ \leq \frac{n(k-1)^{1/2}}{n(k)^{1/2}} \]
\[ \leq T^{-1/2} \]
\[ \leq \delta^{1/2}. \]
Since $\delta$ is arbitrary, (4.6), (4.19), and (4.20) prove (4.2).

Proceeding to the converse, let $\{\beta_n\}$ be as in (4.3). We take $C$ to be $\{\{i\}: i \geq 1 \in \mathbb{Z}\}$, and define $b_n := \beta_n / L \ln n$, $a_n := \min b_j$, so $b_n \to \infty$ and $a_n \uparrow \infty$. Define inductively $R_1 := 1$ and

$$R_n := \min\left(\frac{a_n^{1/4}}{n^{1/4}}, \left(\frac{n(n-1)}{R_{n-1}}\right)^{1/2}, R_{n-1}\right), \quad n \geq 1,$$

so $R_n \to \infty$ and $n/R_n^2 \uparrow \infty$. Let

$$\ell(k) := \left\lceil \exp\left(\exp\left(16k(\log 2)/3R_k^2\right)\right) \right\rceil, \quad k \geq 1,$$

$$j(k) := 2^k \ell(k),$$

$$Q(n) := \frac{1}{2}\ell_n + \frac{1}{2}\ell_{n'} - n,$$

where $\lfloor \cdot \rfloor$ denotes the integer part. Then $\ell(k) \uparrow \infty$, so we can choose a sequence $P(i)$, $i \geq 1$, of p.m.'s on $\mathbb{Z}$ such that exactly $\ell(k)$ of the first $j(k)$ $P(i)$'s are equal to $Q(n)$ for each $1 \leq n \leq 2^k$, for all $k \geq 1$.

Since $\ell(k+1) - \ell(k)$ of the p.m.'s $\{P(i): j(k) < i \leq j(k+1)\}$ are equal to $Q(k)$, we may take the sequence to be such that

$$P(j(k)+1) = P(j(k)+2) = \cdots = P(j(k)+\ell(k+1)-\ell(k)) = Q(k)$$

for all $k \geq 1$. 
For \( k > 1 \) let \( Y_k \) be an r.v. with binomial distribution with parameters \( \ell(k) \) and 1/2. Writing \( \nu_n(i) \) for \( \nu_n({i}) \), we have that \( \{\nu_n(i): i > 1\} \) are independent r.v.'s for fixed \( n \), with

\[
\mathcal{L}(\nu_n(i)) = \mathcal{L}(j(k)^{-1/2}(Y_k - \ell(k)/2)).
\]

Now, using (4.21),

\[
\alpha_n = \sup_{i \geq 1} \frac{1}{4n} \text{card}(\{m: P(m) = Q(i), m \leq n\})
\]

\[
= \frac{1}{4n} \text{card}(\{m: P(m) = Q(1), m \leq n\})
\]

(4.22)

\[
\begin{cases}
(\ell(k) + n - j(k))/4n & \text{if } j(k) \leq n < j(k) + \ell(k+1) - \ell(k) \\
\ell(k+1)/4n & \text{if } j(k) + \ell(k+1) - \ell(k) < n < j(k+1)
\end{cases}
\]

In particular

(4.23) \( \alpha_j(k) = \ell(k)/4j(k) = 2^{-(k+2)} \).

Let

\[
P_k := P[\nu_j(k)(1) \geq R_k u_j(k)]
\]

\[
= Pr[\ell(k)^{-1/2}Y_k - \ell(k)/2 \geq R_k(\ell(j(k))^{1/2}/2)]
\]

using (4.23). Define events \( D_k \) by
Then the $D_k$ are independent, with

$$P(D_k) = 1 - (1-p_k)^{2^{k-1}}.$$  

Applying Lemma 4.2 to the definition of $p_k$, with $n := \ell(k)$, $s = \ell(k)^{1/2}/2$, $\delta := 1/4$, $\rho := R_k(LLj(k))^{1/2}/2$, and $a := \ell(k)^{-1/2}$, we get (for large $k$, as always):

$$p_k \geq \exp(-5R_k^2(LLj(k))/32)$$

$$\geq \exp(-3R_k^2(LL\ell(k))/16)$$

$$\geq \exp(-k \log 2).$$

Hence $\sum_{k=1}^{\infty} 2^{k-1} p_k = \infty$, so by (4.24), $\sum_{k=1}^{\infty} P(D_k) = \infty$.

Thus $P[D_k \text{i.o.}] = 1$, and (4.5) follows.

To prove (4.4), set

$k(n) := k$ for $j(k) \leq n < j(k+1)$, $k \geq 1$.

Using (4.22) and (4.23), we have if

$n \leq j(k(n)) + \ell(k(n)+1) - \ell(k(n))$ that

$$\alpha_n = \frac{1}{4} \ell(k(n)) + n - \frac{j(k(n))}{j(k(n)) + n - j(k(n))} \geq \frac{\ell(k(n))}{4j(k(n))} = \alpha_j(k(n))$$
while if \( n > j(k(n)) + \lambda(k(n)+1) - \lambda(k(n)) \) we have

\[
\alpha_n = \frac{\lambda(k(n)+1)}{4n} > \frac{\lambda(k(n)+1)}{4j(k(n)+1)} = \alpha_j(k(n)+1) = \alpha_j(k(n))/2.
\]

Hence using (4.23)

\[
\log \alpha_n^{-1} = O(\log \alpha_j^{-1}(k(n)))
\]

\[
= O(k(n))
\]

\[
= O(R_{k(n)}^2 \lambda \lambda(k(n)))
\]

\[
= O(a_{k(n)}^{1/2} \lambda \lambda j(k(n)))
\]

\[
= O(a_n^{1/2} \lambda \lambda n)
\]

\[
= o(\beta_n).
\]
V. Weighted Empirical Processes with Truncation

In this chapter we study weighted empirical processes, with the weight at each set depending only on the size of that set; that is, we consider processes \( \nu_n / q \circ \bar{F}(n) \) on \( C \subset \mathcal{C} \), where \( q \) is a non-negative continuous function on \([0,1]\). In particular we will prove a CLT for a modification of such processes.

If \( q(\bar{F}(n)(C)) = 0 \) for some \( C \in \mathcal{C} \) with \( 0 < \bar{F}(n)(C) < 1 \), then \( \nu_n(C)/q(\bar{F}(n)(C)) \) is infinite with positive probability. Therefore we restrict our attention to functions \( q \) with \( q > 0 \) on \((0,1)\).

As discussed in the introduction, we will be interested in the behavior of \( q \) near 0 and 1. Since \( \nu_n(C) = -\nu_n(C^c) \), studying \( \nu_n / q \circ \bar{F}(n) \) on \( C \) is essentially the same as studying it on \( \{C^c : C \in \mathcal{C} \} \). Therefore we need only consider the behavior of \( q \) near 0, not near 1. Thus we define

\[
Q := \{ q \in C[0,1] : q > 0 \text{ on } (0,1) \}
\]

and suppose \( q \in Q \); all our results are easily modified to include the case where both \( q(0) \) and \( q(1) \) may be 0.

Example 5.1
Suppose \( \mathcal{C} \) consists of all subintervals of \( X := [0,1] \), all \( P(i) \) are the uniform law \( P \) on \([0,1]\), and \( q(0) = 0 \). Let \( \{C_k\} \) be a sequence of
intervals shrinking to the (random) point $X_1$. Then
\[ \lim sup_{k} \nu_n(C_k) \geq n^{-1/2} \text{ for all } n, \text{ but } q(P_n(C_k)) \to 0 \]
as $k \to \infty$ for all $n$. Thus $\nu_n/q_0P(n)$ is unbounded a.s.; Theorem 2.9, in particular (2.5), shows we cannot therefore expect a CLT to hold.

To avoid this difficulty we "truncate" the process $\nu_n/q_0P(n)$, replacing it with 0 on "small" sets $C$. Thus for $0 < \tau \leq 1$ and $n \geq 1$ define

\[
\begin{align*}
\tau_{\tau,n}(C) := \begin{cases} 
\nu_n(C)/q(P_n(C)) & \text{if } P_n(C) \geq \tau \\
0 & \text{if } P_n(C) < \tau.
\end{cases}
\end{align*}
\]

We consider sequences $\tau(n) \to 0$ and look for convergence in law of $\tau_{\tau(n),n}$ to a Gaussian process. An alternative way around the difficulty in Example 5.1 will be considered in Chapter 6.

Whenever an expression appears which evaluates as $0/q(0)$ or $0/q^2(0)$ (e.g., $\nu_n(C)/q(P_n(C))$ when $\nu_n(C) = q(P_n(C)) = 0$) we define this by convention to be 0. In all situations where this arises it will be apparent from the proofs that this is consistent with definition by continuity.
For $\mathcal{C} \subseteq \mathcal{A}$, $t > 0$, and $H$ a law on $(X, \mathcal{A})$, define

$$\mathcal{C}_2(t, \mathcal{C}, H) := \{C \in \mathcal{C} : H(C) < t\}.$$ 

For $\delta > 0$ we say $\mathcal{C} \subseteq \mathcal{A}$ is size-$\delta$ measurable (with constants) for $\{P(i), i \geq 1\}$ (or "for $P$") if $\mathcal{C}$, $\mathcal{C}_1(t, \mathcal{C}, \mathcal{P}(n))$, and $\mathcal{C}_2(t, \mathcal{C}, \mathcal{P}(n))$ are each $n$-deviation-measurable (with constants in the case of $\mathcal{C}_2(t, \mathcal{C}, \mathcal{P}(n))$) for $\{P(i), i \geq 1\}$ (or for $P$) for all $t \in (0, \delta]$ and $n \geq 1$.

Given a sequence $\{\psi_n\}$ of $l^\infty(\mathcal{C})$-valued r.v.'s on $(X, \mathcal{A}, \mathcal{P})$ and a law $P$ on $(X, \mathcal{A})$, we say a subspace $D$ of $l^\infty(\mathcal{C})$ is admissible for $\{\psi_n\}, P$, and $\{P(i), i \geq 1\}$ if $\psi_n \in D$ a.s. for all $n$, $\psi_n$ is $\mathcal{P}$-completion measurable into $(D, \mathcal{B}_D)$ for all $n$, and $D$ contains $C_{b,u}(\mathcal{C}, d_P)$.

At times we will need regularity conditions on $q$ to facilitate our proofs; for these occasions we define

$$Q_1 := \{q \in Q : q(t) \text{ non-decreasing and } q(t)/t \text{ non-increasing on } [0, \delta] \text{ for some } \delta > 0\}.$$ 

Define finally

$$\rho_n(A, B) := \text{cov}_P(\nu_n(A), \nu_n(B)) = \frac{1}{n} \sum_{i=1}^{n} P(i)(A \cap B) - P(i)(A)P(i)(B),$$

(5.1)
\( \rho_P(A,B) := \text{cov}(G_P(A), G_P(B)) \)

\[ = P(A \cap B) - P(A)P(B) \]

for all \( A, B \in \mathcal{C} \).

**Theorem 5.2.** Suppose \( q \in \mathcal{Q}_1 \) satisfies

(5.2) \( q^2(t)/t \log \frac{1}{t} \to \infty \) as \( t \to 0 \).

Let \( \mathcal{C} \subseteq \mathcal{Q} \) be a VC class which is size-\( \delta \) measurable for \( P \) for some \( \delta > 0 \). Let \( \tau(n) \to 0 \) with either

(5.3) \( n^{-1/2} \log n = o(q(\tau(n))) \)

or

(5.4) \( n^{-1} \log n = O(\tau(n)) \).

Let \( P \) be a law on \( (X, \mathcal{A}) \), and let \( D \) be a subspace of \( \mathcal{L}^\infty(\mathcal{C}) \) admissible for \( \{\zeta_{\tau(n)}, P\} \), and \( \{P_i\} \). Suppose that

(5.5) for every \( \lambda > 0 \) there exist \( \delta > 0 \) and \( n_0 \) such that \( \widehat{P}(n)(A \Delta B) < \lambda \) whenever \( n \geq n_0 \), \( A, B \in \mathcal{C} \), and \( P(A \Delta B) < \delta \),

(5.6) \( \rho_n(A,B) \to \rho_P(A,B) \) for all \( A, B \in \mathcal{C} \),

and
(5.7) \( q(\overline{P}(n)(C)) + q(P(C)) \) for all \( C \in \mathcal{C} \).

Then \( \xi_{\tau}(n) \) converges in law to \( G_P/q_P \) in \( (D, \mathcal{B}_D) \).

To satisfy (5.6) and (5.7) it is sufficient that
\[
\overline{P}(A \cap B) + P(A \cap B) \quad \text{for all } A, B \in \mathcal{C}.
\]

In the identically distributed case, (5.5), (5.6), and (5.7) always hold.

The condition (5.2) is natural because, by Theorem 2.2 of Dudley (1973) and Lemma 2.5 above, \( \psi(t) := (t \log \frac{1}{t})^{1/2} \) is a sample modulus for \( G_P \). In fact, in the special case of Example 5.1, we have \( G_P([a,b]) = W_0(b) - W_0(a) \), where \( W_0 \) is a Brownian bridge. By a theorem of Lévy (1937, 1954) \( \psi \) is the best possible sample modulus for \( W_0 \), hence also for \( G_P \). Thus \( \psi \) is in some cases the best possible sample modulus for \( G_P \) on a VC class \( \mathcal{C} \).

In the special case of Example 5.1, our Theorem 5.2 is related to Theorem 1.2 of Shorack and Wellner (1982).

If \( q(0) > 0 \) then (5.3) is satisfied with \( \tau(n) = 0 \) for all \( n \). In particular, in case \( q \equiv 1 \), Theorem 5.2 provides a CLT for the unmodified processes \( v_n \), generalizing Theorem 7.1 of Dudley (1978).

We point out the similarity between the conditions (5.2), (5.3), and (5.4) and the hypotheses of Corollary 3.10.
Lemma 5.3. Suppose \( q \in Q \) satisfies

\[
q^2(t)/t \to \infty \quad \text{as} \quad t \to 0.
\]

Let \( C \subseteq Q \) be a VC class which is size-\( \delta \) measurable for some \( \delta > 0 \), let \( \tau(n) \to 0 \), and let \( P \) be a law on \((X,\mathcal{A})\). Suppose (5.5), (5.6), and (5.7) hold, and suppose

\[
(5.9) \quad \text{for every } \theta, \eta > 0 \text{ there exist } \delta > 0 \text{ and } n_1
\]
such that

\[
P^*[\sup\{|\psi_n(C)|/q(\overline{P}(n)(C)) : C \in C, \tau(n) < \overline{P}(n)(C) < \delta\} > \theta] < \eta
\]

for all \( n \geq n_1 \).

If not all \( P_i \) are equal, suppose also that (5.3) holds.

Then (2.5), (2.6), and (2.7) hold for \( \psi_n = \zeta_{\tau(n)}, n \).

Proof. \( C \) is totally bounded for \( d_P \) by Lemma 2.5, so we must verify (2.5) and (2.7).

We begin with (2.7). Using (5.3) in the non-i.i.d. case and Corollary 18.2 of Bhattacharya and Rao (1976), we see that it is sufficient to show

\[
(5.10) \quad \text{cov}(\zeta_{\tau(n)}, n(A), \zeta_{\tau(n)}, n(B)) \to \frac{\rho_P(A,B)}{q(P(A))q(P(B))}
\]
as \( n \to \infty \) for all \( A, B \in C \). Now
\[ (5.11) \quad \text{cov}(\zeta_{t(n)}(n), n(A), \zeta_{t(n)}(n), n(B)) = \begin{cases} \frac{\rho_n(A,B)}{q(n(A))q(n(B))} & \text{if } \bar{P}(n)(A) \geq \tau(n) \text{ and } \\
 & \bar{P}(n)(B) \geq \tau(n) \\
& 0 & \text{otherwise,} \end{cases} \]

so if neither \( \bar{P}(n)(A) < \tau(n) \) infinitely often nor \( \bar{P}(n)(B) < \tau(n) \) infinitely often, then (5.10) follows from (5.6), (5.7) and (5.8). Thus suppose \( \bar{P}(n(k))(A) < \tau(n(k)) \) for some subsequence \( \{n(k)\} \).

Case 1: \( q(0) > 0 \). Since \( \rho_n(k)(A,B) \leq \bar{P}_n(k)(A) \), we have by (5.6) that \( \rho_n(A,B) + 0 = \frac{\rho_n(A,B)}{q(P(A))} \), so (5.10) follows from (5.11) since \( q \) is bounded away from 0.

Case 2: \( q(0) = 0 \). Then \( q(\bar{P}(n(k))(A)) + 0 \) as \( k \to \infty \), so we have by (5.7) that \( q(\bar{P}(n)(A)) + 0 = q(P(A)) \), so \( \bar{P}(n)(A) + 0 \). But

\[
\frac{|\rho_n(A,B)|}{q(\bar{P}(n)(A))q(\bar{P}(n)(B))} \leq \frac{\bar{P}(n)(A)^{1/2}}{q(\bar{P}(n)(A))} \frac{\bar{P}(n)(B)^{1/2}}{q(\bar{P}(n)(B))}
\]

so (5.10) follows from (5.11) and (5.8).
We proceed to (2.5). Fix $\varepsilon > 0$. Suppose there exist $\beta > 0$ and $n_2$ such that
\[(5.12) \quad \mathbb{P}^*\left[\xi_{\tau}(n), n \in B_{\beta, \varepsilon}(\bar{F}(n))\right] < \varepsilon \text{ for all } n \geq n_2.\]

By (5.5) we can find $\delta > 0$ such that $B_{\delta, \varepsilon}(\mathcal{P}) \subseteq B_{\beta, \varepsilon}(\bar{F}(n))$ for large $n$, and (2.5) follows. Thus it is sufficient to prove (5.12).

By (5.9) we can find $\delta > 0$ and $n_1$ such that $\tau(n) < \delta/2$ if $n \geq n_1$ and
\[(5.13) \quad \mathbb{P}^*\left[\sup_{C_2(\delta, \bar{C}, \bar{F}(n))} |\xi_{\tau}(n), n(C)| > \varepsilon/2\right] < \varepsilon/2 \text{ for all } n \geq n_1.\]

By Theorem 3.8, applied with $\beta = 1$, we can choose $M > 0$ such that
\[(5.14) \quad \mathbb{P}\left[\sup_{C} |v_n(C)| > M\right] < \varepsilon/4 \text{ for all } n \geq 1.\]

Let $w := \inf\{q(t) : \delta/2 \leq t \leq 1\}$, $u := \varepsilon w/2$, and $v := V(\mathcal{C}_1(1, \bar{C}, \mathcal{P}))$. Choose $0 < \alpha < \delta/2$ small enough so
\[(5.15) \quad s, t \in [\delta/2, 1], |s-t| < \alpha \text{ imply } |1/q(s) - 1/q(t)| < \varepsilon/2M.\]

We wish to choose a $\beta > 0$ and apply Theorem 3.8 to $\mathcal{C}_1(\beta, \bar{C}, \bar{F}(n))$. Fix $\beta > 0$ small enough so (3.52) holds,
\( \beta < \alpha \), (3.51) holds for \( i = 1 \) (with \( r^{(1)} \) as in (3.50)), \( \mathbb{C} \) is size-\( \beta \) measurable, and

\[(5.16) \quad 4K_1(v) \exp(-u^2/256\beta) \leq \varepsilon/8.\]

Choose \( n_2 > n_1 \) large enough so (3.53) holds, (3.51) holds for \( i = 2 \), and

\[(5.17) \quad 4K_1(v) \exp(-\frac{1}{2}u/1024) \leq \varepsilon/8 \quad \text{for all} \quad n \geq n_2.\]

By Theorem 3.8, (5.16) and (5.17) we have

\[(5.18) \quad P[\sup_{C_1(\mathbb{B},C,F(n))} |\nu_n(C)| > \varepsilon w/2] \leq \varepsilon/4 \quad \text{for all} \quad n \geq n_2.\]

Fix \( n \geq n_2 \) and suppose there exist \( A, B \in \mathbb{C} \) with \( F(n)(A) \geq \delta/2, F(n)(B) \geq \delta/2, F(n)(A \Delta B) < \beta \), and

\[|\xi_{\tau(n),n}(A) - \xi_{\tau(n),n}(B)| > \varepsilon.\]

Then since \( \beta < \alpha \) and \( \tau(n) < \delta/2 \) we have using (5.15):

\[
\varepsilon < \left| \frac{\nu_n(A)}{q(F_n(A))} - \frac{\nu_n(B)}{q(F_n(B))} \right|
\]

\[(5.19) \quad = \left| \frac{\nu_n(A) - \nu_n(B)}{q(F_n(A))} + \nu_n(B) \left( \frac{1}{q(F_n(A))} - \frac{1}{q(F_n(B))} \right) \right|
\]

\[\leq w^{-1} \max( |\nu_n(A \setminus B)|, |\nu_n(B \setminus A)| ) + (\varepsilon/2M) |\nu_n(B)|.\]
Now since $\beta < \delta/2$, if $\overline{P}(n)(A \Delta B) < \beta$ then either

$\overline{P}(n)(A) \geq \delta/2$ and $\overline{P}(n)(B) \geq \delta/2$ or $\overline{P}(n)(A) < \delta$

and $\overline{P}(n)(B) < \delta$. Hence we get from (5.13), (5.19), (5.18), and (5.14):

\[
P^* [\xi_{\tau(n)}, n \in B_{\beta, \varepsilon}(\overline{P}(n))]
\]

\[
\leq P^* [\sup_{C_2(\delta, \overline{C}, \overline{P}(n))} |\xi_{\tau(n)}, n(C)| > \varepsilon/2]
\]

\[
+ P^* [\sup \{ |\xi_{\tau(n)}, n(A) - \xi_{\tau(n)}, n(B)| : A, B \in C, \overline{P}(n)(A) \geq \delta/2, \overline{P}(n)(B) \geq \delta/2, \overline{P}(n)(A \Delta B) < \beta \} > \varepsilon]
\]

\[
\leq \varepsilon/2 + P [\sup_{C_1(\beta, \overline{C}, \overline{P}(n))} |\nu_n(C)| > \varepsilon \lambda/2]
\]

\[
+ P [\sup_{C} |\nu_n(C)| > M]
\]

\[
< \varepsilon
\]

and (5.12) follows.

For $q \in Q_1$ define

\[q^{-1}(t) := \sup \{ s \in [0, 1] : q(s) \leq t \}, \ t \geq 0,\]

where we take $\sup \emptyset$ to be 0.
The following lemma summarizes a calculation, based on facts from O'Reilly (1974), which we will use several times.

**Lemma 5.4.** Let $q \in Q_1$ with $q(0) = 0$ and

\[(5.20) \quad \int_0^1 \exp\left(-\frac{q^2(t)}{t}\right) \frac{dt}{t} < \infty \quad \text{for all } \varepsilon > 0.\]

Let $u > 0$, $\theta > 0$, and $0 < \lambda < 1/3$. Then there exists a constant $\delta_0(u, \theta, \lambda, q)$ such that for $0 < \delta < \delta_0$ and $t_j := q^{-1}(q(\delta) \lambda^j)$,

\[(5.21) \quad \sum_{j=0}^{\infty} \exp(-u q^2(t_j)/t_j) < \theta.\]

Furthermore if $n > 1$ and $N > 0$ (with $\delta$ now arbitrary), and

\[(5.22) \quad u^{-1}(\lambda^{-1} - 1)^{-1} \log 2 \leq M \leq n^{1/2} q(\delta) \lambda^N,\]

then

\[(5.23) \quad \sum_{j=0}^{N} \exp(-u n^{1/2} q(t_j)) \leq 2 \exp(-u M).\]

**Proof.** By (5.20) we can fix $\delta_0$ small enough so $q(t)$ and $q(t)/t$ are monotone in $[0, q^{-1}(q(\delta)/\lambda)]$ and
92.

\[
(5.24) \quad \frac{2}{1+\lambda} \int_0^{q^{-1}(q(\delta_0)/\lambda)} \exp(-uq^2(t)/t) \frac{dt}{t} < \theta.
\]

Since \( q \in Q_1 \) and \( \lambda < 1/3 \) we have

\[
q((1-\lambda)t_{j-1}/2) \geq (1-\lambda)q(t_{j-1})/2 > q(\delta)\lambda^j, \quad j \geq 0,
\]

so

\[
t_j \leq (1-\lambda)t_{j-1}/2
\]

and hence

\[
\frac{2(t_{j-1}-t_j)}{(1+\lambda)t_{j-1}} \geq 1.
\]

Using this and (5.24) we get if \( \delta \leq \delta_0 \):

\[
\sum_{j=0}^{\infty} \exp(-uq^2(t_j)/t_j)
\]

\[
\leq \frac{2}{1+\lambda} \sum_{j=0}^{\infty} \frac{t_{j-1}-t_j}{t_{j-1}} \exp(-u_\lambda^2 q^2(t_{j-1})/t_j)
\]

\[
\leq \frac{2}{1+\lambda} \sum_{j=0}^{\infty} \int_{t_j}^{t_{j-1}} \exp(-u_\lambda^2 q^2(t)/t) \frac{dt}{t}
\]

\[
< \theta.
\]

To prove (5.23), we have, using (3.58) and (5.22) and reversing the order of summation:
\[ \sum_{j=0}^{N} \exp(-uM \lambda^{-j}) \]

\[ \leq \sum_{j=0}^{N} \exp(-uM(1 + (\lambda^{-1} - 1)j)) \]

\[ = \exp(-uM)/(1 - \exp(-(\lambda^{-1} - 1)uM)) \]

\[ \leq 2 \exp(-uM). \]

Lemma 5.5. Suppose \( q \in Q_1 \) satisfies

\[ q^2(t)/t \log \frac{1}{t} \to \infty \quad \text{as} \quad t \to 0. \]

Let \( \mathcal{C} \subset \mathcal{G} \) be a VC class which is size-\( \delta \) measurable for some \( \delta > 0 \). Let \( \tau(n) \to 0 \) with

\[ n^{-1/2} \log n = o(q(\tau(n))). \]

Then (5.9) holds.

Proof. If \( q(0) > 0 \) then (5.9) is an easy consequence of Theorem 3.9. Hence we assume \( q(0) = 0 \).
Fix \( \theta, \eta > 0 \). Let \( v := V(\mathcal{C}) \), fix \( 0 < \lambda < 1/3 \), and let \( M > 0 \) be large enough so

\[
(5.27) \quad 1024(\theta \lambda)^{-1}(\lambda^{-1} - 1)^{-1} \log 2 \leq M
\]

\[
(5.28) \quad 8K_1(v)\exp(-\theta \lambda M/1024) < \eta/2.
\]

Let \( \delta > 0 \) and \( t_j := q^{-1}(q(\delta)\lambda^j), j \geq 0 \); since (5.25) implies (5.20), by Lemma 5.4 and (5.25) we can take \( \delta \) to be small enough so \( q \) is size-\( \delta \) measurable and

\[
(5.29) \quad q \text{ is monotone on } [0,\delta] \text{ and } t_j \in [0,\delta] \text{ for all } j,
\]

\[
(5.30) \quad \theta^2 \lambda^2 q^2(t)/t \geq 36 \log 2 \text{ for all } t \leq \delta,
\]

\[
(5.31) \quad \left[ \frac{\theta^2 \lambda^2 q^2(t)}{512vt \log 2} \right] \geq \log \frac{1}{\log 2} \text{ for all } t \leq \delta,
\]

\[
(5.32) \quad t^{-1/2}q(t) \geq (2^{15}v)^{1/2}/\theta \lambda \text{ for all } t \leq \delta,
\]

and

\[
(5.33) \quad \sum_{j=0}^{\infty} \exp(-\theta^2 \lambda^2 q^2(t_j)/256t_j) < \eta/8K_1(v)
\]

where \([ \cdot ]\) denotes the integer part. By (5.26) we can find \( n_1 \) such that for \( n \geq n_1 \),
(5.34) \( t(n) < \delta \),

(5.35) \( \theta \lambda n^{1/2} q(t(n)) \geq 512 \log 2 \),

(5.36) \( \left[ \frac{\theta \lambda q(t(n)) n^{1/2}}{2048v \log 2} \right] > \frac{\log n}{\log 2} \),

(5.37) \( M < n^{1/2} q(t(n)) \),

and

(5.38) \( (2^{13} v)^{1/2} n^{-1/2} < \theta \lambda q(t(n))/2 \).

Fix \( n \geq n_1 \). By (5.29) and (5.34) there is an integer \( N \geq 0 \) such that

(5.39) \( t_j \geq \tau(n) \) if and only if \( j \leq N \).

Hence by (5.29),

\[
P^*[\sup\{|\nu_n(C)|/q(\bar{P}(n)(C)) : C \subseteq \mathbb{C}, t(n) \leq \bar{P}(n)(C) < \delta\} > \theta]
\leq \sum_{j=0}^{N} P^*[\sup\{|\nu_n(C)| : C \subseteq \mathbb{C}, t_{j+1} \leq \bar{P}(n)(C) < t_j\} > \theta q(t_{j+1})]
\leq \sum_{j=0}^{N} P[\sup_{C} |\nu_n(C)| > \theta \lambda q(t_j)].
\]
Fix \( j \leq N \); we wish to apply Theorem 3.8 to the right side of (5.40), with \( \beta := t_j \) and \( u := \theta \lambda q(t_j) \).

(3.52) follows from (5.30), and (3.53) from (5.35). Let \( r^{(1)} \) and \( r^{(2)} \) be as in (3.50). Then by (5.31),

\[
 r^{(1)} > \log t_j^{-1} / \log 2
\]

so by (5.32),

\[
 (2^{13\nu})^{1/2} - r^{(1)} / 2 \leq (2^{13\nu})^{1/2} t_j^{1/2} \leq \theta \lambda q(t_j) / 2,
\]

and (3.51) follows for \( i = 1 \). By (5.36),

\[
 r^{(2)} > \log n / \log 2
\]

so by (5.38),

\[
 (2^{13\nu})^{1/2} - r^{(2)} / 2 \leq (2^{13\nu})^{1/2} n^{-1/2} \leq \theta \lambda q(\tau(n)) / 2 \leq \theta \lambda q(t_j) / 2
\]

and (3.51) follows for \( i = 2 \). Theorem 3.8 can therefore
be applied to give

\[
\sum_{j=0}^{N} \mathbb{P}\left[ \sup_{C_2(t, C, \bar{P}(n))} \left| v_n(C) \right| > \theta \lambda q(t_j) \right]
\]

\[
(5.41) \leq \sum_{j=0}^{N} 4K_1(v) \exp\left(-\frac{\theta^2 \lambda^2 q^2(t_j)}{256 t_j}\right)
\]

\[
+ \sum_{j=0}^{N} 4K_1(v) \exp\left(-\frac{n^{1/2} \theta \lambda q(t_j)}{1024}\right).
\]

We next apply Lemma 5.4, with \( u := \theta \lambda / 1024 \); (5.22) follows from (5.27), (5.37), and (5.39), so using also (5.33) and (5.28), we have

\[
\sum_{j=0}^{N} 4K_1(v) \exp\left(-\frac{\theta^2 \lambda^2 q^2(t_j)}{256 t_j}\right)
\]

\[
(5.42) + \sum_{j=0}^{N} 4K_1(v) \exp\left(-\frac{n^{1/2} \theta \lambda q(t_j)}{1024}\right)
\]

< \eta.

The lemma now follows from (5.40), (5.41), and (5.42).

Proof of Theorem 5.2. We need only combine Theorem 2.9 with Lemmas 5.3 and 5.5, after noting that (5.2) and (5.4) imply (5.3).
VI. Weighted Empirical Processes Without Truncation

For this chapter we consider only the identically distributed case, where all $P(i)$ are the same law $P$. We will look for relations between $C$, $P$, and $q$ which enable a CLT to hold for the processes $\nu_n/q_0P$ without the truncation at $\tau(n)$ used in Chapter 5.

Fix a law $P$ and a VC class $C \subseteq \mathcal{Q}$ and define

$$E_t := \cup\{C \in C : P(C) < t\},$$

$$a(t) := P(E_t), \quad a_*(t) = P_*(E_t), \quad t \geq 0.$$  

If $C$ contains only a few "small" sets, in the sense that $a(t) \to 0$ as $t \to 0$, then the difficulty illustrated in Example 5.1 may occur only with small probability, and a CLT may hold. Such a CLT, as well as a functional LIL, will be corollaries of the following, our main theorem.

**Theorem 6.1.** Suppose the VC class $C$ is size-$\delta$ measurable with constants for $P$ for some $\delta > 0$. Let $q \in Q_1$ with

$$q^2(t)/a(t) \rightarrow \infty \text{ as } t \to 0,$$

(6.1)
Then there exists a sequence \( \{G_j, j \geq 1\} \) of independent copies of \( G_p/q \circ P \), with sample functions a.s. \( d_p \)-continuous, such that

\[
\text{(6.3)} \quad n^{-1/2} \max \sup_{k \leq n} |k^{1/2} v_k(C)/q(P(C)) - \sum_{j=1}^{k} G_j(C)| \to 0 \quad \text{as } n \to \infty
\]

in probability, and in \( L^p \) for all \( p < 2 \). If

\[
\text{(6.4)} \quad \int_{0}^{\infty} a(q^{-1}(x^{-1/2})) dx < \infty,
\]

then the convergence in (6.3) can be obtained in \( L^2 \) also.

If in addition

\[
\text{(6.5)} \quad \int_{0}^{\infty} a(q^{-1}((xLLx)^{-1/2})) dx < \infty,
\]

then the \( G_j \)'s can be chosen so (instead of (6.3))

\[
\text{(6.6)} \quad \sup_{C} |n^{1/2} v_n(C)/q(P(C)) - \sum_{j=1}^{n} G_j(C)| = o((nLLn)^{1/2}) \quad \text{a.s.}
\]

From Lemmas 2.12 and 2.11 the following is then immediate.
Corollary 6.2. Under the hypotheses of Theorem 6.1
(including (6.5)), \( \{v_n/qoP\} \) satisfies the compact and
functional LIL's in \( \ell^\infty(\mathcal{C}) \), and

\[
\lim \sup_n \sup_{C} |v_n(C)/(2LLn)^{1/2}q(P(C)) = \sup_{C} P(C)^{1/2}/q(P(C)) \quad \text{a.s.}
\]

Theorem 6.3. Suppose the VC class \( \mathcal{C} \) is size-\( \delta \) measurable
with constants for \( P \) for some \( \delta > 0 \). Let \( q \in Q_1 \) with

(6.7) \( q^2(t)/a(t) \to \infty \) as \( t \to 0 \),

(6.8) \( \int_0^1 \exp(-\varepsilon q^2(t)/t) \frac{dt}{t} < \infty \) for all \( \varepsilon > 0 \).

Let \( D \) be a subspace of \( \ell^\infty(\mathcal{C}) \) admissible for \( \{v_n/qoP\} \)
and \( P \). Then \( v_n/qoP \) converges in law to \( G_P/qoP \) in
\( (D, \mathcal{B}_b) \) and the latter process has sample paths a.s. in
\( C_{b,u}(\mathcal{C},d_P) \).

Conversely if \( v_n/qoP \) converges in law in \( (D, \mathcal{B}_b) \)
to a Gaussian process with sample paths a.s. in \( C_{b,u}(\mathcal{C},d_P) \),
then

(6.9) \( \int_0^1 \exp(-\varepsilon q^2(t)/t) \frac{dt}{t} < \infty \) for all \( \varepsilon > 0 \)

for some \( \tilde{q} \in Q_1 \) with \( \tilde{q}oP = qoP \).
and if \( q(0) = 0 \) then

\[
(6.10) \quad \frac{q^2(t)}{a_*(t)} \to \infty \text{ as } t \to 0. \quad \square
\]

The first half of Theorem 6.3 is an immediate consequence of Theorem 6.1, Lemma 2.12, and Theorem 2.8. We prove the converse half in the following two propositions.

Let \( A_t \) and \( F_t \) be sets in \( \mathcal{A} \) with \( A_t \subseteq E_t \subseteq F_t \), \( P(A_t) = P_*(E_t) \), and \( P(F_t) = P^*(E_t) \).

Proposition 6.4. If \( \nu_n / q \circ P \) on the VC class \( \mathcal{C} \) converges in law in \( (D, \mathcal{B}_D) \) to a Gaussian process with sample paths a.s. in \( C_{b,u}(\mathcal{C}, d_p) \) for some \( D \subseteq l^\infty(\mathcal{C}) \), and \( q(0) = 0 \), then

\[
(6.11) \quad \frac{q^2(t)}{a_*(t)} \to \infty \text{ as } t \to 0. \]

Proof. We may assume \( a_*(t) > 0 \) for all \( t > 0 \); otherwise
(6.11) is obvious. Suppose first \( a_*(0+) > 0 \). We may assume \( A_1 \supseteq A_1/2 \supseteq A_1/3 \supseteq \ldots \). Let \( A := \bigcap_m A_1/m \). If \( X_i \in A \) for some \( 1 \leq i \leq n \), then \( X_i \in C_m \) for some \( C_m \in \mathcal{C} \) with \( P(C_m) < 1/m \), for each \( m \geq 1 \). Hence

\[
\limsup_m \frac{\nu_n(C_m)}{q(P(C_m))} = \infty. \quad \text{Thus}
\]
102.

\[ P[\sup_{C} |v_{n}(C)| / q(P(C)) = \infty] \geq P[X_{i} \in A \text{ some } i \leq n] \]

for all \( n \), which contradicts the assumption that \( v_{n}/q\circ P \)
is \( \ell^{\infty}(C) \)-valued, since \( P(A) = a_{*}(0+) > 0 \).

Hence we assume \( a_{*}(0+) = 0 \). Suppose (6.11) does not hold. Then there exist \( M > 1/2 \) and \( t_{i} \to 0 \) with \( q^{2}(t_{i}) \leq M a_{*}(t_{i}) \). If \( i \) is sufficiently large (an assumption made tacitly in all inequalities for the remainder of this proof) then there exists \( n_{i} \) with \( (8n_{i}M)^{-1} \leq a_{*}(t_{i}) \leq (4n_{i}M)^{-1} \). If \( P(C) \leq t_{i} \) then

\( (6.12) \ P(C) \leq a_{*}(P(C)) \leq a_{*}(t_{i}) \leq (4n_{i}M)^{-1} \leq (2n_{i})^{-1} \)

and

\( (6.13) \ q^{2}(P(C)) \leq q^{2}(t_{i}) \leq M a_{*}(t_{i}) \leq (4n_{i})^{-1} \).

If \( X_{j} \in A_{t_{i}} \) for some \( j \leq n_{i} \) then there exists \( C \in \mathcal{C} \) with \( P(C) \leq t_{i} \) and \( X_{j} \in C \), so by (6.12) and (6.13),

\[ \frac{v_{n_{i}}(C)}{q(P(C))} \geq \frac{n_{i}^{1/2}(n_{i}^{-1} - (2n_{i})^{-1})}{(4n_{i})^{-1/2}} = 1. \]

It follows that for each \( \theta > 0 \) there are arbitrarily large \( n_{i} \) for which
103.

$$\mathbb{P}_* \left[ \sup \{ |\nu_{n_i} (C)|/q(P(C)) : C \in \mathbb{C}, P(C) < \theta \} \geq 1 \right]$$

$$\geq \mathbb{P}[X_j \in A_{t_i} \text{ for some } j \leq n_i]$$

(6.14) $$= 1 - (1 - P(A_{t_i}))^{n_i}$$

$$\geq 1 - (1 - (8n_i M)^{-1})^{n_i}$$

$$\geq (1 - \exp(-1/8M))/2.$$ 

From its covariance we see that the limiting Gaussian process must be $G_{p}/q\cdot P$. Because this process is $l^\infty(\mathbb{C})$-valued, $\text{var}(G_{p}(C)/q(P(C))) = P(C)(1 - P(C))/q^2(P(C))$ must be bounded in $C$, so

(6.15) $P(C)/q(P(C)) \rightarrow 0$ as $P(C)(C \in \mathbb{C}) \rightarrow 0.$

Fix $\theta > 0$ and $n > 1$. Since $a_*(t) > 0$ for all $t > 0$, it follows from (6.15) that we can find $C \in \mathbb{C}$ such that

$$P(C) < \theta, n^{1/2}P(C)/q(P(C)) \leq 1/2,$$

and

$$\mathbb{P}[X_j \in C \text{ for some } j \leq n] \leq (1 - \exp(-1/8M))/4.$$ 

Thus for each $\theta > 0$ and $n > 1$, 

104.

\[ P^* \{ \inf \{ |v_n(C)| / q(P(C)) : C \in \mathcal{C}, P(C) < \theta \} \geq 1/2 \} \]

(6.16)
\[ \leq (1 - \exp(-1/8M))/4. \]

From (6.14) and (6.16) we see that (2.5) cannot hold for \( \psi_n = v_n/q_0 P \), so the convergence in law cannot occur. \( \square \)

Proposition 6.5. If \( v_n/q_0 P \) on the VC class \( \mathcal{C} \) converges in law in \( (D, \mathbb{B}_D) \) to a Gaussian process with sample paths a.s. in \( C_{b,u}(\mathbb{C}, d_p) \) for some subspace \( D \) of \( l^\infty(\mathbb{C}) \), then

\[ \int_0^1 \exp(-\varepsilon q^2(t)/t) \frac{dt}{t} < \infty \] for all \( \varepsilon > 0 \) for some \( \tilde{q} \in Q_1 \) with \( \tilde{q} \circ P = q_0 P \).

Proof. If \( q(0) > 0 \) or \( a_*(t) = 0 \) for some \( t > 0 \), then (6.17) is trivial, so we assume \( q(0) = 0 \) and \( a_*(t) > 0 \) for \( t > 0 \). By Proposition 6.4 we have \( a_*(0+) = 0 \), so \( \mathcal{C} \) contains sets of arbitrarily small positive probability.

Suppose (6.17) does not hold. We will show that there exists a sequence \( \{C_i\} \) in \( \mathcal{C} \) and a constant \( u > 0 \) such that

\[ P(C_{i+1}) \leq P(C_i)/3 \]
and

\[(6.19) \sum_{i=1}^{\infty} \exp(-uq^2(P(C_i)/P(C_i))) = \infty.\]

Fix \(\delta > 0\) such that \(q(t)\) and \(q(t)/t\) are monotone on \([0, \delta]\), and let \(T := \{P(C) : C \in \mathcal{C}\}\). Then we can write

\[(0,1) \setminus T = \bigcup_{0 < n < N} (r_n, s_n),\]

a union of disjoint open intervals, for some \(0 < N < \infty\).

Let \(I := \{n \geq 0 : r_n \leq s_n/3\}\), \(M = \text{card}(I)\), and \(S := \bigcup_{n \in I} (r_n, s_n)\). Then we can write \(S = \bigcup_{0 < m < M} (a_m, b_m)\), a union of disjoint intervals, with \(b_0 > b_1 > \ldots\). Let

\[s := \begin{cases} 
\inf(S \cap [0, \delta]) & \text{if } S \cap [0, \delta] \neq \emptyset \\
\delta & \text{if } S \cap [0, \delta] = \emptyset.
\end{cases}\]

Since \(a_*(0+) = 0\) but \(a_*(t) > 0\) for \(t > 0\), no \(a_m\) can be 0. Thus either \(M = \infty\) or \(s > 0\).

Case 1: \(M = \infty\). Then \(b_m \to 0\), so \(b_m < \delta\) for \(m \geq \text{some } m_0\). Define

\[c_m := q(b_m)a_m/q(a_m), \ m \geq m_0.\]
Then since $b_m \leq \delta_1$, we have $a_m \leq c_m \leq b_m$. Define

$$
\tilde{q}(t) = \begin{cases} 
q(t) & \text{if } t \not\in S \\
q(b_m) & \text{if } t \in [c_m, b_m) \\
q(a_m)t/a_m & \text{if } t \in [a_m, c_m].
\end{cases}
$$

Then $\tilde{q} \in Q_1$, so there exists $\varepsilon > 0$ such that

$$
\lim_{t \to 0} \frac{\exp(-\varepsilon q^2(t)/t)}{t} = 0.
$$

Thus we get three subcases.

Case 1a. The first sum on the right side of (6.20) is infinite. Since $b_m \in T$, we have $a_*(b_m+) \geq b_m$. By Proposition 6.4 we have $q^2(t)/a_*(t) \to \infty$ as $t \to 0$ if $q(0) = 0$. It follows that $q^2(b_m)/b_m \to \infty$ as $m \to \infty$, so $q^2(b_m)/b_m \geq \varepsilon^{-1}\log 2$ for all $m$ greater than or equal to some $m_1 \geq m_0$. We have then
\[
\sum_{m=m_1}^{\infty} \int_{0}^{b_m} \exp(-\varepsilon q^2(b_m)/t) \frac{dt}{t} = \sum_{m=m_1}^{\infty} \left( \sum_{k=0}^{2-k-b_m} \exp(-\varepsilon q^2(b_m)/t) \frac{dt}{t} \right) 
\]

\[
(6.21) \leq \sum_{m=m_1}^{\infty} \left( \sum_{k=0}^{\infty} (2-k-b_m)(2-k-b_m) \exp(-2k \varepsilon q^2(b_m)/b_m) \right) 
\]

\[
\leq \sum_{m=m_1}^{\infty} \left( \sum_{k=0}^{\infty} \exp(-(k+1) \varepsilon q^2(b_m)/b_m) \right) 
\]

\[
\leq \sum_{m=m_1}^{\infty} 2 \exp(-\varepsilon q^2(b_m)/b_m). 
\]

Now for each \( m \geq 1 \) we can find a set \( C_m \in \mathbb{C} \) with \( b_m \leq P(C_m) \leq b_m - 1/3 \) and \( q^2(P(C_m))/P(C_m) \geq q^2(b_m)/2b_m \).

(6.18) and (6.19), with \( u = \varepsilon/2 \), then follow from (6.21).

Case lb. The second sum on the right side of (6.20) is infinite. Similarly to Case la, we have \( q^2(a_m)/a_m \to \infty \) as \( m \to \infty \), so \( a_m/\varepsilon q^2(a_m) < 1 \) for \( m \) greater than or equal to some \( m_2 \geq m_0 \). Hence

\[
\sum_{m=m_2}^{\infty} \int_{0}^{a_m} \exp(-\varepsilon q^2(a_m)/a_m) \frac{dt}{t} \frac{dt}{t} 
\]

\[
(6.22) \leq \sum_{m=m_2}^{\infty} \left( \sum_{x=a_m}^{\infty} \exp(-\varepsilon q^2(a_m)/a_m) dx \right) 
\]

\[
\leq \sum_{m=m_2}^{\infty} \exp(-\varepsilon q^2(a_m)/a_m). 
\]
Now for each \( m \geq 0 \) we can find \( c_m \in \mathbb{C} \) with
\[
3a_{m+1} \leq P(c_m) \leq a_m \quad \text{and} \quad q^2(P(c_m)) \geq q^2(a_m)/2a_m.
\]
(6.18)
and (6.19), with \( u = \varepsilon/2 \), then follow from (6.22).

Case lc. The third sum on the right side of (6.20) is infinite. For \( m \geq m_0 \) let
\[
W_m := \begin{cases} 
[b_{m+1}, a_m) \cap \{3^{2^l}b_{m+1} \leq 0 \} & \text{if } b_{m+1} < a_m \\
\{b_{m+1}\} & \text{if } b_{m+1} = a_m.
\end{cases}
\]

Let \( \{t_k, k \geq 0\} \) be a sequence consisting of the set
\[
\bigcup_{m \geq m_0} W_m
\]
in decreasing order. Then
\[
\begin{align*}
\sum_{m=m_0}^{\infty} t_k \in [b_{m+1}, a_m] 
& \int_0^{9t_k} \exp(-\varepsilon q^2(t)/t) \frac{dt}{t} \\
& \leq \sum_{k=0}^{\infty} 8 \exp(-\varepsilon q^2(t_k)/9t_k).
\end{align*}
\]
(6.23)

Now \( \inf ([t_k, \infty) \cap T) < 3t_k, \ k \geq 1 \), so we can find
\( C_k \in \mathbb{C} \) with \( t_k \leq P(C_k) < 3t_k \) and \( P(C_k) \) in the same
interval \( [b_{m+1}, a_m] \) as \( t_k \). Then \( P(C_{k+1}) \leq P(C_k)/3 \)
and \( q^2(P(C_k))/P(C_k) \leq 3q^2(t_k)/t_k \) for all \( k \), so (6.18)
holds, and (6.19), with \( u = \varepsilon/3 \), follows from (6.23).

Case 2. \( s > 0 \). Let \( v_m := s/9^m, m \geq 0 \). For some
\( \varepsilon > 0 \) we have
\[ \int_0^\infty \exp(-\varepsilon q^2(t)/t) \frac{dt}{t} = \sum_{m=0}^{\infty} \int_0^{v_m} \exp(-\varepsilon q^2(t)/t) \frac{dt}{t} \]
\[ \leq \sum_{m=0}^{\infty} (v_m - v_{m+1}) v_m^{-1} \exp(-\varepsilon q^2(v_m)/9v_m) \]
\[ = 8 \sum_{m=0}^{\infty} \exp(-\varepsilon q^2(v_m)/9v_m). \]

Now as in case lc we have \( \inf([v_m, \infty) \cap T) < 3v_m \), so we can find \( C_m \in \mathbb{C} \) with \( v_m \leq P(C_m) < 3v_m = v_{m-1}/3, \ m \geq 1. \)

Then (6.18) and (6.19), with \( u = \varepsilon/3 \), follow as in case lc.

Thus we have \( \{C_m\} \) and \( u \) satisfying (6.18) and (6.19) in all cases. Define

\[ D_m := C_m \setminus \bigcup_{j>m} C_j, \ m \geq 1, \]

so the \( D_m \) are disjoint. Since \( P(C_j) \leq 3^{-(j-m)} P(C_m) \) for \( j > m \), we have

\[ (6.24) \ P(D_m) \geq P(C_m) - \sum_{j>m} P(C_j) \geq P(C_m)/4. \]

Let \( W \) be a mean-0 Gaussian process indexed by \( \Omega \), with covariance \( \mathbb{E}W(A)W(B) := P(A \cap B) \). Then
(6.25) \(\mathcal{L}(C_p(C)) = \mathcal{L}(W(C) - P(C)W(X)), C \subseteq C\).

Let \(p_m := P[|W(D_m)| / q(P(C_m)) > u^{1/2}/2], m \geq 1\). Now
\(a^*(P(C_m)^+) > P(C_m)\), so by Proposition 6.4,

(6.26) \(q(P(C_m))/P(C_m)^{1/2} \to \infty\) as \(m \to \infty\).

Using the fact that \(x^{-1} \geq \exp(-x^2/16\pi)\) for large \(x\), we have for large \(m\):

\[
P_m = P[|W(D_m)| / P(D_m)^{1/2} > u^{1/2} q(P(C_m)) P(C_m)^{1/2} / P(D_m)^{1/2}] \geq P[|W(D_m)| / P(D_m)^{1/2} > u^{1/2} q(P(C_m)) / P(C_m)^{1/2}]
\]

\[
\geq ((2\pi)^{1/2} 2u^{1/2} q(P(C_m)) P(C_m)^{-1/2})^{-1} \exp(-u q^2(P(C_m))/2P(C_m)) \geq \exp(-u q^2(P(C_m))/P(C_m)).
\]

Hence by (6.19), \(\sum_{m=1}^{\infty} p_m = \infty\). Since \(W(D_m), m \geq 1\), are independent, it follows that

(6.27) \(P[|W(D_m)| / q(P(C_m)) > u^{1/2}/2 \text{ i.o.}] = 1\).

We can therefore define \(\tau(j), j \geq 1\), to be the
$j^\text{th}$ index $m$ such that $|W(D_m)/q(P(D_m)) > u^{1/2}/2$. Then $\tau(j)$ is an r.v., and $[\tau(j) = k]$ is an event in the $\sigma$-algebra generated by $W(D_1), \ldots, W(D_k)$, for all $k \geq 1$ so $1_{[\tau(j) = k]}$ is independent of $W(C_k \setminus D_k)$, since $C_k \setminus D_k$ is disjoint from $D_1, \ldots, D_k$ and $W$ is independent on disjoint sets. Fix $\lambda > 0$. Now

$$W(C_m) = W(D_m) + W(C_m \setminus D_m) \text{ a.s. for all } m \geq 1,$$

so by (6.26) and Chebyshev's inequality, for $k \geq$ some $k_0$ we have

$$P[|W(C_\tau(j))|/q(P(C_\tau(j))) > u^{1/2}/4|\tau(j) = k]$$

$$\geq P[|W(C_k \setminus D_k)|/q(P(C_k)) \leq u^{1/2}/4|\tau(j) = k]$$

$$\geq P[|W(C_k \setminus D_k)| \leq \lambda P(C_k \setminus D_k)^{1/2}]$$

$$\geq 1 - \lambda^{-2}.$$  

Since $\tau(j) \geq j$, and $k \geq k_0$ is arbitrary in (6.28), it follows that

$$P[|W(C_\tau(j))|/q(P(C_\tau(j))) > u^{1/2}/4] \geq 1 - \lambda^{-2}, j \geq k_0$$
so since $\lambda$ is arbitrary,

$$\mathbb{P}[|W(C_m)/q(P(C_m)) > u^{1/2}/4 \ i.o.] = 1.$$  

From (6.25) and (6.26) it then follows that

$$(6.29) \quad \mathbb{P}[|G_p(C_m)/J(P(C_m)) > u^{1/2}/8 \ i.o.] = 1.$$  

However, for each $n \geq 1$ we have

$$\sum_{m=1}^{\infty} \mathbb{P}[P_n(C_m) \neq 0] = \sum_{m=1}^{\infty} 1 - (1 - P(C_m))^n$$

$$\leq \sum_{m=1}^{\infty} 1 - (1 - 3^{-m+1})^n$$

$$\leq \sum_{m=1}^{\infty} 3^{-m+1}n$$

$$< \infty$$

so $P_n(C_m) \to 0$ as $m \to \infty$ a.s. When combined with (6.26) this shows

$$|v_n(C_m)/q(P(C_m)) \to 0 \ as \ m \to \infty \ a.s.$$  

This and (6.29) show the convergence in law cannot occur. $\blacksquare$

The next three lemmas will be used in the proof of Theorem 6.1. The first is derived from a proof in O'Reilly (1974).
Lemma 6.6. If \( q \in Q_1 \) and

\[
\int_0^1 \exp(-\varepsilon q^2(t)/t) \frac{dt}{t} < \infty \quad \text{for all } \varepsilon > 0,
\]

then \( q^2(t)/t \to \infty \) as \( t \to 0 \).

Proof. Suppose \( q^2(t)/t \neq \infty \). Choose \( \delta > 0 \) such that \( q \) is monotone on \([0, \delta]\). Then there exist \( M > 0 \) and \( t_i, i \geq 1 \), with \( t_{i+1} \leq t_i/2 \), \( t_1 \leq \delta \), and \( q^2(t_i) \leq Mt_i \).

Hence for \( t_i/2 \leq t \leq t_i \) we have \( q^2(t)/t \leq Mt_i/t \leq 2M \), so

\[
\int_0^1 \exp(-q^2(t)/t) \frac{dt}{t} \geq \sum_{i=1}^{\infty} \int_{t_i/2}^{t_i} \exp(-2M) \frac{dt}{t} = \infty.
\]

Lemma 6.7. Suppose \( C \subset A \) and \( q \in Q_1 \). Let \( g \) and \( h \) be nonnegative functions on \([0, \infty)\), with \( h \) strictly increasing, \( h(0) = 0 \), and

\[
q(x) \sim h^{-1}(x) \quad \text{as } x \to \infty.
\]

Then

\[
\int_0^\infty a(q^{-1}(1/h(x))) dx < \infty
\]

if and only if there exists a measurable r.v. \( F \) on
(X, \mathcal{A}, P) \text{ such that both }

(6.33) \quad 1_C/q(P(C)) \leq F \text{ for all } C \in \mathcal{C}

and

(6.34) \quad \int g(F) \, dP < \infty.

Proof. By (6.31), (6.34) is equivalent to \( \int h^{-1}(F) \, dP < \infty \), so we may assume \( g = h^{-1} \).

Suppose first (6.32) holds. Let \( T \) be the set of points of discontinuity of \( a(t) \), \( S \) the set of rational numbers, and \( U := S \cup T \). For \( t \geq 0 \) let

\[ H_t := \cap \{ F_s : s > t, s \in U \}. \]

Then \( H_t \in \mathcal{A} \) since \( U \) is countable, and \( H_s \subseteq H_t \) for \( s < t \). Recalling that \( E_t \subseteq F_t \) and \( P^*(E_t) = P(F_t) \), we see that \( E_t \subseteq H_t \) and \( P^*(E_t) = P(H_t) = a(t) \) for all \( t \).

Define

\[ F(x) := \frac{1}{\inf\{q(t) : x \in H_t \}}, \quad x \in X. \]

Then for \( u \geq 0 \),

(6.35) \quad F(x) > u \text{ if and only if } x \in H_t \text{ and } q(t) < \frac{1}{u} \text{ for some } t > 0.
Since $q$ is continuous and the sets $H_t$ increase with $t$, we may take the $t$ in (6.35) to be rational; it follows that $F$ is measurable.

If $u > 0$ and $l_C(x)/q(P(C)) > u$, then $q(P(C)) < 1/u$ and $x \in C$, so $q(t) < 1/u$ for some $t$ with $x \in H_t$, so $F(x) > u$. From this (6.33) follows. Using (6.35) we have

$$
\int g(F) dP = \int_{0}^{\infty} P[h^{-1}(F) > u] du
$$

$$
\leq \int_{0}^{\infty} P[F > h(u)] du
$$

$$
\leq \int_{0}^{\infty} P(H^{-1}_{q^{-1}(1/h(u))}) du
$$

$$
= \int_{0}^{\infty} a(q^{-1}(1/h(u)))
$$

$$
< \infty.
$$

Conversely assume (6.33) and (6.34) hold. If $q(0) > 0$, then $a(q^{-1}(1/h(x))) = a(0) = 0$ for large $x$, and (6.32) holds; hence we assume $q(0) = 0$. Then there exists $K > 0$ such that for $u \geq K$, $t < q^{-1}(1/h(u))$ implies $q(t) \leq 1/h(u)$. If $u \geq K$ and $x \in \mathbb{E}^{-1}_{q^{-1}(1/h(u))}$, then $x \in C$ for some $C \in \mathcal{C}$ with $q(P(C)) \leq 1/h(u)$, so $F(x) \geq l_C(x)/q(P(C)) > h(u)$. Hence
Lemma 6.8. Suppose the VC class $\mathcal{C}$ is size-$\delta$ measurable with constants for $\mathbb{P}$ for some $\delta > 0$, and let $q \in Q_1$ with

(6.36) $q^2(t)/a(t) \to \infty$ as $t \to 0,$

(6.37) $\int_0^1 \exp(-\varepsilon q^2(t)/t) \frac{dt}{t} < \infty$ for all $\varepsilon > 0.$

Let $\theta > 0$ and $0 < \eta < 1;$ then there exist $\delta > 0$ and $n_1$ such that for $n \geq n_1,$

(6.38) $\mathbb{P}^*\left[\sup\{|v_n(C)|/q(P(C)) : C \in \mathcal{C}, P(C) < \delta\} > \theta\right] < \eta.$

Proof. If $q(0) > 0$ then (6.38) is an easy consequence of Theorem 3.8. Hence we assume $q(0) = 0.$ By (6.36), $a(0+) = 0,$ and clearly we may assume $a(t) > 0$ for all $t > 0.$

Let $v := V(\mathcal{C}),$ fix $0 < \lambda < 1/3,$ and let $M > 0$ be large enough so
\[\begin{align*}
(6.39) \quad & 2048 (\theta \lambda n^{1/2})^{-1} (\lambda^{-1-1})^{-1} \log 2 \leq M, \\
(6.40) \quad & (16K_1(\nu) + 8) \exp(-\theta \lambda n^{1/2}M/2048) < \eta/4.
\end{align*}\]

Let \( r \) be the least integer such that \((2^{13} \nu)^{1/2} 2^{-r/2} \leq 1/2.\)

Let \( \delta > 0, \) and \( t_j := q^{-1}(q(\delta) \lambda^j), \ j \geq 0, \) where \( q^{-1} \) is as defined above Lemma 5.4. By (6.36), (6.37), and Lemmas 5.4 and 6.6, we can take \( \delta \) to be small enough so \( \sim \) is size-\( \delta \) measurable with constants for \( P \) and

\[\begin{align*}
(6.41) \quad & q \text{ is increasing on } [0, \delta] \text{ and } t_j \in [0, \delta] \text{ for all } j, \\
(6.42) \quad & \theta^2 \lambda^2 q^2(t)/16a(t+) \geq 1 \text{ for all } t \leq \delta, \\
(6.43) \quad & \theta^2 \lambda^2 q^2(t)/8t \geq 36 \log 2 \text{ for all } t \leq \delta, \\
(6.44) \quad & \frac{\theta^2 \lambda^2 q^2(t)}{4096tv \log 2} \geq r \text{ for all } t \leq \delta, \\
(6.45) \quad & \sum_{j=0}^{\infty} \exp(-\theta^2 \lambda^2 q^2(t_j)/2048t_j) < \eta/(16K_1(\nu) + 8).
\end{align*}\]

Let \( a_n := \sup \{t : a(t) \leq \eta/2n\}, n \geq 1. \) By (6.36) we have \( n^{1/2}q(a_n) \to \infty \text{ as } n \to \infty, \) so we can find \( n_1 \) such that for \( n \geq n_1, \)

\[\begin{align*}
(6.46) \quad & \theta \lambda n^{1/2} q(a_n)/2 \geq 512 \log 2,
\end{align*}\]
\[
(6.47) \quad \frac{\theta \lambda n^{1/2} q(a_n)}{4096 \sqrt{\log 2}} \geq r,
\]

\[
(6.48) \quad n^{1/2} q(a_n) \geq M,
\]

\[
(6.49) \quad \theta \lambda n^{1/2} q(a_n)/6 \geq \eta^{1/2},
\]

\[
(6.50) \quad a_n < \delta.
\]

Fix \( n \geq n_1 \). By (6.50) and (6.41) there is an integer \( N \geq 0 \) such that

\[
(6.51) \quad t_j \geq a_n \quad \text{if and only if} \quad j \leq N.
\]

If \( C \in \mathcal{C} \), \( P(C) < t_{N+1} \), and \( P_n(C) = 0 \), then by (6.43),

\[
|v_n(C)|/q(P(C)) = n^{1/2} P(C)/q(P(C))
\]

\[
\leq (n \alpha(t_{N+1}))^{1/2} P(C)^{1/2}/q(P(C))
\]

\[
\leq (n/2)^{1/2} (\lambda^2 \theta^2 / 288 \log 2)^{1/2}
\]

\[
\leq \theta.
\]

It follows that
\begin{align*}
\mathbb{P}^*\left[\sup\{\nu_n(C)/q(P(C)) : C \in \mathcal{C}, P(C) < t_{N+1}\} > \theta \right] \\
\leq \mathbb{P}^*\left[P_n(C) > 0 \text{ for some } C \in \mathcal{C} \text{ with } P(C) < t_{N+1}\right] \\
(6.52) \\
\leq na(t_{N+1}) \\
< n/2.
\end{align*}

Define
\begin{align*}
g(t) := \max(a(t)/t,1), \\
h(t) := \lambda q(t) \min(g(t)^{1/2},(na(t+))^{1/2})
\end{align*}
for \( t > 0. \) Then
\begin{align*}
\mathbb{P}^*\left[\sup\{\nu_n(C)/q(P(C)) : C \in \mathcal{C}, t_{N+1} \leq P(C) < \delta\} > \theta \right] \\
\leq \sum_{j=0}^{N} \mathbb{P}\left[\sup\{\nu_n(C) : C \in \mathcal{C}, t_{j+1} \leq P(C) < t_j\} > \theta q(t_{j+1})\right] \\
(6.53) \\
\leq \sum_{j=0}^{N} \mathbb{P}\left[\nu_n(F_{t_j}) > \theta h(t_j)/2\right] \\
+ \sum_{j=0}^{N} \mathbb{P}\left[\sup_{C \in \mathcal{C}_2(t_j,\mathcal{C},P)} |\nu_n(C)| > \theta \lambda q(t_j) |\nu_n(F_{t_j})| < \theta h(t_j)/2\right].
\end{align*}

Fix \( 0 \leq j \leq N, \) and let \( k_1, k_2 \) be nonnegative integers such that
\begin{equation}
(6.54) \quad |\nu_n(F_{t_j})| \leq \theta h(t_j)/2 \text{ if and only if } k_1 \leq nP_n(F_{t_j}) \leq k_2.
\end{equation}
120.

Define the measure $H(\cdot) := P(\cdot \mid F^t_j)$ on $\mathcal{C}$, and let $\mathbb{H} := H^\infty$. For $k \geq 1$ let $L_k$ be the (random) set of the $k$ least indices $i \geq 1$ such that $X_i \in F^t_j$, and define the following functions on $(X^\infty, \mathcal{C}^\infty)$:

$$H_k := \frac{1}{k} \sum_{i \in L_k} \delta_{X_i},$$

$$H'_k := \frac{1}{k} \sum_{i=1}^{k} \delta_{X_i},$$

$$H_{k-1} := \frac{1}{k} \sum_{i=k+1}^{2k} \delta_{X_i},$$

$$\mu_k := k^{1/2} (H_k - H),$$

$$\mu'_k := k^{1/2} (H_k - H)',$$

$$\mu^0_k := k^{1/2} (H_k - H_k').$$

($H_k$ will be well defined a.s. in our context.) Then on the set $[nP_n(F^t_j) = k]$ we have

(6.55) $C \in \mathcal{C}$, $P(C) < t_j$ implies $kH_k(C) = nP_n(C)$.

Also

(6.56) $H_k = \tilde{H}_k \mathbb{H}$-a.s.
121.

Fix \( k_1 \leq k \leq k_2 \); if \( C \in \mathcal{C} \) and \( P(C) < t_j \) then by (6.54),

\[
|kH(C) - nP(C)| = \frac{nP(C)}{P(F_{t_j})} |k - P(F_{t_j})| \leq n^{1/2} |v_n(F_{t_j})| / g(t_j)
\]

(6.57)

\[
\leq n^{1/2} \theta h(t_j) / 2g(t_j)^{1/2}
\]

\[
\leq n^{1/2} \theta \lambda q(t_j) / 2.
\]

By assumption, \( C \) is size-\( \delta \) measurable with constants for \( P \). Let \( || \cdot || \) denote the sup norm for functions on \( \mathcal{C}_2(t_j, C, P) \); then

\[
||k^{1/2} (P_k - \frac{P}{P(F_{t_j})})||_j
\]

\[
||v_k^0||_j = ||v_k^0||_j
\]

\[
||H_{2k} - H||_j = ||P_{2k} - \frac{P}{P(F_{t_j})}||_j
\]

so since \( H^{2n} \ll P^{2n} \) it follows that \( \mathcal{C}_2(t_j, C, P) \) is \( k \)-deviation-measurable for \( H \). Hence by (6.55), (6.56), and (6.57),
We wish to apply Theorem 3.8 to the right side of (6.58), with \( D := C_2(t_j, C, P) \), \( u := \sqrt[2]{\theta n/k} q(t_j)/2 \), and \( \beta := 1/g(t_j) \). We may assume \( k \neq 0 \). Let

\[ J := \{ j : 0 \leq j \leq N, \; \theta h(t_j) < 2n^{1/2}a(t_j) \} \]

**Case 1:** \( j \in J \). Then by (6.54),

\[
\frac{n}{k} \geq \left( a(t_j) + \theta h(t_j)/2n^{1/2} \right)^{-1} \geq \left( 2a(t_j) \right)^{-1} \quad \text{so by (6.42)}
\]

\[
u^2 \geq \frac{\theta^2 \lambda^2 q^2(t_j)}{8a(t_j)} \geq 1,
\]

while by (6.43),

\[
u^2/\beta \geq \frac{\theta^2 \lambda^2 q^2(t_j)}{8t_j} \geq 36 \log 2,
\]

which establishes (3.52); by (6.46),
(6.61) \( k^{1/2} u \geq \theta \lambda n^{1/2} q(t_j)/2 \geq 512 \log 2, \)

which establishes (3.53) with \( k \) in place of \( n \). Let \( r^{(1)} \) and \( r^{(2)} \) be as in (3.50), but with \( n \) replaced by \( k \); by (6.60), the left half of (6.61), (6.44), and (6.47) we have \( r^{(1)} \geq r \) and \( r^{(2)} \geq r \), so (3.51) follows from (6.59). Thus we may apply Theorem 3.8 and the left halves of (6.60) and (6.61) to get

\[
\mathbb{H}[|\|\mu_k\|_j > \theta \lambda (n/k)^{1/2} q(t_j)/2]
\]

(6.62) \[
\leq 4K_1(v) \exp(-\theta^2 \lambda^2 q^2(t_j)/2048 t_j)
\]

\[
+ 4K_1(v) \exp(-\theta \lambda n^{1/2} q(t_k)/2048) \text{ if } j \in J.
\]

Case 2: \( j \notin J \). Then by (6.54),

\[
n/k \geq (a(t_j) + \theta h(t_j)/2n^{1/2})^{-1}
\]

(6.63) \[
\geq n^{1/2}/\theta h(t_j)
\]

\[
\geq (\theta \lambda q(t_j) a(t_j +)^{1/2})^{-1}
\]

so by (6.42),
(6.64) \[ u^2 > \frac{\theta \lambda q(t_j)}{4a(t_j+)^{1/2}} \geq 1, \]

while by the second inequality in (6.63) and (6.46),

\[ \frac{u^2}{\beta} > \frac{\theta \lambda^2 n^{1/2} q^2(t_j) g(t_j) / 4h(t_j)}{e^{n^{1/2} q(t_j)/4}} > 36 \log 2. \]

Also (6.61) is valid in this case also. Let \( r^{(1)} \) and \( r^{(2)} \) again be as in (3.50), but with \( n \) replaced by \( k \);

by the second inequality in (6.65), the first in (6.61),

and (6.47) we have \( r^{(1)} > r \) and \( r^{(2)} > r \), so (3.51) follows from (6.64). Thus we may apply Theorem 3.8, the second inequality in (6.65), and the first in (6.61) to get

\[ H[|u_k||_j > \theta \lambda (n/k)^{1/2} q(t_j) / 2] \]

\[ \leq 8k_1(v) \exp(-\theta \lambda n^{1/2} q(t_j) / 2048) \text{ if } j \not\in J. \]

We turn next (for both cases) to the \( j^{th} \) term of the first sum on the right side of (6.53). Define
125.

\[ w(s) := \frac{s}{2(1 + s/3)}, \quad s \geq 0, \]

\[ \gamma := \theta h(t_j)/2n^{1/2}a(t_j)(1 - a(t_j)), \]

\[ \tau := \theta n^{1/2}h(t_j)/2. \]

Then \( w(s) \geq \min(s/6, 1) \), and

\[ \tau = \theta \lambda n^{1/2}q(t_j)\min(g(t_j)^{1/2}/2, (na(t_j)+)^{1/2}/2) \]

\[ \geq \theta \lambda n^{1/2}n^{1/2}q(t_j)/4, \]

so by Bernstein's inequality (see Hoeffding (1963)) and (6.49),

\[ \mathbb{P}[|\sqrt{n} \{F_n(t_j)\} - \theta h(t_j)/2| > \theta h(t_j)/2] \leq 2 \exp(-\tau w(\gamma)) \]

\[ \leq 2 \exp(-\tau) + 2 \exp(-\tau \gamma/6) \]

\[ \leq 2 \exp(-\theta \lambda n^{1/2}n^{1/2}q(t_j)/4) \]

\[ + 2 \exp(-\theta^2 h^2(t_j)/24a(t_j)) \]

(6.67)

\[ \leq 2 \exp(-\theta \lambda n^{1/2}n^{1/2}q(t_j)/4) \]

\[ + 2 \exp(-\theta^2 \lambda^2 q^2(t_j)/24t_j) \]

\[ + 2 \exp(-\theta^2 \lambda^2 nq^2(t_j)/24) \]

\[ \leq 4 \exp(-\theta \lambda n^{1/2}n^{1/2}q(t_j)/4) \]

\[ + 2 \exp(-\theta^2 \lambda^2 q^2(t_j)/24t_j). \]
Since $j \leq N$ and $k_1 \leq k \leq k_2$ are arbitrary, from (6.53), (6.54), (6.58), (6.62), (6.66), and (6.67) we get

$$\mathbb{P}^* \left[ \sup \{|\nu_n(C)|/q(P(C)) : C \in \mathcal{C}, \quad t_{N+1} \leq P(C) < \delta \} > \theta \right]$$

\[
\leq \sum_{j=0}^{N} 4 \exp(-\theta \lambda \eta^{1/2} n^{1/2} q(t_j)/4) \\
+ \sum_{j=0}^{N} 2 \exp(-\theta^2 \lambda^2 q(t_j)/24 t_j) \\
+ \sum_{j=0}^{N} 4K_1(v) \exp(-\theta^2 \lambda^2 q(t_j)/2048 t_j) \\
+ \sum_{j=0}^{N} 8K_1(v) \exp(-\theta \lambda n^{1/2} q(t_j)/2048) \\
\leq (8K_1(v) + 4) \sum_{j=0}^{N} \exp(-\theta \lambda n^{1/2} q(t_j)/4) \\
+ (4K_1(v) + 2) \sum_{j=0}^{N} \exp(-\theta^2 \lambda^2 q(t_j)/2048 t_j).
\]

Since (5.22) with $u := \theta \lambda n^{1/2}/2048$ follows from (6.39), (6.41), (6.51), and (6.48), we can apply Lemma 5.4, (6.45), and (6.40) to the right side of (6.68) to bound it by $\eta/2$. In combination with (6.52), this proves the lemma. \hfill \blacksquare

**Proof of Theorem 6.1.** We use Theorem 2.13. Define a map $\psi : \mathcal{C} \rightarrow L^2(X, \mathcal{A}, \mathbb{P})$ by $\psi(C) = 1_C$, and a map $\phi$ from the
closure $\overline{\psi(C)}$ to $L^2(X, \mathcal{A}, P)$ by $\phi(1_C) = 1_C/P(P(C))$.

Let $\tilde{T} := \phi(\psi(C))$, and $f_C := \phi(\psi(C))$, $C \in C$. Since $||\phi(1_C)||_2 = P(C)^{1/2}/q(P(C))$, by Lemma 6.6 $\phi$ is continuous. But $\psi(C)$ is totally bounded, so $\overline{\psi(C)}$ is compact, so $\phi$ is a homeomorphism. Hence $\tilde{T}$ is totally bounded, and given $\gamma > 0$ there exists $\delta > 0$ such that

(6.69) $f_A, f_B \in \tilde{T}$, $\int (f_A - f_B)^2 dP < \delta^2$ imply $P(A \Delta B) < \gamma$.

Let $\tau(n)$ be 0 for all $n$; then (5.8) and (5.9) follow from Lemmas 6.6 and 6.8, so by Lemma 5.3 we have (2.5) holding for $\psi_n = \nu_n/q^oP$. Hence using (6.69), given $\epsilon > 0$ we can find $\gamma, \delta > 0$ and $n_0$ such that for $n \geq n_0$,

$$P*\left[ \sup \{ |f - g| \nu_n |: f, g \in \tilde{T}, \int (f - g)^2 dP < \delta^2 \} > \epsilon \right]$$

(6.70) $\leq P*\left[ \nu_n/q^oP \in B_{\gamma, \epsilon}^o(P) \right]$

$< \epsilon$,

i.e. (2.13) holds. Thus we get $Y_j, j \geq 1$, satisfying (2.14) - (2.16); let $G_j := Y_j \circ \phi \circ \psi$. Then $\mathcal{L}(G_j) = \mathcal{L}(G_{1}/q^oP)$ and $G_j$ has a.s. $d_P$-continuous sample paths; the remainder of the theorem follows from Theorem 2.13 and Lemma 6.7,
using the latter first with 
\( h(x) = x^{1/2}, g(x) = x^2, \) then
with 
\( h(x) = (x \ll x)^{1/2}, g(x) = x^2/\ll x. \)

Following Dudley (1978), let 
\( D_0(\mathcal{C}, d_p) \) denote the
space all functions in \( \ell^\infty(\mathcal{C}) \) which are the sum of a
d\( p \)-continuous function and a finite linear combination
\[ \sum_{i=1}^m b_i \delta_{x_i} \] of point masses, where \( x_i \in X. \) For \( q \in C[0,1] \)
let 
\[ D_0(\mathcal{C}, p, q) := \{ f/q \circ P : f \in D_0(\mathcal{C}, d_p), f/q \circ P \in \ell^\infty(\mathcal{C}) \}. \]

Example 6.9. Let \( d \geq 1, X := [0,1]^d, \mathcal{C}_x := [0,x] = \prod_{i=1}^d [0,x_i] \)
for \( x \in X, \mathcal{C} := \mathcal{C}_d := \{ C_x : x \in X \}, \) and let \( P \) be the uniform
law on the Borel sets \( \mathcal{G} \) of \( X. \) Then 
\( E_t = \{ x \in X : x_1 \cdots x_d < t \}, \)
so 
\( a(t) = a_*(t). \) Here \( \nu_n(C_x) \) is the normalized empirical
distribution function at \( x. \) We claim that
\[ a(t) = \sum_{j=0}^{d-1} (\log \frac{1}{t})^j/j!, \quad 0 < t \leq 1. \]

Proceeding by induction on \( d, \) we suppose this to be true
for \( d = k, \) and let \( a_i(t) \) be the function \( a(t) \) correspond-
ing to \( C_i \) and \( P. \) Then
\[ a_{k+1}(t) = t + \int_0^1 a_k(t/x) dx \]
\[ = t + \sum_{j=0}^k \int_0^{1/\log (x/t)} \frac{1}{x} \left( \log \frac{1}{x} \right)^j dx \]
\[ = t + \sum_{j=0}^k \int_0^{\frac{1}{\log (1/t)}} \frac{\log (1/t)}{u} u^j du \]
\[ = t \sum_{j=0}^{k+1} (\log \frac{1}{t})^{j+1}/(j+1)! \]
and the claim follows. Hence

\[ a(t) \sim t(\log (1/t))^{d-1}/(d-1)! \text{ as } t \to 0, \]

so (6.1) implies (6.2) if \( d \geq 2 \). Let \( b_d(t) := t(\log (1/t))^{d-1} \), and let \( V \) be a countable subset of \( X \) such that for each \( x \in X \) there exists a sequence \( \{x(n)\} \) in \( V \) with \( x_i^{(n)} \to x_i \) as \( n \to \infty \) for each \( i \leq d \). Applying Lemma 2.10 with \( \mathcal{D} := \mathcal{C}_1(t, \mathcal{C}, P), \mathcal{C}_2(t, \mathcal{C}, P), \) and 

\[ \mathcal{S}_0 := \{C_x : x \in V, P(C_x) < t\} \text{ or } \{C_x \setminus C_y : x, y \in V, P(C_x \setminus C_y) < t\}, \]

we see that \( \mathcal{C} \) is size-1 measurable with constants for \( P \), and that \( D_0(\mathcal{C}, P, q) \) is admissible for \( \{v_n/q_0 P\} \) and \( P \) if \( v_n/q_0 P \) is \( \mathcal{L}(\mathcal{C}) \)-valued a.s. for all \( n \). Hence Theorem 6.1 and Corollary 6.2 give the following, which was proved in part by O'Reilly (1974) in the case \( d = 1 \), when it is also related to the functional LIL of James (1975).

**Corollary 6.10.** Let \( q \in \mathcal{Q}_1 \) and let \( P \) be the uniform law on \([0,1]^d\). Then \( v_n/q_0 P \) converges in law in \((D_0(\mathcal{C}, P, q), \mathcal{B}_p))\) to \( G_p/q_0 P \), and the latter has a.s. uniformly \( d_p \)-continuous sample paths, if and only if

\[ \int_0^\infty \exp(-\varepsilon q^2(t)/t) \frac{dt}{t} < \infty \text{ for all } \varepsilon > 0 \text{ if } d = 1 \]

\[ q^2(t)/(t(\log (1/t))^{d-1} \to \infty \text{ as } t \to 0 \text{ if } d \geq 2. \]
If \( \int_0^\infty b_d(q^{-1}((xLLx)^{-1/2}))dx < \infty \) and (6.71) holds, then \( \{\nu_n/q^oP\} \) satisfies the compact and functional LIL's in \( \ell^\infty(\mathbb{C}) \).

When \( d = 1 \), as discussed by O'Reilly (1974), the condition (6.71) for convergence in law is the same as that for sample continuity of the limiting Gaussian process \( G_{P}/q^oP \). This is not true for \( d > 2 \), however; by Theorem 2.2 of Pruitt and Orey (1973), the function \( q(t) = (t \log t)^{1/2} \) makes \( G_{P}/q^oP \) sample-continuous but does not satisfy (6.71).

For \( C = \{C^C: C \in \mathbb{C}_d\} \), it is easy to show that \( a(t) \sim dt \) as \( t \to 0 \).

**Example 6.11.** Let \( d \geq 1 \), \( X := \mathbb{R}^d \), and \( C^b_\omega := \{x \in \mathbb{R}^d: x \cdot \omega \geq b\} \) for \( \omega \) in the sphere \( S^{d-1} \) and \( b \in \mathbb{R} \). Let \( P \) be a nondegenerate normal law on \( \mathbb{R}^d \), and let \( C \) be the class \( \{C^b_\omega: (\omega, b) \in S^{d-1} \times \mathbb{R}\} \) of all closed half spaces in \( \mathbb{R}^d \).

Then there is an affine map \( A: \mathbb{R}^d \to \mathbb{R}^d \) with \( \mathcal{L}(A) = N(0, I) \). Since \( A \) maps \( C \) one-to-one onto itself, \( a(t) \) is the same for \( P \) as for \( N(0, I) \). Hence we may assume \( P = N(0, I) \).

Let \( \phi \) be the distribution function on \( \mathbb{R} \) of \( N(0,1) \), and let \( \chi^2_d \) be a chi-squared r.v. with \( d \) degrees of freedom. Now \( P(C^b_\omega) < t \) if and only if \( b > \phi^{-1}(1-t) \), so

\[ E_t = \{x \in \mathbb{R}^d: ||x|| > \phi^{-1}(1-t)\} \]. Let \( r_t := \phi^{-1}(1-t) \);

since \( r_t \sim (2 \log(1/t))^{1/2} \) and \( r_t^{-1}\exp(-r_t^2/2) \sim (2\pi)^{1/2}t \).
as \( t \to 0 \), it follows that as \( t \to 0 \),

\[
a(t) = P[|X_i|^2 > r_t^2] = Pr[x^2_d > r_t^2] \sim C_1 r_t^{d-2} \exp(-r_t^2/2).
\]

where \( C_1 \) and \( C_2 \) are constants.

Let \( V \) be a countable dense subset of \( S^{d-1} \times \mathbb{R} \); similarly to Example 6.9, we see that \( C \) is size-1 measurable with constants for \( P \), and that \( D_0(C,P,q) \) is admissible for \( \{v_n/q\circ P\} \) and \( P \) provided \( v_n/q \circ P \) is \( \ell^\infty(C) \)-valued a.s. for all \( n \). Thus Theorem 6.1 and Corollary 6.2 give the following.

**Corollary 6.12.** Let \( q \in Q_1 \), let \( P \) be a nondegenerate normal law on \( \mathbb{R}^d \), and let \( C \) be the class of all closed half spaces in \( \mathbb{R}^d \). Then \( v_n/q \circ P \) converges in law in \( (D_0(C,P,q),C_0) \) to \( G_P/q \circ P \), and the latter has a.s. uniformly \( d_P \)-continuous sample paths, if and only if

\[
\int_0^\infty \exp(-\epsilon q^2(t)/t) \frac{dt}{t} < \infty \text{ for all } \epsilon > 0 \text{ if } d = 1
\]

\[
(6.72)
\]

\[
q^2(t)/t(\log \frac{1}{t})^{(d-1)/2} \to \infty \text{ as } t \to 0 \text{ if } d \geq 2.
\]
132.

If \( \int_{0}^{\infty} f_{d}(q^{-1}\{(x\log x)^{-1/2}\}) \, dx < \infty \) and (6.72) holds, where

\[ f_{d}(t) = t(\log \frac{1}{t})^{(d-1)/2}, \]

then \( \{v_{n}/q^{o}P\} \) satisfies the compact and functional LIL's in \( \ell^{\infty}(\mathbb{C}) \). [□]

The functions \( a(t) \) which can occur for some \( a(0) = 0 \) and \( a(1) \leq 1 \). In fact, let \( a(t) \) be any left-continuous increasing function on \([0,1]\) with \( a(0) = 0 \) and \( a(1) \leq 1 \). Then letting \( P \) be the uniform law on \([0,1]\) and \( C := \{[\alpha,\beta]: 0 \leq \alpha < \beta \leq a(\beta-\alpha)\} \cup \{[0,t]: a(t+) = t\} \) gives rise to the given function \( a(t) \). Note that in general \( a \) is left-continuous if \( E_{t} \) is measurable for all \( t \).
INDEX OF NOTATION

<table>
<thead>
<tr>
<th>Notation</th>
<th>Page</th>
<th>Notation</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q</td>
<td>12</td>
<td>K₁(v)</td>
<td>36</td>
</tr>
<tr>
<td>A \ B</td>
<td>15</td>
<td>K₂(v)</td>
<td>50</td>
</tr>
<tr>
<td>Aₙ</td>
<td>101</td>
<td>K₃(v)</td>
<td>62</td>
</tr>
<tr>
<td>a(t)</td>
<td>98</td>
<td>Lₓ</td>
<td>68</td>
</tr>
<tr>
<td>a⋆(t)</td>
<td>98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A(v)</td>
<td>16</td>
<td>L∞(x)</td>
<td>12</td>
</tr>
<tr>
<td>Aᵦ</td>
<td>20</td>
<td>mᵦ(n)</td>
<td>13</td>
</tr>
<tr>
<td>B</td>
<td>18</td>
<td>N(ε, C, H)</td>
<td>16</td>
</tr>
<tr>
<td>b</td>
<td>18</td>
<td>P</td>
<td>12</td>
</tr>
<tr>
<td>B(x, r)</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>42</td>
<td>P(i)</td>
<td>12</td>
</tr>
<tr>
<td>Bₜ, ε(P)</td>
<td>24</td>
<td>P_j, N</td>
<td>42</td>
</tr>
<tr>
<td>C</td>
<td>13</td>
<td>P[λ,m]</td>
<td>71</td>
</tr>
<tr>
<td>C({•})</td>
<td>27</td>
<td>P_n</td>
<td>12</td>
</tr>
<tr>
<td>C_b,u(C, d_p)</td>
<td>18</td>
<td>p'_n</td>
<td>25</td>
</tr>
<tr>
<td>C₁(t, C, H)</td>
<td>47</td>
<td>p_(1)</td>
<td>42</td>
</tr>
<tr>
<td>C₂(t, C, H)</td>
<td>83</td>
<td>p_(2)</td>
<td>42</td>
</tr>
<tr>
<td>d_p</td>
<td>16</td>
<td>P(n)</td>
<td>13</td>
</tr>
<tr>
<td>D₀(C, d_p)</td>
<td>128</td>
<td>P(n)</td>
<td>26</td>
</tr>
<tr>
<td>Eₜ</td>
<td>98</td>
<td>P(n, 2)</td>
<td>58</td>
</tr>
<tr>
<td>F_c</td>
<td>21</td>
<td>P(N)</td>
<td>42</td>
</tr>
<tr>
<td>G_p</td>
<td>17</td>
<td>Pr</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pr[N]</td>
<td>42</td>
</tr>
<tr>
<td>Notation</td>
<td>Page</td>
<td>Notation</td>
<td>Page</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>----------</td>
<td>------</td>
</tr>
<tr>
<td>Pr*</td>
<td>24</td>
<td>ρ_n</td>
<td>83</td>
</tr>
<tr>
<td>Pr</td>
<td>24</td>
<td>ρ_p</td>
<td>84</td>
</tr>
<tr>
<td>q</td>
<td>81</td>
<td>σ(i)</td>
<td>42</td>
</tr>
<tr>
<td>q^{-1}</td>
<td>90</td>
<td>τ(i)</td>
<td>42</td>
</tr>
<tr>
<td>Q</td>
<td>42,81</td>
<td>Q</td>
<td>42</td>
</tr>
<tr>
<td>Q_1</td>
<td>83</td>
<td>V_C</td>
<td>14</td>
</tr>
<tr>
<td>V(ϖ)</td>
<td>14</td>
<td>v_1(v)</td>
<td>50</td>
</tr>
<tr>
<td>x</td>
<td>12</td>
<td>X_i</td>
<td>12</td>
</tr>
<tr>
<td>x(n)</td>
<td>32</td>
<td>x(n)</td>
<td>32</td>
</tr>
<tr>
<td>y</td>
<td>42</td>
<td>α(ϖ,n)</td>
<td>32</td>
</tr>
<tr>
<td>β(ϖ,n)</td>
<td>57</td>
<td>Δ_C(F)</td>
<td>13</td>
</tr>
<tr>
<td>ζ_t,n</td>
<td>82</td>
<td>v_n</td>
<td>13</td>
</tr>
<tr>
<td>v_0</td>
<td>25</td>
<td>v_0</td>
<td>42</td>
</tr>
<tr>
<td>v_n</td>
<td>42</td>
<td>v_(j)</td>
<td>42</td>
</tr>
</tbody>
</table>
REFERENCES


136.


137.


About the Author

Kenneth Alexander was born March 3, 1958 and raised in Seattle, Washington, where he attended Shoreline High School and the University of Washington, receiving his B.S. in 1979. Since 1979 he has studied at M.I.T. under a National Science Foundation Graduate Fellowship. In August 1982 he will be married to Chris Czarnecki.