The Annals of Probability
1987, Vol. 15, No. 3, 837–870

THE GLIVENKO–CANTELLI PROBLEM

BY MICHEL TALAGRAND

University of Paris VI and The Ohio State University

We give a new type of characterization of the Glivenko–Cantelli classes. In the case of a class $\mathscr{C}$ of sets, the characterization is closely related to the configuration that the sets of $\mathscr{C}$ can have. It allows one to decide simply whether a given class is a Glivenko–Cantelli class. The characterization is based on a new measure theoretic analysis of sets of measurable functions. This analysis also gives an approximation theorem for Glivenko–Cantelli classes, sharpenings of the Vapnik–Červonenkis criteria and the value of the asymptotic discrepancy for classes that are not Glivenko–Cantelli. An application is given to the law of large numbers in a Banach space for functions that need not be random variables.

1. Notation and main results. Let (Ω, Σ, P) be a complete probability space. On Ω^n we denote by P^n the product probability. On $\Omega^\infty = \Omega^{\mathbb{N}}$, we denote by P^∞ the product probability. The corresponding outer (resp. inner) probabilities are denoted by P^*, P^{n*}, $P^{\infty *}$ (resp. P_*, P^n_*, P^∞_*).

We denote by $\mathscr{L}^1 = \mathscr{L}^1(P)$ the set of measurable functions f such that $E(|f|) = \int |f|\, dP < \infty$. We shall *not* identify functions in $\mathscr{L}^1$ with their classes in L^1. We say that a subset Z of $\mathscr{L}^1$ is order bounded if there is $u \in \mathscr{L}^1$, $u \geq 0$, such that for each f in Z we have $|f| \leq u$ *everywhere*.

Given $s = (s_i) \in \Omega^\infty$, $n \in \mathbb{N}$, we consider the empirical measure

$$Q_n(s) = \frac{1}{n} \sum_{i \leq n} \delta_{s_i},$$

so, if $f \in \mathscr{L}^1$, we have

$$Q_n(s)(f) = \frac{1}{n} \sum_{i \leq n} f(s_i).$$

Given a set $Z \subset \mathscr{L}^1$, and $s \in \Omega^\infty$, we define the discrepancy $D_n(s)$ of Z by

$$D_n(s) = \sup_{f \in Z} |Q_n(s)(f) - E(f)|.$$

We shall say that Z is a Glivenko–Cantelli class if $D_n(s) \to 0$ a.s. (Note that for possibly nonmeasurable functions, a.s. convergence does not imply convergence in probability.) A class of sets is called a Glivenko–Cantelli class if the class of the indicator functions of its sets is a Glivenko–Cantelli class. This name comes from the fact that the Glivenko–Cantelli theorem can be reformulated by saying that the set of intervals $(-\infty, t]$ of $\mathbb{R}$ is a Glivenko–Cantelli class for each probability P on $\mathbb{R}$.

Received October 1984; revised September 1986.

AMS 1980 *subject classifications*. Primary 60F15, 60B12, 28A20; secondary 28A51, 60F05.

Key words and phrases. Uniform law of large numbers, empirical process, empirical discrepancy, Pettis norm.

The study of the Glivenko–Cantelli property of classes of subsets of the Euclidean space, in relation with their geometrical properties, has attracted long-standing attention. These classes include convex sets, half spaces, etc. See [3] for a discussion. These results have been obtained by special methods.

For uniformly bounded classes, and with some measurability conditions, Vapnik and Červonenkis [11] obtained a characterization of Glivenko–Cantelli classes in the abstract setting. Their result has been recently extended to order bounded classes by Giné and Zinn [4]. Given $s \in \Omega^\infty$, consider the set

$$U = \{(f(s_1), \dots, f(s_n));\ f \in Z\} \subset \mathbb{R}^n.$$

Provide $\mathbb{R}^n$ with the sup norm (resp. the norm $\|x\| = (1/n)\Sigma_{i \le n}|x_i|$), and denote by $N_n^\infty(Z, \varepsilon, s)$ [resp. $N_n^1(Z, \varepsilon, s)$] the smallest number of closed balls of radius ε that covers U. [Note that $N_n^1(Z, \varepsilon, s) \le N_n^\infty(z, \varepsilon, s)$.] Then under some measurability conditions, Z is a Glivenko–Cantelli class if and only if Z is order bounded and for each $\varepsilon > 0$

$$\lim \frac{1}{n} \int \log N_n^\infty(Z, \varepsilon, s)\, dP^\infty(s) = 0$$

[resp. $\limsup_n N_n^1(Z, \varepsilon, s) < \infty$ a.s.].

In practice, this result is not satisfactory, since the quantities involved in this criterion can be computed only with difficulty, when such computation is at all possible. Some other sufficient conditions, of much easier use, are surveyed in [3], but they are by no means necessary. So there remains the need for a simple, necessary and sufficient criterion. The criterion we propose is necessary and sufficient in the most general case, but it in the end is based on what is the common property of the geometries of the various Glivenko–Cantelli classes of subsets of $\mathbb{R}^n$. It is precisely this feature that makes its application easy. The basic notion is as follows.

DEFINITION 1. A set Z of functions on Ω is called *stable* if for each $\alpha < \beta$, and each set $A \in \Sigma$ with $P(A) > 0$, there exists $n > 0$ such that

$$P^{2n*}(\{(s_1, \dots, s_n, t_1, \dots, t_n) \in A^{2n};\ \exists f \in Z;$$
$$\forall i \le n,\ f(s_i) < \alpha,\ f(t_i) > \beta\}) < P(A)^{2n}.$$

(A set of functions on a probability space might be called stable if every finite subset has a stable joint distribution. The definition just given, however, is unrelated to stable laws.)

The main result of this work is

THEOREM 2. *For a subset Z of $\mathscr{L}^1$, the following are equivalent:*

(a) *Z is a Glivenko–Cantelli class and $\{E(f);\ f \in Z\}$ is bounded;*
(b) *Z is stable and order bounded.*

As a first illustration, let us prove the extended Blum–DeHardt law of large numbers ([1], 6-1-5) that is, Z is a Glivenko–Cantelli class if for each $\varepsilon > 0$, there exists a finite family $f_1, \dots, f_p, g_1, \dots, g_p$ of $\mathscr{L}^1$, such that for $i \le p$, $f_i \le g_i$, $\int (g_i - f_i)\, dP \le \varepsilon$ and that for each $f \in Z$, there exists $i \le p$ with $f_i \le f \le g_i$. First, taking $\varepsilon = 1$, we see that Z is order bounded and clearly $Z \subset \mathscr{L}^1$. Suppose that Z is not stable. Then there is a set $A \in \Sigma$, with $P(A) > 0$ and $\alpha < \beta$ such that for each $n > 0$, we have

$$P^{2n*}\big(\{(s_1, \dots, s_n, t_1, \dots, t_n) \in A^{2n};\ \exists\, f \in Z;\ \forall\, i \le n,\ f(s_i) < \alpha,\ f(t_i) > \beta\}\big) = P(A)^{2n}.$$

Let $\varepsilon < (\beta - \alpha)P(A)$. Consider a finite family $f_1, \dots, f_p, g_1, \dots, g_p$ of $\mathscr{L}^1$ such that $f_i \le g_i$ and $\int (g_i - f_i)\, dP \le \varepsilon$ for $i \le p$. For $i \le p$, if the set $A \cap \{g_i \le \beta\}$ has positive measure let $B_i = A \cap \{g_i \le \beta\}$. If $A \cap \{g_i \le \beta\}$ is negligible, then by choice of ε, $A \cap \{f_i \ge \alpha\}$ has positive measure, and we set $B_i = A \cap \{f_i \ge \alpha\}$, so either $B_i \subset \{g_i \le \beta\}$ or $B_i \subset \{f_i \ge \alpha\}$. Let

$$C = \big\{(s_1, \dots, s_p, t_1, \dots, t_p) \in A^{2p};\ \forall\, i \le p,\ s_i, t_i \in B_i\big\}.$$

Then $P^{2p}(C) > 0$ and $C \subset A^{2p}$; so there exists $(s_1, \dots, s_p, t_1, \dots, t_p) \in C$ and $f \in Z$ with $f(s_i) < \alpha$, $f(t_i) > \beta$ for each $i \le p$. If $B_i \subset \{g_i \le \beta\}$, then $f(t_i) > g_i(t_i)$. If $B_i \subset \{f_i \ge \alpha\}$, then $f(s_i) < f_i(s_i)$. In particular for no i do we have $f_i \le f \le g_i$; this contradiction concludes the proof.

As it stands, Definition 1 is not as appealing as the reader might have expected. It however simplifies if we make some measurability assumptions.

DEFINITION 3. We say that Z satisfies condition (M) if for each $\alpha < \beta$ and each $n \in \mathbb{N}$, the set

$$\big\{(s_1, \dots, s_n, t_1, \dots, t_n) \subset \Omega^{2n};\ \exists\, f \in Z, \forall\, i \le n,\ f(s_i) < \alpha,\ f(t_i) > \beta\big\}$$

is measurable.

PROPOSITION 4. *If Z satisfies condition* (M), *Z fails to be stable if and only if there exist $\alpha < \beta$ and $A \in \Sigma$, with $P(A) > 0$, such that for each n, and P^n almost each $s \in A^n$, for each subset I of $\{1, \dots, n\}$, there is $f \in Z$ with $f(s_i) < \alpha$ for $i \in I$ and $f(s_i) > \beta$ for $i \notin I$.*

Roughly speaking, this means that for the generic choice $s_1, \dots, s_n$ in A, the functions of Z oscillate wildly over $s_1, \dots, s_n$. The name stable attempts to convey the idea that such a wild oscillation does not occur.

It seems worthwhile to reformulate Theorem 2 in the most important case.

THEOREM 5. *Let Z be a class of measurable sets that satisfies* (M). *Then Z fails to be a Glivenko–Cantelli class if and only if there exists $A \in \Sigma$, $P(A) > 0$, P atomless on A [that is, if $B \subset A$, $B \in \Sigma$, $P(B) > 0$, there is $C \subset B$, $C \in \Sigma$, $0 < P(C) < P(B)$] such that for each n, for P^n almost every choice of $s_1, \dots, s_n$ in A, each subset of $\{s_1, \dots, s_n\}$ is the trace on $\{s_1, \dots, s_n\}$ of a set of Z.*

Note that the condition "P atomless on A" is equivalent to saying that for almost each choice of $s_1, \dots, s_n$ in A, the points $s_1, \dots, s_n$ are distinct.

Condition (M) cannot be dropped from this statement. Let us compare this result with the result of Vapnik–Červonenkis [11]. They prove that if Z is not a Glivenko–Cantelli class, then there exists $a > 0$, such that for n large, and most of the choices of $(s_1, \dots, s_n) \in \Omega^n$, there is a subset I of $\{x_1, \dots, x_n\}$ with card $I \geq an$, such that each subset of I is the trace on I of an element of Z. The gain in Theorem 5 is that we know that actually one can take $I = A \cap \{x_1, \dots, x_n\}$ for some fixed A with $P(A) > 0$. Roughly speaking, Theorem 5 means that the trace on A of the sets of Z can produce arbitrarily complicated patterns. So, in checking that a class is a Glivenko–Cantelli class, it is enough to check that this pathology occurs on no set of positive measure. It might seem a drawback to have to inspect every set of positive measure. However, the pathology that we look for is so strong that it is easy to recognize.

Many classes of subsets of $\mathbb{R}^n$ which are known to be Glivenko–Cantelli classes for some probability have some geometrical properties which prevent their sets to be of complicated shape (although, on the other hand, any disjoint family of measurable sets is a Glivenko–Cantelli class). It is this geometrical property which is used to apply the criteria of Theorem 5. As a specific example, let us outline a proof that if μ is a probability on $\mathbb{R}^n$, such that $\mu(\partial C) = 0$ for each convex set C, then the class Z of closed convex sets is a Glivenko–Cantelli class [2] and [6]. Let $A \in \Sigma$ with $\mu(A) > 0$. We can as well suppose A compact and self-supported (that is, for U open, $U \cap A \neq \varnothing$, $\mu(A \cap U) > 0$). Let C be the convex hull of A. Since $\mu(\partial C) = 0$, we can choose $x_0 \in A \setminus \partial C$. We can then pick $x_1, \dots, x_{n+1} \in A$ and $\alpha_1, \dots, \alpha_{n+1} > 0$, of sum 1, with $x_0 = \sum_{i \leq n+1} \alpha_i x_i$. We can choose x_i so that the simplex with vertices x_i has nonempty interior. Next, we can choose for $0 \leq i \leq n+1$ a neighborhood V_i of x_i, such that if for $0 \leq i \leq n+1$, s_i belongs to V_i, then s_0 belongs to the interior of the simplex of vertices $s_1, \dots, s_{n+1}$. Since A is self-supported, we have $P^{n+2}(B) > 0$, where

$$B = \left\{(s_0, s_1, \dots, s_{n+1}) \in A^{n+2}; \forall\, i, 0 \leq i \leq n+1,\ s_i \in V_i\right\}.$$

However, for $(s_0, s_1, \dots, s_{n+1}) \in B$, each convex set that contains $s_1, \dots, s_{n+1}$ also contains s_0 so $\{s_1, \dots, s_{n+1}\}$ is not the trace on $\{s_0, \dots, s_{n+1}\}$ of a convex set. This concludes the proof.

Consider (for simplicity) a Glivenko–Cantelli class Z of sets, that satisfies condition (M). For $A \in \Sigma$, $P(A) > 0$, let $n(A)$ be the largest integer n such that

$$P^n(\{(s_1, \dots, s_n) \in A^n; \text{each subset of } \{s_1, \dots, s_n\} \text{ is the trace of a set in } Z\}) = P(A)^n.$$

In the example in $\mathbb{R}^n$ that we have described, we had $n(A) \leq n + 1$ for each A. It is not always the case that the numbers $n(A)$ are bounded, but it may be of interest to note that for each $\delta > 0$, we have $\sup\{n(A); P(A) \geq \delta\} < \infty$. This question is addressed at the beginning of Section 2, as an introduction to our methods.

We will analyze in great detail the behavior of $D_n(s)$ when Z is not a Glivenko–Cantelli class, and show that Z is a Glivenko–Cantelli class if and only if an array of equivalent (some of them formally weaker) conditions hold. Among these conditions, the following approximation theorem might be the most important.

THEOREM 6. *Let $Z \subset \mathscr{L}^1$ be order bounded. Then Z is a Glivenko–Cantelli class if and only if for each $\varepsilon > 0$ there is a finite subalgebra $\mathscr{A}$ of Σ such that*

$$\limsup_n \int^* \sup_{f \in Z} \left(\frac{1}{n} \sum_{i \le n} |f(s_i) - E(f|\mathscr{A})(s_i)| \right) dP^\infty(s) \le \varepsilon.$$

The point of this theorem is that the absolute values are inside the summation, so there is no cancellation.

The following is a sharpening of the Vapnik–Červonenkis result. Theorem 6 will be proved as part of Theorem 22. Then it will take a few more pages to prove Theorem 7.

THEOREM 7. (a) *Let $Z \subset \mathscr{L}^1$. If Z is stable, and if there is a $u \ge 0$ such that $|f| \le u$ for each $f \in Z$ and $\log(u + 1) \in \mathscr{L}^1$, then for each $\varepsilon > 0$,*

$$\lim_n \int^* \frac{1}{n} \log N_n^\infty(Z, \varepsilon, s)\, dP^\infty(s) = 0.$$

(b) *Let $Z \subset \mathscr{L}^1$ be stable and order bounded. Then for each $\varepsilon > 0$ there is a constant $c(\varepsilon) > 0$ such that*

$$\limsup_n N_n^1(Z, \varepsilon, s) \le c(\varepsilon) \quad a.s.$$

(c) *Conversely, if $Z \subset \mathscr{L}^1$ satisfies condition* (M), *and if for each $\varepsilon > 0$ and each $a > 0$ we have*

$$\limsup_n P_*^\infty \left\{ \frac{1}{n} \log N_n^1(Z, \varepsilon, \cdot) < a \right\} > 0,$$

then Z is stable.

The measurability condition in (c) cannot be dropped.

We will also use Theorem 6 to prove a new stability property of Glivenko–Cantelli classes, and a comparison principle for stochastic processes that satisfy the law of large numbers.

In Section 4, we apply our result to the study of the law of large numbers for Banach space valued functions. If E is a Banach space, say that the map ϕ from Ω to E is properly measurable if the set

$$Z_\phi = \{x^* \circ \phi;\ x^* \in E^*,\ \|x^*\| \le 1\}$$

is stable. Properly measurable functions have already proved to be an effective tool in the study of Pettis integration. They are studied in great detail in this respect in [9]. Here we show

THEOREM 8. *The sequence $(1/n)\Sigma_{i \le n}\phi(s_i)$ converges in norm for almost each (s_i) in Ω^∞ if and only if ϕ is properly measurable and $\int^* \|\phi\|\, dP < \infty$. In* *that case, ϕ is Pettis integrable, and* $(1/n)\Sigma_{i \le n}\phi(s_i)$ *converges a.s in norm to the Pettis integral of ϕ.*

Now a few comments about the history of the methods and the results. The notion of stable sets was developed by D. H. Fremlin and the author in the course of their study of pointwise compact sets of measurable functions, as being the most natural criterion for pointwise compactness. The author then undertook in [9] (where more historical comments can be found) a systematic study of this notion and its application to various questions of measure theory, mostly Pettis integration. One natural question about stable sets is as follows: Is the convex hull of a uniformly bounded stable set still stable? This is far from clear from the definition. To prove it, the author established Theorem (11-1-1) of [9]. In the language of the present paper, this result means that a uniformly bounded set of functions Z is stable if and only if D_n goes to zero in probability. Due to his ignorance of probability theory, the author became aware that he had essentially solved the Glivenko–Cantelli problem only after he discussed the material of [5] with Professor Hoffmann–Jørgensen. In the present paper, we undertake a systematic development of the ideas of Theorem (11-1-1) of [9], which was considered there as a purely technical point, with the aim of obtaining the sharpest possible result, and the ambition of settling the Glivenko–Cantelli problem as well as related questions. For example, we have taken care to approach the problem so that we cannot only describe Glivenko–Cantelli classes, but also exactly describe the behavior of the asymptotic discrepancy D_n for any bounded class of functions. These and other refinements increase the technicalities. It thus might be helpful to guide the reader. Section 2 develops some aspects of the theory of stable sets. The main results, on which the paper relies, are Theorems 16 and 17. They are by far the hardest of the paper, and their proof is not probabilistic. After understanding the statements of these theorems, the reader may like to delay reading the proofs in Section 2 and go directly to Section 3, to see how these theorems relate to the Glivenko–Cantelli problem. The arguments there are much more familiar. [Reading of the proof of the delicate Theorem 20, which is needed only to add the equivalent condition (VIII) in Theorem 22, should also be delayed.] Then the reader should visit Section 4. The simplicity of the statements there is a welcome reward for all the earlier technical work. Only then, when motivation has been found, should the reader go through the proofs of Section 2. The arguments there are self-contained, and we have tried to give enough details that the proofs can be checked step by step. However, the underlying ideas are intricate, and the technical devices are not quite standard (although purely measure-theoretic). The reader would be better armed to penetrate the ideas of the proofs with some previous exposure to related (but simpler) ideas, as can be found in Chapters 8 and 9 of [9]. The main idea of the approach actually goes back to [8], Theorem 5.

2. Structure results for sets of functions. Given a complete probability space (Ω', Ξ, Q) and $X \subset \Omega'$, a measurable cover A of X is a measurable set A with $A \supset X$ and $Q(A) = Q^*(X)$. Given a real-valued function g on Ω', we denote by g^* the essential infimum of the measurable functions $f: \Omega \to \mathbb{R}$ such that $f \geq g$. We have $\int^* g\,dP = \int g^*\,dP$ when both these quantities are defined.

For two sets $A, B \subset \Omega'$, we write $A \supset_{\text{ess}} B$ if $Q(B \setminus A) = 0$.

Let $A \in \Sigma$ with $P(A) > 0$. Let u, v be two functions on Ω. Throughout the paper we will use the following notation

$$\begin{aligned} B_{k,l}(A, u, v) &:= B_{k,l}(Z, A, u, v) \\ &= \{(s_1, \dots, s_k, t_1, \dots, t_l) \in A^{k+l}; \exists f \in Z, \forall i \leq k, \\ &\qquad f(s_i) < u(s_i), \forall j \leq l, f(t_j) > v(t_j)\}. \end{aligned}$$

We will often consider $B_{k,l}(A, \alpha, \beta)$, where $\alpha < \beta$ are constant.

We first prove Proposition 4. Assume that Z is not stable and satisfies condition (M). There is $A \in \Sigma$ with $P(A) > 0$ and $\alpha < \beta$ such that for all n, $P^{2n*}(B_{n,n}) = P(A)^{2n}$, where $B_{k,l} = B_{k,l}(A, \alpha, \beta)$.

Condition (M) implies that $B_{n,n}$ is measurable, so $P^{2n}(B_{n,n}) = P^{2n*}(B_{n,n}) = P(A)^{2n}$. Let $k, l \in \mathbb{N}$ and $n = \max(k, l)$. Let π be the map from A^{2n} to A^{k+l} that sends $(s_1, \dots, s_n, t_1, \dots, t_n)$ to $(s_1, \dots, s_k, t_1, \dots, t_l)$. Then $B_{n,n} \subset \pi^{-1}(B_{k,l})$, and this forces $P^{k+l}(B_{k,l}) = P(A)^{k+l}$. It is now obvious that for each n

$$\begin{aligned} P^n(\{(s_1, \dots, s_n) \in A^n; \forall I \subset \{1, \dots, n\}, \exists f \in Z; \forall i \in I, f(s_i) < \alpha, \\ \forall i \notin I, f(s_i) > \beta\}) = P(A)^n, \end{aligned}$$

and this concludes the proof.

This might be the proper time to observe that a finite set of functions is stable if and only if they are measurable.

We now start an auxiliary discussion that is not needed for the main chain of arguments (which starts with Lemma 12) but that will shed some light on the use of the criterion of Theorem 5.

Let us fix a stable set of functions Z and $\alpha < \beta$. For each n, let $B_{n,n} = B_{n,n}(\Omega, \alpha, \beta)$, and let C_n be a measurable cover of $B_{n,n}$. For a set $A \in \Sigma$, with $P(A) > 0$, we have $P^{2n*}(B_{n,n}(A, \alpha, \beta)) = P(A)^{2n}$ if and only if $A^{2n} \subset_{\text{ess}} C_n$. We denote by $\mathscr{F}_n$ the family of measurable subsets of Ω with this property, and we note that $\mathscr{F}_{n+1} \subset \mathscr{F}_n$.

LEMMA 9. *Let (A_k) be a sequence in $\mathscr{F}_n$ and v be a weak cluster point of (1_{A_k}) in $L^2(P)$. Then $\{v > 0\} \in \mathscr{F}_n$.*

PROOF. In $L^2(P^{2n})$, the function $v^{\otimes 2n}$ given by $v^{\otimes 2n}(s_1, \dots, s_{2n}) = \prod_{i \leq 2n} v(s_i)$ is a weak cluster point of $1_{A_k^{2n}}$. It follows that $\{v > 0\}^{2n} \subset_{\text{ess}} C_n$, so $\{v > 0\} \in \mathscr{F}_n$. □

PROPOSITION 10. $\lim_n \sup_{A \in \mathscr{F}_n} P(A) = 0$.

PROOF. Otherwise there is a sequence (A_n) with $P(A_n) \geq \delta > 0$ and $A_n \in \mathscr{F}_n$ for each n. Let v be a weak cluster point of 1_{A_n} in $L^2(P)$. Then $P(\{v > 0\}) \geq \delta > 0$ and $\{v > 0\} \in \mathscr{F}_n$ for each n, which contradicts the fact that Z is stable. □

Proposition 11. *Fix $n > 0$. Let $\delta = \sup_{A \in \mathscr{F}_n} P(A)$. Then for each $\varepsilon > 0$ there is a finite family $\mathscr{G}_n$ of measurable sets, such that $P(B) \le \delta + \varepsilon$ for $B \in \mathscr{G}_n$, and that for each $A \in \mathscr{F}_n$, there is $B \in \mathscr{G}_n$ with $P(A \setminus B) \le \varepsilon$.*

Proof. Otherwise, we can construct by induction a sequence (A_k) of $\mathscr{F}_n$, such that whenever D is a set of the algebra generated by $A_1, \dots, A_{k-1}$ with $P(D) \le \delta + \varepsilon$, we have $P(A_k \setminus D) \ge \varepsilon$. Denote by Ξ the algebra generated by the sets (A_k). Let v be a weak cluster point in $L^2(P)$ of 1_{A_k}. Since by Lemma 9, $\{v > 0\} \in \mathscr{F}_n$, we have $P(\{v > 0\}) \le \delta$. It follows that there is a k and a set D in the algebra generated by $A_1, \dots, A_k$ with $P(D) \le \delta + \varepsilon$ and $P(\{v > 0\} \setminus D) < \varepsilon/2$. This implies $P(A_l \setminus D) \le \varepsilon/2$ for some $l > k$, a contradiction. □

We now start the main chain of lemmas. Let $J = [-1/2, 1/2]$. We denote by λ Lebesgue's measure on J. Let $\Omega' = \Omega \times J$. For $v \in J^{\Omega}$, let

$$S(v) = \{(s,t) \in \Omega';\ t \le v(s)\};\ A_n(v) = S(v)^n \subset \Omega'^n.$$

Let $\nu = P \otimes \lambda$.

Lemma 12. *Let $\alpha \in \mathbb{R}$ and let (v_n) be a sequence of measurable functions valued in $[-3/8, 1/2]$ with $E(v_n) \ge \alpha$. Then there is a measurable function v valued in $[-3/8, 1/2]$ with $E(v) \ge \alpha$ and*

$$\forall\, k,\ A_k(v) \subset_{\text{ess}} \limsup_n A_k(v_n).$$

Proof. By taking a subsequence, we can assume that (v_n) converges weakly in $L^2(P)$ to some function v and that $1_{S(v_n)}$ converges weakly in $L^2(\nu)$ to some function f. The function v is valued in $[-3/8, 1/2]$. We first show that $f > 0$ a.e. on $S(v)$. Otherwise, there is a measurable set $C \subset S(v)$ with $\nu(C) > 0$ and $f = 0$ on C. So, there is $\varepsilon > 0$ and a measurable $D \subset \Omega$ with $P(D) > 0$ such that for $s \in D$ we have

$$\lambda(\{t \in J;\ (s,t) \in C\}) > \varepsilon.$$

For each n, set

$$D_n = \{s \in D;\ v_n(s) \ge v(s) - \varepsilon/2\}.$$

Since $-1/2 \le v_n \le 1/2$, we have

$$\begin{aligned}
\int_D v_n\,dP &\le \int_{D \setminus D_n} v_n\,dP + \frac{1}{2} P(D_n) \\
&\le \int_{D \setminus D_n} v\,dP - \frac{\varepsilon}{2} P(D \setminus D_n) + \frac{1}{2} P(D_n) \\
&\le \int_D v\,dP - \frac{\varepsilon}{2} P(D \setminus D_n) + P(D_n) \\
&\le \int_D v\,dP - \frac{\varepsilon}{2} P(D) + \left(1 + \frac{\varepsilon}{2}\right) P(D_n).
\end{aligned}$$

So

$$\left(1+\frac{\varepsilon}{2}\right)P(D_n) \geq \frac{\varepsilon}{2}P(D) + \int_D (v_n - v)\, dP$$

and this shows that $\liminf_n P(D_n) > 0$. Note that for $s \in D_n$,

$$\begin{aligned}\lambda(\{t \in J;\ (s,t) \in C;\ -1/2 \leq t \leq v_n(s)\}) &\\ \geq \lambda(\{t \in J, (s,t) \in C, t \leq v(s)\}) - \varepsilon/2 &\\ = \lambda(\{t \in J, (s,t) \in C\}) - \varepsilon/2 \geq \varepsilon - \varepsilon/2 \geq \varepsilon/2.\end{aligned}$$

We have that

$$\begin{aligned}\nu(S(v_n) \cap C) &\geq \int_{D_n} \lambda(\{t \in J;\ (s,t) \in C,\ -1/2 \leq t \leq v_n(s)\})\, dP(s)\\ &\geq \varepsilon P(D_n)/2.\end{aligned}$$

On the other hand, $\nu(S(v_n) \cap C) \to \int_C f\, d\nu = 0$. This contradiction proves the claim.

Let us fix k. The weak limit in $L^2(\nu^k)$ of $1_{S^k(v_n)}$ is the function $f^{\otimes k}$ given by $f^{\otimes k}(x_1, \ldots, x_k) = f(x_1) \times \cdots \times f(x_k)$, so it follows that $\{f > 0\}^k \subset_{\text{ess}} \limsup_n S^k(v_n)$ and in particular $S(v)^k \subset_{\text{ess}} \limsup_n S^k(v_n)$, that is $A_k(v) \subset_{\text{ess}} \limsup_n A_k(v_n)$. The proof is complete. □

We now fix a set Z of functions. We do not assume Z to be stable, or even the functions in Z to be measurable. Let $V := V(Z) := V(Z, P)$ be the set of all measurable functions v such that

$$\forall\, n,\ P^{n*}(\{(s_1, \ldots, s_n) \in \Omega^n;\ \exists\, f \in Z, \forall\, i \leq n,\ f(s_i) > v(s_i)\}) = 1.$$

Let $d(Z) = \sup\{E(v);\ v \in V(Z)\}$. [If no such v exists, we set $d(Z) = -\infty$.] We now suppose $Z \subset [-1/4, 1/2]^{\Omega}$. Let $A_n = \bigcup_{v \in Z} A_n(v)$; and let B_n be a measurable cover of A_n.

Lemma 13. *Let $a > d(Z)$. Then there are $n \in \mathbb{N}$ and $\zeta > 0$ such that for each measurable v valued in $[-3/8, 1/2]$ we have*

$$\nu^n(A_n(v) \setminus B_n) < \zeta \Rightarrow E(v) < a.$$

Proof. Otherwise, for each n, there is a measurable function v_n valued in $[-3/8, 1/2]$ with $\int v_n\, dP \geq a$ and $\nu^n(A_n(v_n) \setminus B_n) \leq 2^{-4n}$. Let $m \leq n$. We first show that

$$\nu^m(A_m(v_n) \setminus B_m) \leq 2^{-n}. \tag{1}$$

Let π be the natural projection of Ω'^n on Ω'^m. Obviously, $\pi(A_n) \subset A_m$, so $\pi(A_n) \subset B_m$ and $A_n \subset \pi^{-1}(B_m)$. Since $\pi^{-1}(B_m)$ is measurable and B_n is a measurable cover of A_n, we have $B_n \subset_{\text{ess}} \pi^{-1}(B_m)$, or $B_n \subset_{\text{ess}} B_m \times \Omega'^{n-m}$. We have

$$\begin{aligned}(A_m(v_n) \setminus B_m) \times S(v_n)^{n-m} &= A_m(v_n) \times S(v_n)^{n-m} \setminus (B_m \times S(v_n)^{n-m})\\ &= A_m(v_n) \times S(v_n)^{n-m} \setminus (B_m \times \Omega'^{n-m})\\ &= A_n(v_n) \setminus (B_m \times \Omega'^{n-m}) \subset_{\text{ess}} A_n(v_n) \setminus B_n.\end{aligned}$$

It follows that

$$\nu^m(A_m(v_n)\setminus B_m) \le 2^{-4n}\nu(S(v_n))^{m-n}.$$

Since v_n is valued in $[-3/8, 1/2]$, we have $\nu(s(v_n)) \ge 2^{-3}$, so (1) follows.

From Lemma 12, there exists a measurable function v valued in $[-3/8, 1/2]$ with $E(v) \ge a$ and $A_m(v) \subset_{\text{ess}} \limsup_n A_m(v_n)$ for each m. It follows from (1) that $A_m(v) \subset_{\text{ess}} B_m$ for each m. Let $v' = (v - (a - d(Z))/2) \vee (-1/2)$, so $E(v') > d(Z)$. Since $v \ge -3/8$, we have $v(s) > v'(s)$ everywhere. We show that for each m we have $P^{m*}(K_m) = 1$, where

$$K_m = \{(s_1, \dots, s_m) \in \Omega^m; \exists\, f \in Z, \forall\, i \le m,\ f(s_i) > v'(s_i)\}.$$

This will contradict the definition of $d(Z)$ and finish the proof. Let $C \subset \Omega^m$ with $P^m(C) > 0$. We have to show that $K_m \cap C \neq \varnothing$. Let

$$D = \{((s_1, t_1), \dots, (s_m, t_m)) \in \Omega'^m;\ (s_1, \dots, s_m) \in C,\\ \forall\, i \le m,\ v'(s_i) < t_i < v(s_i)\},$$

so $D \subset A_m(v)$ and

$$\nu^m(D) \ge \int_C \prod_{i=1}^{m} (v(s_i) - v'(s_i))\, dP^m(s_1, \dots, s_m).$$

Since $v'(s) < v(s)$ for each s, we have $\nu^m(D) > 0$, and since $A_m(v) \subset_{\text{ess}} B_m$ we have $D \subset_{\text{ess}} B_m$.

Since B_m is a measurable cover of A_m, we have $A_m \cap D \neq \varnothing$. So there is $f \in Z$ with $A_m(f) \cap D \neq \varnothing$. Let $((s_1, t_1), \dots, (s_m, t_m)) \in D \cap A_m(f)$. For $i \le m$, we have $v'(s_i) < t_i$ and $t_i < f(s_i)$, so $v'(s_i) < f(s_i)$. This shows that

$$(s_1, \dots, s_m) \in C \cap K_m.$$

The proof is complete. □

We say that a probability is *atomic* if it has finite support.

LEMMA 14. *Let $\delta > 0$. Then there is a finite subalgebra $\mathscr{A}$ of Σ, $n \in \mathbb{N}$, a set $D \in \Sigma^n$, and $\eta > 0$ such that whenever μ is an atomic probability on (Ω, Σ) that satisfies*

(2) $\quad |\mu^n(D) - P^n(D)| < \eta; \qquad \forall\, A$ *an atom of* $\mathscr{A}$, $|\mu(A) - P(A)| < \eta,$

then for each $f \in Z$, there is an $\mathscr{A}$-measurable function f' with $\int f'\, d\mu \le d(Z) + \delta$ and $\mu(\{f > f'\} \cap A) \le \delta\mu(A)$ for each atom A of $\mathscr{A}$.

PROOF. We can assume $\delta/3 < 1/8$. From Lemma 13 there is $n \in \mathbb{N}$ and $0 < \zeta < 1$ such that for each measurable v valued in $[-3/8, 1/2]$ we have

$$\nu^n(A_n(v)\setminus B_n) < \zeta \Rightarrow E(v) < d(Z) + \delta/3.$$

Let $\gamma = (\delta/2)^n\zeta/2$. A measurable set in a finite product of measure spaces can always be approximated by a set measurable with respect to a product of finite algebras. So, there is a finite subalgebra $\mathscr{A}$ of Σ such that each atom of $\mathscr{A}$ has positive measure, and a partition $\mathscr{P}$ of J into right closed intervals of length $< \delta/3$, such that if $\mathscr{B}$ denotes the algebra generated by $\mathscr{P}$ there is $C \in (\mathscr{A} \times \mathscr{B})^n$ such that $\nu^n(C \vartriangle B_n) < \gamma^2/8$. Let

$$D = \{(s_1, \dots, s_n) \in \Omega^n;$$
$$\lambda^n(\{t_1, \dots, t_n\} \in J^n; ((s_1, t_1), \dots, (s_n, t_n)) \in B_n \setminus C) > \gamma/2\},$$

so $P^n(D) \leq \gamma/4$ by Fubini. We can assume that $B_n \in (\Sigma \times \text{Borel})^n$, and then we have $D \in \Sigma^n$. We note that for each probability μ, if $\mu^n(D) < \gamma/2$ we have

$$(\mu \times \lambda)^n(B_n \setminus C) \leq \mu^n(D) + (\gamma/2)\mu^n(\Omega^n \setminus D) < \gamma.$$

Let

$$\eta = \inf\left(\frac{\gamma}{4}, \frac{\delta}{3}\inf\{P(A), A \text{ an atom of } \mathscr{A}\}\right).$$

Since we assumed that each atom of $\mathscr{A}$ has positive measure, we have $\eta > 0$. Let $f \in Z$. Let μ be an atomic probability that satisfies (2). Define f' in the following way: The value of f' on an atom A of $\mathscr{A}$ is given by

$$\inf\{t \in [-3/8, 1/2]; \mu(A \cap \{f > t\}) \leq \delta\mu(A)\}.$$

(We note that since μ is atomic, the set $A \cap \{f > t\}$ is μ-measurable.) It follows that $\mu(\{f > f'\} \cap A) \leq \delta\mu(A)$. Define an $\mathscr{A}$-measurable function v in the following way: If L denotes the set of ends of intervals of $\mathscr{P}$,

$$v(t) = \max\{x; x < f'(t); x \in L\}.$$

We have $f' - v < \delta/3$, so we have $E(f' - v) < \delta/3$. Since $f \geq -1/4$, and since $\delta/3 < 1/8$, v is valued in $[-3/8, 1/2]$. We show now that

$$\nu^n(A_n(v) \setminus B_n) < \zeta.$$

This will imply $E(v) < d(Z) + \delta/3$ so $E(f') \leq d(Z) + 2\delta/3$. The definition of η and (2) imply that for each atom A of $\mathscr{A}$, $|\mu(A) - P(A)| \leq \delta P(A)/3$, so $|E(f') - \int f' \, d\mu| \leq \delta/3$, so this will finish the proof. Since $\gamma \leq \zeta/2$, and since $\nu^n(C \setminus B_n) < \gamma$, it is enough to show that $\nu^n(A_n(v) \setminus C) < \zeta/2$. We note that $A_n(v)$ is $(\mathscr{A} \times \mathscr{B})^n$-measurable. Let $C_1, \dots, C_n$ be atoms of $\mathscr{A}$, and $I_1, \dots, I_n$ be atoms of $\mathscr{B}$ such that

$$U := \prod_{i \leq n} (C_i \times I_i) \subset A_n(v) \setminus C.$$

For $i \leq n$, let $D_i = C_i \cap \{f \geq v\}$. Since $v < f'$, the definition of f' shows that $\mu(D_i) \geq \delta\mu(C_i)$. Moreover, for each $i \leq n$,

$$|\mu(C_i) - P(C_i)| < \eta \leq (\delta/3)P(C_i) < P(C_i)/2,$$

so

$$\mu(C_i) \geq P(C_i) - |\mu(C_i) - P(C_i)| \geq P(C_i) - P(C_i)/2 \geq P(C_i)/2,$$

so $\mu(D_i) \geq (\delta/2)P(C_i)$. Since $C_i \times I_i \subset S(v)$, we have $v \geq \sup I_i$ on C_i. Since

$f \geq v$ on D_i, we have $D_i \times I_i \subset S(f)$. It follows that

$$\prod_{i\leq n}(D_i \times I_i) \subset A_n(f), \quad \text{so } \prod_{i\leq n}(D_i \times I_i) \subset A_n(f) \cap U.$$

We now have

$$\begin{aligned}(\delta/2)^n \nu^n(U) &= (\delta/2)^n \prod_{i\leq n} P(C_i)\lambda(I_i)\\ &\leq \prod_{i\leq n}\mu(D_i)\lambda(I_i) = (\mu\times\lambda)^n\Big(\prod_{i\leq n}(D_i \times I_i)\Big)\\ &\leq (\mu\times\lambda)^n(A_n(f)\cap U).\end{aligned}$$

Summation over all the atoms U of $(\mathscr{A}\times\mathscr{B})^n$ that are contained in $A_n(v)\setminus C$ gives

$$\begin{aligned}(\delta/2)^n \nu^n(A_n(v)\setminus C) &\leq (\mu\times\lambda)^n(A_n(f)\setminus C)\\ &\leq (\mu\times\lambda)^n(B_n\setminus C) < \gamma \leq (\delta/2)^n\zeta/2,\end{aligned}$$

so $\nu^n(A_n(v)\setminus C) < \zeta/2$. The proof is complete. □

COROLLARY 15. *Let $a > d(Z)$. Then there exists $n \in \mathbb{N}$, $\eta > 0$, a finite subalgebra $\mathscr{A}$ of Σ, and $D \in \Sigma^n$, such that for each atomic probability μ on (Ω, Σ) that satisfies* (2), *we have $\int f\,d\mu \leq a$ for each $f \in Z$.*

We now need a result that relates the properties of a set Z of measurable functions and of the set $Z^c = \{f - E(f);\ f \in Z\}$. It is possible to approach this result through a "symmetrized" version of Lemma 14, but we choose a more direct way.

THEOREM 16. (a) *Let Z_1, Z_2 be uniformly bounded sets of measurable functions. Then $d(Z_1 \cup Z_2) \leq \max(d(Z_1), d(Z_2))$.*

(b) *Let Z be a uniformly bounded set of measurable functions, and let $Z^c = \{f - E(f);\ f \in Z\}$. Then if Z is stable, $d(Z^c) = 0$.*

(c) *Let Z be a uniformly bounded set of functions. Assume that Z is not stable. Let A with $P(A) > 0$, $\alpha < \beta$ be such that for each n,*

$$P^{2n*}(B_{n,n}(A, \alpha, \beta)) = P(A)^{2n}.$$

Then there exist two measurable functions u, v with $E(v) > E(u) + (\beta - \alpha)P(A)/3$ such that for each n, we have $P^{2n*}(B_{n,n}(\Omega, u, v)) = 1$.

(d) *If Z, Z^c are as in* (b), *Z^c is stable whenever Z is stable.*

PROOF. (a) Consider a bounded function v with $E(v) > \max(d(Z_1), d(Z_2))$. For $i = 1, 2$ and each n, denote

$$K_{n,i} = \{(s_1, \ldots, s_n) \in \Omega^n;\ \exists f \in Z_i, \forall j \leq n,\ f(s_j) > v(s_j)\}.$$

Since $E(v) > d(Z_i)$, we see easily that for some n and some $\delta < 1$, we have $P^{n*}(K_{n,i}) \leq \delta$ for $i = 1, 2$. Let p be large enough that $\delta^p < 1/2$. Let $m = pn$.

Since $K_{m,i} \subset (K_{n,i})^p$, we have $P^{m*}(K_{m,i}) < \delta^p$ for $i = 1,2$, so $P^{m*}(K_{m,1} \cup K_{m,2}) < 1$. In other words,

$$P^{m*}(\{(s_1,\dots,s_m) \in \Omega^m;\ \exists\, f \in Z_1 \cup Z_2, \forall\, i \le n,\ f(s_i) > v(s_i)\}) < 1.$$

This proves the result.

(b) *Step* 1. We can suppose $Z \subset B = \{f \in \mathscr{L}^\infty,\ \|f\|_\infty \le 1\}$. Since $d(Z) \ge E(f)$ for any f in Z, we have $d(Z^c) \ge 0$. Suppose, if possible, that $d(Z^c) > 0$. Then there exists v with $E(v) > 0$ and for all n, $P^{n*}(K_n) = 1$, where

$$K_n = \{(s_1,\dots,s_n) \in \Omega^n;\ \exists\, f \in Z, \forall\, i \le n,\ f(s_i) - E(f) > v(s_i)\}.$$

We first prove the following claim: There is h in B, such that for each weak neighborhood V of h in $L^2(P)$, and each n, we have

$$P^{n*}(\{(s_1,\dots,s_n) \in \Omega^n;\ \exists\, f \in V \cap Z, \forall\, i \le n,\ f(s_i) - E(f) > v(s_i)\}) = 1.$$

Otherwise, we can cover B by a finite collection of weakly open sets $(V_j)_{j\le k}$ such that

$$P^{n_j*}\Big(\big\{(s_1,\dots,s_{n_j}) \in \Omega^{n_j};\ \exists\, f \in V_j \cap Z, \forall\, i \le n_j,\ f(s_i) - E(f) > v(s_i)\big\}\Big) < 1,$$

for some $n_j \in \mathbb{N}$. Let $n = \max\{n_j;\ j \le k\}$. Then there is a $\delta < 1$ with

$$P^{n*}\big(\{(s_1,\dots,s_n) \in \Omega^n;\ \exists\, f \in V_j \cap Z, \forall\, i \le n,\ f(s_i) - E(f) > v(s_i)\}\big) < \delta,$$

for $j \le k$. Choose p in $\mathbb{N}$ with $\delta^p < 1/k$. Let $m = np$. Then for $j \le k$, we have $P^{m*}(K_{m,j}) < \delta^p$, where

$$K_{m,j} = \big\{(s_1,\dots,s_m) \in \Omega^m;\ \exists\, f \in V_j \cap Z, \forall\, i \le m,\ f(s_i) - E(f) > v(s_i)\big\}.$$

Since $K_m \subset \bigcup_{j\le k} K_{m,j}$, we get $P^{m*}(K_m) < 1$. This contradiction concludes the proof of the claim.

Step 2. Let v and h be as above. Let $a = E(v)/3$, so $E(v + E(h) - a) > E(h + a)$ and $a \le 1/3$. We show that $P^{2n*}(L_n) = 1$, where

$$L_n = \big\{(s_1,\dots,s_n,t_1,\dots,t_n) \in \Omega^{2n};\ \exists\, f \in Z, \forall\, i \le n,$$
$$f(t_i) < a + h(t_i),\ f(s_i) > v(s_i) + E(h) - a\big\}.$$

Let $C \subset \Omega^{2n}$ with $P^{2n}(C) > 0$. We will show that $C \cap L_n \neq \varnothing$. Let $\delta = (3a/14)^n$. There exist measurable sets $B_1,\dots,B_{2n}$, of positive measure, with

$$P^{2n}\Big(C \cap \prod_{i\le 2n} B_i\Big) > (1-\delta) P^{2n}\Big(\prod_{i\le 2n} B_i\Big).$$

For $s = (s_1,\dots,s_n) \in \Omega^n$, let

$$C(s) = \Big\{(t_1,\dots,t_n) \in \Omega^n;\ (s_1,\dots,s_n,t_1,\dots,t_n) \in C \cap \prod_{i\le 2n} B_i\Big\}.$$

Let

$$D = \Big\{(s_1,\dots,s_n) \in \Omega^n;\ P^n(C(s)) > (1-\delta) P^n\Big(\prod_{n<i\le 2n} B_i\Big)\Big\},$$

so $P^n(D) > 0$.

Step 1 shows that $P^{n*}(T_n) = 1$, where

$$T_n = \Big\{(s_1, \dots, s_n) \in \Omega^n;\ \exists\, f \in Z,\ |E(f) - E(h)| < a,$$
$$\forall\, i \le n,\ f(s_i) > v(s_i) + E(f),$$
$$\forall\, n < i \le 2n,\ \int_{B_i} f\,dP \le \int_{B_i} h\,dP + aP(B_i)/2\Big\}.$$

Since $P^n(D) > 0$, $P^{n*}(T_n) = 1$ we have $D \cap T_n \ne \varnothing$. So there exist $s = (s_1, \dots, s_n) \in D$ and $f \in Z$, which satisfy the following conditions:

$$\forall\, i \le n, \qquad f(s_i) > v(s_i) + E(f); \qquad |E(f) - E(h)| < a,$$
$$\forall\, n < i \le 2n, \qquad \int_{B_i} f\,dP \le \int_{B_i} h\,dP + aP(B_i)/2.$$

For $n < i \le 2n$, let $D_i = B_i \cap \{f < h + a\}$. We get, since $-1 \le f,\ h \le 1$ and $a \le 1/3$,

$$\int_{B_i} f\,dP \ge \int_{B_i \setminus D_i} (h + a)\,dP - P(D_i) \ge \int_{B_i} h\,dP + aP(B_i) - \tfrac{7}{3}P(D_i).$$

Since $\int_{B_i} f\,dP \le \int_{B_i} h\,dP + aP(B_i)/2$, we get $P(D_i) \ge 3aP(B_i)/14 = \delta^{1/n}P(B_i)$, so

$$P^n\Big(\prod_{n<i\le 2n} D_i\Big) > \delta P^n\Big(\prod_{n<i\le 2n} B_i\Big).$$

Since

$$P^n(C(s)) > (1 - \delta)P^n\Big(\prod_{n<i\le 2n} B_i\Big),$$

we have $C(s) \cap \prod_{n<i\le 2n} D_i \ne \varnothing$, so there is $(t_1, \dots, t_n) \in C(s)$ such that $f(t_i) < h(t_i) + a$ for $i \le n$. Since $E(f) \ge E(h) - a$, we have $f(s_i) > v(s_i) + E(h) - a$ for $i \le n$, so $(s_1, \dots, s_n, t_1, \dots, t_n) \in L_n$.

Since $(t_1, \dots, t_n) \in C(s)$, we have $(s_1, \dots, s_n, t_1, \dots, t_n) \in C$, so $L_n \cap C \ne \varnothing$. Since C is arbitrary, we have $P^{2n*}(L_n) = 1$.

Since $E(v + E(h) - a) > E(h + a)$, there exists a set A with $P(A) > 0$, and $\alpha < \beta$ with $A \subset \{v + E(h) - a > \beta\} \cap \{h + a < \alpha\}$. It follows from $P^{2n*}(L_n) = 1$ that

$$P^{2n*}\big(\{(s_1, \dots, s_n, t_1, \dots, t_n) \in A^{2n};\ \exists\, f \in Z, \forall\, i \le n,$$
$$f(t_i) < \alpha,\ f(s_i) > \beta\}\big) = P(A)^{2n}.$$

This shows that Z is not stable and finishes the proof.

(c) Take $A \subset \Omega$ with $P(A) > 0$ and $\alpha < \beta$ such that for each n we have $P^{2n*}(B_{n,n}(A, \alpha, \beta)) = P(A)^{2n}$. We leave to the reader to use the method of (b), step 1, to show that there is a function h such that for each weak neighborhood V of h in $L^2(P)$ and all n we have $P^{2n*}(B_{n,n}(Z \cap V, A, \alpha, \beta)) = P(A)^{2n}$. Let $a = (\beta - \alpha)P(A)/3$. Let u be given by $u = h + a$ on $\Omega \setminus A$ and $u = \alpha$ on A. Let

v be given by $v = h - \alpha$ on $\Omega \setminus A$ and $v = \beta$ on A. Then $E(v) \geq E(u) + \alpha$. For two subsets I, J of $\{1, \dots, n\}$, let

$$K_{I,J} = \left\{(s_1, \dots, s_n, t_1, \dots, t_n) \in \Omega^{2n};\ s_i \in A \Leftrightarrow i \in I;\ t_j \in A \Leftrightarrow j \in J\right\}.$$

We also leave to the reader to use the method of (b), step 2, to show that for each n, I and J,

$$P^{2n*}\big(\{(s_1, \dots, s_n, t_1, \dots, t_n) \in K_{I,J};\ \exists\, f \in Z, \forall\, i \leq n,\\ f(s_i) < u(s_i),\ f(t_i) > v(t_i)\}\big) = P^{2n}(K_{I,J}).$$

Summation over I, J gives $P^{2n*}(B_{n,n}(\Omega, u, v)) = 1$, and this completes the proof.

(d) If Z is stable, then obviously $-Z = \{-f;\ f \in Z\}$ is stable, so by (b), $d(Z^c) = d(-Z^c) = 0$. If Z^c is not stable, by (c) there exist two measurable functions u, v with $E(u) < E(v)$ such that for each n

$$P^{2n*}\big(\{(s_1, \dots, s_n, t_1, \dots, t_n) \in \Omega^{2n};\\ \exists\, f \in Z^c, \forall\, i \leq n,\ -f(s_i) > -u(s_i),\ f(t_i) > v(t_i)\}\big) = 1.$$

This shows that $0 = d(Z^c) \geq E(v)$, and $0 = d(-Z^c) \geq E(-u)$, so $E(v) - E(u) \leq 0$, a contradiction.

This completes the proof. □

The following result goes a long way toward understanding the structure of stable sets.

THEOREM 17. *Let Z be a uniformly bounded set of measurable functions. Then the following are equivalent*:

(a) *Z is stable*;

(b) *for each $\varepsilon > 0$, there exists a finite subalgebra $\mathscr{A}$ of Σ and a finite family $\mathscr{F}$ of pairs (n, D), for $n \in \mathbb{N}$, $D \in \Sigma^n$, and $\gamma > 0$ such that for each atomic probability μ on (Ω, Σ) that satisfies*

$$\forall\, (n, D) \in \mathscr{F}, \qquad |\mu^n(D) - P^n(D)| < \gamma,$$

we have

$$\forall\, f \in Z, \qquad \int |f - E(f|\mathscr{A})|\, d\mu < \varepsilon.$$

PROOF. (a) ⇒ (b). For a function g, write $g^+ = \max(g, 0)$, $g^- = \max(-g, 0)$, so $g = g^+ - g^-$, $|g| = g^+ + g^-$. We first note the following consequence of Lemma 14. If T is a uniformly bounded set, with $d(T) \leq 0$, and $\delta > 0$, there is $\eta > 0$ and a finite family $\mathscr{G}$ of pairs (n, D) for $n \in \mathbb{N}$, $D \in \Sigma^n$ such that for each atomic probability measure μ on (Ω, Σ) that satisfies

$$\forall\, (n, D) \in \mathscr{G}, \qquad |\mu^n(D) - P^n(D)| < \eta,$$

for each f in T we have $f \leq f_1 + h$, where f_1 is $\mathscr{A}$-measurable, $\int f_1\, d\mu \leq \delta$ and $\int |h|\, d\mu \leq \delta$.

We first apply this observation to $Z^c = \{f - E(f);\ f \in Z\}$, with $\delta = \varepsilon/11$. From Theorem 16(b), we have that $d(Z^c) = 0$. (We shall not use explicitly the fact that Z^c itself is stable.) We then find a finite family $\mathscr{G}$ of pairs (n, D), for $n \in \mathbb{N}$, $D \in \Sigma^n$, such that for each atomic probability measure μ on (Ω, Σ) that satisfies

$$\forall (n, D) \in \mathscr{G}, \qquad |\mu^n(D) - P^n(D)| < \eta,$$

for each f in Z^c we have $f \le f_1 + h$ where f_1 is $\mathscr{A}$-measurable, $\int f_1\, d\mu \le \delta$ and $\int |h|\, d\mu \le \delta$.

We now enumerate the atoms of $\mathscr{A}$ as $A_1, \dots, A_k$. We can assume that they are all of positive measure. For $1 \le j \le k$, consider the probability P_j on (Ω, Σ) given by $P_j(B) = P^{-1}(A_j)P(B \cap A_j)$ for $B \in \Sigma$. It is obvious from the definition that Z is stable on (Ω, Σ, P_j). So $-Z$ is also stable on (Ω, Σ, P_j). It follows from Theorem 16(b) applied to $-Z$ in (Ω, Σ, P_j) that if we set

$$T_j = \left\{-f + P^{-1}(A_j)E\left(f1_{A_j}\right);\ f \in Z\right\},$$

then $d(T_j) = 0$ (where the basic probability is still P_j).

Consider now

$$Z_j = \left\{(-f + E(f|\mathscr{A}))1_{A_j};\ f \in Z\right\}.$$

We show first that, for the probability P, we have $d(Z_j) \le 0$. Let $v \in V(Z_j, P)$, so we also have $v \in V(Z_j, P_j)$. The restrictions of Z_j and T_j to A_j coincide. Since $P_j(A_j) = 1$, we have $v \in V(T_j, P_j)$. Since $d(T_j) = 0$ for the basic probability space P_j, we have $\int_{A_j} v\, dP \le 0$. Since the functions of Z_j are zero outside A_j, we have $v < 0$ a.s. outside A_j, so $E(v) < \int_{A_j} v\, dP \le 0$. This shows that $d(Z_j) \le 0$.

We now use the form of Lemma 14 that was spelled out at the beginning of the proof, to obtain $\eta_j > 0$ and a finite family $\mathscr{G}_j$ of pairs (n, D), for $n \in \mathbb{N}$, $D \in \Sigma^n$, such that for each atomic probability measure μ on (Ω, Σ) that satisfies

$$\forall (n, D) \in \mathscr{G}_j, \qquad |\mu^n(D) - P^n(D)| < \eta_j,$$

we have $\int g\, d\mu \le \delta/k$ for each f in Z_j. So we have $\int(-f + E(f|\mathscr{A}))1_{A_j}\, d\mu \le \delta/k$ for each f in Z. Consider now the collection $\mathscr{F}$ of all pairs (n, D) that belong either to $\mathscr{G}$ or to one of the $\mathscr{G}_j$, $1 \le j \le k$, or are of the type $(1, A)$ for A atom of $\mathscr{A}$. Let $0 < \alpha \le \min(\eta, \eta_j, 1 \le j \le k)$. We can assume that α is small enough that for each function f in Z, for each probability measure μ on (Ω, Σ) such that $|\mu(A) - P(A)| < \alpha$ for each atom A of $\mathscr{A}$, we have $|\int E(f|\mathscr{A})\, d\mu - E(f)| \le \delta$. Consider now an atomic probability measure μ on (Ω, Σ) such that $|\mu^n(D) - P^n(D)| < \alpha$ for each $(n, D) \in \mathscr{F}$. Fix f in Z. Then we have $f - E(f) \le f_1 + h$ where, f_1 is $\mathscr{A}$-measurable, $\int f_1\, d\mu \le \delta$, $\int |h|\, d\mu \le \delta$ and for $1 \le j \le k$ we have $\int_{A_j}(f - E(f|\mathscr{A}))\, d\mu \ge -\delta/k$.

Let $f_2 = f_1 + E(f) - E(f|\mathscr{A})$, so f_2 is $\mathscr{A}$-measurable and $\int f_2\, d\mu \le 2\delta$. Let $g = f - E(f|\mathscr{A})$, so $g \le f_2 + h$. Integrating over A_j, we find that the integral of f_2 over A_j is $\ge -\delta/k - \int_{A_j}|h|\, d\mu$. So we have $\int f_2^-\, d\mu \le 2\delta$, so $\int f_2^+\, d\mu = \int f_2^-\, d\mu + \int f_2\, d\mu \le 4\delta$, so since $g^+ \le f_2^+ + h^+$, we have $\int g^+\, d\mu \le 5\delta$. Since for each $1 \le j \le k$, we have $\int_{A_j} g\, d\mu \ge -\delta/k$, we have $\int g\, d\mu \ge -\delta$. This shows that $\int^{-} g\,d\mu = \int^{+} g\,d\mu - \int g\,d\mu \le 6\delta$. So we have $\int |f - E(f|A)|\,d\mu = \int |g|\,d\mu \le 11\delta = \varepsilon$ and the proof is complete.

(b) $\Rightarrow$ (a). If Z is not stable, from Theorem 16(c) there are measurable functions u, v with $E(u) < E(v)$ such that for each n, $P^{2n*}(B_{n,n}(\Omega, u, v)) = 1$. Let $\varepsilon > 0$ with $6\varepsilon < E(v) - E(u)$ and let $\mathscr{A}$ be a finite subalgebra of Σ and $\mathscr{F}$ be a finite family of pairs (n, D) and $\gamma > 0$ such that for each atomic probability μ on (Ω, Σ) we have

$$\forall\,(n, D) \in \mathscr{F},\quad |\mu^n(D) - P^n(D)| < \gamma \Rightarrow \int |f - E(f|\mathscr{A})|\,d\mu < \varepsilon, \quad \forall f \in Z.$$

For each m and $s \in \Omega^{2m}$, let

$$\mu_{1,s} = m^{-1} \sum_{i \le m} \delta_{s_i}, \qquad \mu_{2,s} = m^{-1} \sum_{m < i \le 2m} \delta_{s_i}.$$

It follows from Lemma 18 (to be proved later) that for m large enough we have $P^{2m}(B) > 0$, where

$$B = \Big\{ s \in \Omega^{2m};\ \forall\, i = 1,2,\ \forall\,(n, D) \in \mathscr{F},\ |\mu_{i,s}^n(D) - P^n(D)| < \gamma;$$
$$|\mu_{i,s}(u) - E(u)| < \varepsilon,\ |\mu_{i,s}(v) - E(v)| < \varepsilon,\ \forall f \in L^\infty,\ \|f\| \le 1,$$
$$\Big|\int E(f|\mathscr{A})\,d\mu_{i,s} - E(f)\Big| < \varepsilon \Big\}.$$

It follows that there is $s \in B$ and $f \in Z$ with

$$\forall\, i \le m,\ f(s_i) < u(s_i); \qquad \forall\, i,\ m < i \le 2m,\ f(s_i) > v(s_i),$$

so we have

$$\varepsilon \ge \int |f - E(f|\mathscr{A})|\,d\mu_{1,s} \ge \int E(f|\mathscr{A})\,d\mu_{1,s} - \int u\,d\mu_{1,s} \ge E(f) - E(u) - 2\varepsilon,$$

$$\varepsilon \ge \int |f - E(f|\mathscr{A})|\,d\mu_{2,s} \ge \int v\,d\mu_{2,s} - \int E(f|\mathscr{A})\,d\mu_{2,s} \ge E(v) - E(f) - 2\varepsilon,$$

so $E(v) - E(u) \le 6\varepsilon$, a contradiction. This completes the proof. □

3. Laws of large numbers. The following is a special case of the convergence of V statistics. The simple proof is included for completeness.

LEMMA 18. *Let $A \subset \Omega^k$ be measurable. Then*

$$P^\infty \text{ a.s.} \qquad \lim_n Q_n^k(t)(A) = P^k(A).$$

PROOF. Let

$$Q_n'(t) = \frac{1}{n(n-1)\cdots(n-k+1)} \sum \delta_{t_{i_1}} \times \delta_{t_{i_2}} \times \cdots \times \delta_{t_{i_k}},$$

where the summation is over all the choices $(i_1, \dots, i_k)$ of distinct integers $\le n$.

It is enough to show that $Q'_n(t)(A) \to P^k(A)$ a.s., since $\|Q^k_n(t) - Q'_n(t)\| \to 0$, where $\|\cdot\|$ is the total variation norm. Let Σ_n be the subalgebra of Σ^∞, generated by the sets that are invariant by permutation of the first n coordinates. It is easy to check that $(Q'_n(\cdot)(A))_n$ is a martingale with respect to the decreasing filtration (Σ_n). So it converges a.s. to its expectation $P^k(A)$. □

Now let Z be a set of (not necessarily measurable) real-valued functions on Ω. For $s \in \Omega^\infty$, let

$$G_n(s) = \sup_{f \in Z} \frac{1}{n} \sum_{i \le n} f(s_i).$$

The relevance of this quantity to the study of the empirical discrepancy of a subset of $\mathscr{L}^1$ is as follows. Suppose that Z is an order-bounded set of measurable functions. Let $Z' = Z^c \cup (-Z^c)$. Then the empirical discrepancy of Z,

$$D_n(x) = \sup_{f \in Z} \left| \frac{1}{n} \sum_{i \le n} f(s_i) - E(f) \right|,$$

coincides with the function $G_n(x)$ relative to the set Z'. So, the forthcoming Theorems 19 and 20 can be seen as results about the empirical discrepancy. We have stated them separately since they are valid for sets of functions that need not be Glivenko–Cantelli classes. It is also useful to note that by Theorem 16(a), $d(Z') = \max(d(Z^c), d(-Z^c))$. (This fact will become clearer in light of the proof of Theorem 19 below.)

Since Z is uncountable, even when Z consists of measurable functions, there is no reason why $G_n(s)$ should be measurable, so it is natural to consider the function $G_n^*(s)$. Also of interest are the functions $\limsup G_n(s)$ and $\liminf G_n(s)$. As these functions do not depend on the first coordinates, the zero one law shows that $(\limsup G_n)^*$, $(\limsup G_n)_*$, $(\liminf G_n)^*$ and $(\liminf G_n)_*$ are all constant a.s. In general, the numbers $(\limsup G_n)_*$ and $(\liminf G_n)_*$ bear little signification, as the following example shows: Let (Ω, Σ, P) be $[0,1]$ with Lebesgue's measure, let $X \subset \Omega$ with $P^*(X) = 1$, $P^*(\Omega \setminus X) = 1$, and let

$$Z = \{ f \colon \Omega \to \{0,1\};\ f \text{ has finite support, } f = 0 \text{ on } X \}.$$

Using the fact that $P^{\infty *}(X^{\mathbb{N}}) = P^{\infty *}((\Omega \setminus X)^{\mathbb{N}}) = 1$, it is clear that $(\limsup G_n)_* = (\liminf G_n)_* = 0$, while

$$(\liminf G_n)^* = (\limsup G_n)^* = 1, \qquad \limsup G_n^*(s) = 1 \quad \text{a.s.}$$

So in that case $(\liminf G_n)_*$ and $(\limsup G_n)_*$ do not carry much information about the behavior of Z. As we shall see, the quantities $(\limsup G_n)^*$ and $(\liminf G_n)^*$ are much more instructive.

THEOREM 19. *Let Z be a set of (not necessarily measurable) functions on Ω. Suppose there is $u \in \mathscr{L}^1(P)$ such that $|f| \le u$ for $f \in Z$. Then*

(a) $\lim G_n^*(s) = d(Z)$, *where the convergence is a.s. and in* $L^1(P^\infty)$;
(b) $(\limsup G_n)^* = d(Z)$.

PROOF. (a) Let $\varepsilon > 0$. Let a be large enough that $\|u'\|_1 < \varepsilon$, where $u' = u - u \wedge a$. For $f \in L^1$, let $f^a = (f \wedge a) \vee (-a)$, so $|f - f^a| \le u'$ for $f \in Z$. Let $Z' = \{f^a;\ f \in Z\}$.

We now check that $d(Z') < d(Z) + 2\varepsilon$. Let $v \in V(Z')$. Let $\eta > 0$. Define w by $w = v$ on $\{u \le a\}$ and $w = -u - \eta$ on $\{u > a\}$. We then have $|w - v| \le 2u' + \eta$, so $E(w) \ge E(v) - 2Eu' - \eta$. On the other hand, if $f^a(s) > v(s)$, then either $u(s) \le a$, in which case $f(s) = f^a(s) > v(s) = w(s)$ or $u(s) > a$, in which case $f(s) \ge -u(s) > w(s)$. This shows that $w \in V(Z)$, so $d(Z) \ge E(w) \ge E(v) - 2Eu' - \eta$. It follows that $d(Z') \le d(Z) + 2Eu' < d(Z) + 2\varepsilon$ and this proves the claim. We set

$$G'_n(s) = \sup_{f \in Z} \frac{1}{n} \sum_{i \le n} f^a(s_i).$$

Let $u'_n(s) = (1/n)\sum_{i \le n} u'(s_i)$ and $u_n(s) = (1/n)\sum_{i \le n} u(s_i)$. The usual law of large numbers implies that $u'_n(s) \to E(u')$ [resp. $u_n(s) \to E(u)$] a.s. and in L^1. We have $G_n(s) \le G'_n(s) + u'_n(s)$. It follows from Corollary 15 [replacing n by k, then setting $\mu = Q_n(s)$] and Lemma 18 (for $A = D$) that for each $\eta > 0$, there is a set $A \subset \Omega^\infty$ with $P^\infty(A) \ge 1 - \eta$ and m in $\mathbb{N}$ with

$$\forall\, s \in A, \forall\, n \ge m, \qquad G'_n(s) \le d(Z) + 2\varepsilon,$$

so we have

$$\forall\, n \ge m, \qquad G^*_n(s) \le 1_A(d(Z) + 2\varepsilon) + 1_{\Omega^\infty \setminus A} u_n(s) + u'_n(s).$$

It is routine to conclude that $G^*_n(t) \vee d(Z) - d(Z)$ goes to zero in L^1 and a.s.

On the other hand, for $\varepsilon > 0$ there is a measurable function v with $E(v) > d(Z) - \varepsilon$, such that $v \in V(Z)$. It follows that if $v_n(s) = (1/n)\sum_{i \le n} v(s_i)$, then $P^{\infty *}(\{G_n > v_n\}) = 1$, so $G^*_n \ge v_n$. Since v_n goes to $E(v)$ in L^1 and a.s., we have $d(Z) - d(Z) \wedge G^*_n(t) \to 0$ a.s. and in L^1, and the proof is complete.

(b) Since $(\limsup G_n)^* \le \limsup G^*_n$, we have $(\limsup G_n)^* \le d(Z)$. Let $\varepsilon > 0$. Let v be measurable with $E(v) > d(Z) - \varepsilon$, such that for each n, $P^{n*}(A_n) = 1$, where

$$A_n = A_n(Z, v) = \{(s_1, \dots, s_n) \in \Omega^n;\ \exists\, f \in Z;\ \forall\, i \le n,\ f(s_i) > v(s_i)\}.$$

Since $f \ge -u$ for f in Z, we can assume $v \ge -u - 1$. Let (n_k) be a sequence with $\lim n_k/n_{k+1} = 0$. Let $A \subset \Omega^\infty$ be given by

$$A = \left\{s \in \Omega^\infty;\ \forall\, k, (s_{n_{k-1}+1}, \dots, s_{n_k}) \in A_{n_k - n_{k-1}}\right\},$$

so $P^{\infty *}(A) = 1$. Let

$$A' = A \cap \{s \in \Omega^\infty;\ Q_n(s)(u) \to E(u)\},$$

so $P^{\infty *}(A') = 1$. We show that for $s \in A'$ we have $\limsup G_n(s) \ge d(Z) - \varepsilon$. Since $P^{\infty *}(A') = 1$ it will follow that $(\limsup G_n)^* \ge d(Z)$. So let $s \in A'$. There is $f \in Z$ with $f(s_i) > v(s_i)$ for $n_{k-1} < i \le n_k$, so we have

$$G_{n_k}(s) \ge \frac{1}{n_k}\sum_{i\le n_k} f(s_i) \ge \frac{1}{n_k}\sum_{i\le n_k} v(s_i) - \frac{1}{n_k}\sum_{i\le n_{k-1}} v(s_i) - \frac{1}{n_k}\sum_{i\le n_{k-1}} u(s_i)$$

$$\ge \frac{1}{n_k}\sum_{i\le n_k} v(s_i) - \frac{2}{n_k}\sum_{i\le n_{k-1}} (u(s_i)+1)$$

$$= \frac{1}{n_k}\sum_{i\le n_k} v(s_i) - 2\frac{n_{k-1}}{n_k}\big(Q_{n_{k-1}}(s)(u)+1\big),$$

so $\limsup_k G_{n_k}(s) \ge E(v) \ge d(Z) - \varepsilon$. The proof is complete. □

REMARK. Assume that for each measurable function v, the sets $A_n = A_n(Z, v)$ are measurable. Then it is easy to show that actually we have $\liminf_n G_n(s) = d(Z)$ P^∞ a.s. Indeed let $\varepsilon > 0$, and let v be a measurable function with $E(v) > d(Z) - \varepsilon$, and such that for each n, $P^n(A_n) = 1$. Consider $A \subset \Omega^\infty$ given by

$$A = \{s \in \Omega^\infty; \forall\, n, \exists\, f \in Z, \forall\, i \le n,\ f(s_i) > v(s_i)\}.$$

Then A is measurable and $P^\infty(A) = 1$. Let

$$A' = A \cap \{s \in \Omega^\infty; Q_n(s)(v) \to E(v)\},$$

so $P^\infty(A') = 1$. For s in A' and n in $\mathbb{N}$, there is f in Z such that $f(s_i) > v(s_i)$ for $i \le n$; so

$$G_n(s) \ge \frac{1}{n}\sum_{i\le n} f(s_i) \ge \frac{1}{n}\sum_{i\le n} v(s_i),$$

so $\liminf_n G_n(s) \ge E(v)$, and this proves the result.

It is not true in general, as Example 21 will show, that we have $(\liminf G_n)^* = d(Z)$. At the expense of considerable extra work, the next theorem will show that $(\liminf G_n)^*$ still carries information. The result is presented at this point for logical reasons, but the reader is not advised to go into this kind of extreme refinement before feeling comfortable with the simpler results.

THEOREM 20. *Let $Z \subset \mathscr{L}^1(P)$ be order bounded, such that $E(f) = 0$ for each $f \in Z$. Then if $d(Z) > 0$ we have* $(\liminf G_n)^* > 0$ *a.s.*

PROOF. *Step* 1. Let v be measurable bounded with $a = E(v) > 0$ and such that for each n, $P^{n*}(M_n) = 1$, where $M_n = A_n(Z, v)$. Let $u \in \mathscr{L}^1$ with $|f| \le u$ for each $f \in Z$ and let $b = E(u)$. Let $k \in \mathbb{N}$ with $k \ge 1 + 16b/a$ and $k \ge 2$.

Step 2. For $p \in \mathbb{N}$, $f \in Z$, let

$$F_p(f) = \Big\{(s_1,\dots,s_{k^p}) \in \Omega^{k^p}; \exists\, j \le k^p, \sum_{i\le j} f(s_i) \le -ak^{p-1}/4\Big\}.$$

We shall prove that $\sum_p \sup_{f\in Z} P^{k^p}(F_p(f)) < \infty$. Let β be large enough so that $E(u') < a/16k$, where $u' = u - u \wedge \beta$. Let

$$F_p = \Big\{(s_1,\dots,s_{k^p}) \in \Omega^{k^p}; \sum_{i\le k^p} u'(s_i) \ge ak^{p-1}/8\Big\}.$$

For $f \in Z$, let f' be the truncation of f at $-\beta$ and β. Since $|f - f'| \le u'$, we have $|E(f')| \le a/16k$.

For $(s_i, \dots, s_{k^p}) \notin F_p$, we have for $j \le k^p$,

$$\sum_{i \le j} (f'(s_i) - E(f')) \le \sum_{i \le j} f(s_i) + \sum_{i \le k^p} u'(s_i) + k^p |E(f')|$$
$$\le \sum_{i \le j} f(s_i) + ak^{p-1}/8 + ak^{p-1}/16.$$

So we have $F_p(f) \subset F_p \cup F_p'(f)$, where

$$F_p'(f) = \left\{(s_1, \dots, s_{k^p}) \in \Omega^{k^p};\ \exists\, j \le k^p, \sum_{i \le j} (f'(s_i) - E(f')) \le -ak^{p-1}/16\right\}.$$

The sequence $\sum_{i \le j}(f'(s_i) - E(f'))$ is a martingale, so

$$P^{k^p}\big(F_p'(f)\big) \le \frac{16}{ak^{p-1}} \left\| \sum_{i \le k^p} (f'(s_i) - E(f')) \right\|_1.$$

Since $|f'| \le \beta$, computation shows that

$$\left\| \sum_{i \le k^p} (f'(s_1) - E(f')) \right\|_2 \le 2\beta k^{p/2},$$

so

$$\sum_p \sup_{f \in Z} P^{k^p}\big(F_p'(f)\big) < \infty.$$

Let $l(p) = \sum_{i < p} k^i$. Let

$$F_p' = \left\{ s \in \Omega^\infty;\ \sum_{l(p) < i \le l(p+1)} u'(s_i) \ge ak^{p-1}/8 \right\}.$$

Let $s \in \limsup_p F_p'$, and p with $s \in F_p'$. We have

$$Q_{l(p+1)}(s)(u') \ge ak^{p-1}/8l(p+1) = ak^{p-1}(k-1)/8(k^{p+1}-1),$$

so

$$\limsup_n Q_n(s)(u') \ge a(k-1)/8k^2 \ge a/16k > E(u').$$

This shows that $P^\infty(\limsup_p F_p') = 0$, so $\sum P^\infty(F_p') < \infty$, so $\sum P^{k^p}(F_p) < \infty$. The step is complete.

Step 3. Let

$$B_p = \left\{ (s_i)_{i \le k^p};\ \sum_{i \le k^p} u(s_i) \ge 2bk^p \right\},$$

$$C_p = \left\{ (s_i)_{i \le k^p};\ \sum_{i \le k^p} v(s_i) \le ak^p/2 \right\}.$$

Using the fact that $k \ge 1 + 16b/a$, the method used to show that $\sum P^{k^p}(F_p) < \infty$

shows that $\Sigma P^{k^p}(B_p) < \infty$, $\Sigma P^{k^p}(C_p) < \infty$. Let

$$d_p = P^{k^p}(B_p) + P^{k^p}(C_p) + \sup_{f \in Z} P^{k^p}(F_p(f)).$$

We have $\Sigma d_p < \infty$.

Step 4. Let $A' \subset \Omega^\infty$ with $P^\infty(A') > 0$. Then we can find a decreasing sequence (A_n) of sets of Ω^∞, each depending only on finitely many coordinates, and such that if $A = \bigcap_n A_n$, we have $P^\infty(A) > 0$, $A \subset A'$. Let $q \in \mathbb{N}$ with $\Sigma_{i > q} d_i < P^\infty(A)$.

For each $p \geq 1$, let $m(p) = k^{q+p}$ and $n(p) = \Sigma_{1 \leq i \leq p} m(i)$. For each p, we write $\Omega^\infty = \Omega^{n(p)} \times T_p$, where $T_p = \Omega^{]n(p), \infty[}$. Let Q_p be the product probability on T_p, so $P^\infty = P^{n(p)} \otimes Q_p$. For each $t \in \Omega^{n(p)}$, let

$$A_p(t) = \{s \in T_p;\ (t, s) \in A\}.$$

By induction over $p \geq 1$, we construct points $(s_i)_{n(p-1) < i \leq n(p)}$ of Ω and $f_p \in Z$, such that the following hold:

$$\text{(3)} \qquad \text{If} \quad t_p = (s_i)_{i \leq n(p)}, \quad \text{then } Q_p(A_p(t_p)) > \sum_{i > p+q} d_i.$$

$$\text{(4)} \qquad \text{If} \quad s^p = (s_{n(p-1)+1}, \dots, s_{n(p)}) \in \Omega^{m(p)}, \quad \text{then} \quad s^p \notin B_{p+q} \cup C_{p+q} \cup F_{p+q}(f_{p-1}).$$

$$\text{(5)} \qquad \text{For} \quad n(p-1) < i \leq n(p), \quad \text{we have } f_p(s_i) > v(s_i).$$

The first step is almost identical to the general step, so we just perform the step to $p + 1$. Let

$$U = \Big\{s' \in \Omega^{m(p+1)};\ Q_{p+1}(A_{p+1}((t_p, s'))) > \sum_{i > p+q+1} d_i\Big\}.$$

Condition (3) and Fubini's theorem show that $P^{m(p+1)}(U) > d_{p+q+1}$. It follows that there is $s' = (s'_1, \dots, s'_{m(p+1)}) \in \Omega^{m(p+1)}$ with

$$s' \in (M_{m(p+1)} \cap U) \setminus (B_{p+q+1} \cup C_{p+q+1} \cup F_{p+q+1}(f_p)).$$

For $n(p) < i \leq n(p+1)$, we set $s_i = s'_{i-n}(p)$ and we pick f_{p+1} to satisfy (5). The construction is complete.

Step 5. Let $s = (s_i) \in \Omega^\infty$ be the sequence constructed in step 4. Condition (3) shows that for each p there exist $t_{n(p)+1}, \dots$ in Ω such that $(s_1, \dots, s_{n(p)}, t_{n(p)+1}, \dots)$ belongs to A. Since each A_n depends on finitely many coordinates, we have $s \in A_n$ for each n, so $s \in A$.

We now show that $\liminf G_n(s) \geq a(k-1)/8k^2$. Since A' was arbitrary, this will show that $(\liminf G_n)^* \geq a(k-1)/8k^2$ and will complete the proof. Let $n \in \mathbb{N}$ and p with $n(p) < n \leq n(p+1)$. We have $G_n(s) \geq n^{-1}\Sigma_{i \leq n} f_p(s_i)$. For $l < p - 1$, from (4) we have

$$\sum_{n(l) < i \leq n(l+1)} |f_p(s_i)| \leq \sum_{n(l) < i \leq n(l+1)} u(s_i) \leq 2bm(l+1),$$

and so

$$\sum_{i \leq n(p-1)} f_p(s_i) \geq -2bn(p-1).$$

From (4) and (5) we have

$$\sum_{n(p-1)<i\le n(p)} f_p(s_i) \ge \sum_{n(p-1)<i\le n(p)} v(s_i) \ge am(p)/2.$$

From (4) we get

$$\sum_{n(p)<i\le n} f_p(s_i) > -am(p)/4.$$

So, finally,

$$G_n(s) \ge [am(p)/4 - 2bn(p-1)]/n(p+1)$$

and easily $\liminf G_n(s) \ge a(k-1)/8k^2$. The proof is complete. □

EXAMPLE 21. There exists a set Z of measurable functions valued in $\{0,1\}$ with $d(Z)=1$ and $(\liminf G_n(s))^* = \frac{1}{2}$ a.s.

PROOF. Let (Ω, Σ, P) be $[0,1]$ with Lebesgue's measure, and let (Ω_n) be a partition of Ω with $P^*(\Omega_n)=1$ for each n. Let Z_n be the set of functions 1_A for $A \subset \Omega_n$ and card $A \le n$. Let $Z = \bigcup Z_n$. It is clear that $G_n^* = 1$ a.s. for each n, so $d(Z)=1$.

Let $s \in \Omega^\infty$ and $m \in \mathbb{N}$. Suppose that $G_m(s) \ge 1/2$. Then there is p such that for some $A \subset \Omega_p$, with card $A \le p$, we have

$$\operatorname{card}\{i \le m;\ s_i \in A\} \ge m/2.$$

So we have $p \ge m/2$. Let q be the largest integer $\le 2p$ such that

$$\operatorname{card}\{i \le q;\ s_i \in \Omega_p\} \ge q/2.$$

So $q \ge m$. Let $r = q+1$. If $q = 2p$, for f in Z_p, we have

$$Q_r(s)(f) \le p/(q+1) < 1/2.$$

If $q < 2p$, we have $\operatorname{card}\{i \le r;\ s_i \in \Omega_p\} < r/2$, so

$$Q_r(s)(f) < r/2r = 1/2.$$

For $n \ne p$, $f \in Z_n$, since $\Omega_n \cap \Omega_p = \varnothing$, we have

$$Q_r(s)(f) \le (r - q/2)/r \le (r+1)/2r = 1/2 + 1/2r.$$

So, since $r > q \ge m$, we have

$$\inf_{k>m} G_k(s) \le 1/2 + 1/2m,$$

so $\liminf G_r(s) \le 1/2$ for each s, and $(\liminf G_r)^* \le 1/2$. Now let

$$L = \{s \in \Omega^\infty;\ \forall\, n, i,\ n! < i \le (n+1)!,\ s_i \in \Omega_n\}.$$

Then $P^{\infty *}(L) = 1$. We leave to the reader the easy proof that $\liminf G_r(s) = 1/2$ for $s \in L$. This shows that $(\liminf G_n)^* \ge 1/2$ and finishes the proof. □

We now come to the various characterizations of Glivenko–Cantelli classes.

Theorem 22. *Let Z be a subset of $\mathscr{L}^1$. The following are equivalent.*

(I) *The quantities $|E(f)|$, for $f \in Z$, are bounded, and Z is a Glivenko–Cantelli class, that is, $\lim_n D_n(s) = 0$, P^∞ a.s.*

(II) *Z is stable and order bounded.*

(III) *Z is order bounded, and for each $\varepsilon > 0$, there exists a finite subalgebra $\mathscr{A}$ of Σ such that*

$$\limsup_n \Big(\sup_{f \in Z} Q_n(s)(|f - E(f|\mathscr{A})|) \Big)^* \le \varepsilon, \quad P^\infty \text{ a.s.}$$

(IV) *For $s \in \Omega^\infty$, there exists a number $a(s) \ge 0$, such that for $s \in \Omega^\infty$ and $f \in Z$ there exists a number $b(s, f)$, with $|b(s, f)| \le a(s)$ such that*

$$\lim_n \sup_{f \in Z} |Q_n(s)(f) - b(s, f)| = 0 \quad \text{a.s.}$$

(V) *Z is order bounded, and $\lim_n D_n^*(s) = 0$, P^∞ a.s.*

(VI) *Z is order bounded, and $\lim_n E(D_n^*(s)) = 0$.*

(VII) *Z is order bounded, and for each $\varepsilon > 0$ we have*

$$\limsup_n P_*^\infty(\{D_n < \varepsilon\}) > 0.$$

(VIII) *Z is order bounded, and $\liminf_n D_n(s) = 0$, P^∞ a.s.*

Moreover, even when Z is any set of (not necessarily measurable) functions on Ω, condition (IV) *implies that $Z \subset \mathscr{L}^1$ (and hence implies all the other conditions).*

Remarks. (1) Note that the assertion $\lim_n D_n^*(s) = 0$ a.s. is a priori much stronger than the assertion $\lim_n D_n(s) = 0$ a.s.

(2) If Z consists of the constant functions, then Z is a Glivenko–Cantelli class that is not order bounded, so some kind of boundedness is necessary in condition (I).

(3) Condition (IV) is not a far-fetched refinement, but exactly what is needed to prove Theorem 26.

Proof. It is obvious that (V) implies (I) and that (I) implies (IV). Let $Z^c = \{f - E(f);\ f \in Z\}$, $Z' = Z^c \cup (-Z^c)$. If we apply Theorem 19 to Z', we see that (V), (VI) and (VII) are equivalent, and that they are equivalent to $d(Z') = 0$. If we apply Theorem 20 to Z', we see that (VIII) is also equivalent to $d(Z') = 0$, so (V) to (VIII) are equivalent. Theorems 16(a) and (b) show that (II) implies $d(Z') = 0$ so (II) implies (V) to (VIII).

When Z is order bounded and $\mathscr{A}$ is finite,

$$\limsup_n \sup_{f \in Z} |Q_n(s)(E(f|\mathscr{A}) - E(f))| = 0, \quad P^\infty \text{ a.s.}$$

and one deduces easily that (III) implies (I). To see that (II) implies (III), one uses the same truncation argument as in the proof of Theorem 19, to deduce (III) from Theorem 17 and Lemma 18.

To finish the proof of the theorem we show that if Z is any set of functions (not necessarily measurable) that satisfies (IV), then $Z \subset \mathscr{L}^1$ and Z satisfies (I). We first show that there is $u \in \mathscr{L}^1$ such that $|f| \le u$ for each $f \in Z$. Condition (IV) implies that

$$\limsup_n \sup_{f\in Z} \left| \frac{1}{n} \sum_{i\le n+1} f(s_i) - \frac{n+1}{n} b(s,f) \right| = 0, \quad P^\infty \text{ a.s.},$$

so

$$\limsup_n \sup_{f\in Z} \left| n^{-1} f(s_{n+1}) - n^{-1} b(s,f) \right| = 0, \quad P^\infty \text{ a.s.}$$

Since $|b(s,f)| \le a(s)$, we have

$$\limsup_n \sup_{f\in Z} \left| n^{-1} b(s,f) \right| = 0, \quad P^\infty \text{ a.s.},$$

so

$$\limsup_n \sup_{f\in Z} \left| n^{-1} f(s_n) \right| = 0, \quad P^\infty \text{ a.s.}$$

Let $X_n = \{\omega \in \Omega;\ \sup_{f\in Z} |f(\omega)| \ge n\}$, and let $U_n \in \Sigma$ be a measurable cover of X_n. Let $u = 1 + \sum_{n\ge 1} 1_{U_n}$. We have $|f| \le u$ when $f \in Z$, and u is measurable. We show now that $\sum P(U_n) < \infty$. Otherwise, there is an increasing sequence (k_n) such that

$$P^\infty(\{s \in \Omega^\infty;\ \exists i,\ k_n < i \le k_{n+1};\ s_i \in U_i\}) \ge 1 - 2^{-n}.$$

Now notice that in any probability space, for sets $X_1, X_2, \dots, X_n$, if U_i is a measurable cover of X_i, $\bigcup_{i\le n} U_i$ is a measurable cover of $\bigcup_{i\le n} X_i$, so we have

$$P^{\infty *}(\{s \in \Omega^\infty;\ \exists\, i,\ k_n < i \le k_{n+1};\ s_i \in X_i\}) \ge 1 - 2^{-n}.$$

So $P^{\infty *}(M) > 0$, where

$$M = \{s \in \Omega^\infty;\ \forall\, n, \exists\, i,\ k_n < i \le k_{n+1},\ s_i \in X_i\}.$$

For $s \in M$, we have $\limsup_n \sup_{f\in Z} |n^{-1} f(s_n)| \ge 1$. This contradiction finishes the proof that $u \in \mathscr{L}^1$.

We show now that each function f in Z is measurable. Suppose, if possible, that for some f in Z we have $f^* \ne f_*$ on a set of positive measure. Then there exist two measurable functions v, w with $|v|, |w| \le u + 1$, $E(v) < E(w)$, $v > f_*$, $w < f^*$. So, if $A = \{f \ge w\}$, $B = \{f \le v\}$, we have $P^*(A) = P^*(B) = 1$. Let n_k be a sequence with $\lim n_k/n_{k-1} = \infty$. For $n_{2p} < n \le n_{2p+1}$, let $C_n = A$; for $n_{2p+1} < n \le n_{2p+2}$, let $C_n = B$. Let $C = \prod C_n$. Then $P^{\infty *}(C) = 1$. Let

$$C' = \Big\{ s \in C;\ \lim_n Q_n(s)(v) = E(v);$$
$$\lim_n Q_n(s)(w) = E(w);\ \lim_n Q_n(s)(u) = E(u) \Big\}.$$

Then $P^{\infty *}(C') = 1$. Let $t \in C'$. We show that $a_k = (1/n_k) \sum_{i\le n_k} f(t_i)$ does not converge. Suppose, if possible, that it converges to some a. We have

$$a_{2k+1} \geq \frac{n_{2k}}{n_{2k+1}} a_{2k} + \frac{1}{n_{2k+1}} \sum_{n_{2k} < i \leq n_{2k+1}} w(t_i)$$

$$\geq \frac{n_{2k}}{n_{2k+1}} a_{2k} + \frac{1}{n_{2k+1}} \sum_{i \leq n_{2k+1}} w(t_i) - \frac{1}{n_{2k+1}} \sum_{i \leq n_{2k}} (u(t_i) + 1).$$

This shows that $a \geq E(w)$. Similarly, one sees that $a \leq E(v)$, a contradiction. This proves that $f_* = f^*$ a.s., so f is measurable.

We show now that Z is stable. Otherwise, there is $A \in \Sigma$, $P(A) > 0$ and $\alpha < \beta$ with $P^{2n*}(B_{n,n}(A, \alpha, \beta)) = P(A)^{2n}$ for each n. Let $a = (\beta - \alpha)P(A)/9$. Let $b > \max(|\alpha|, |\beta|)$ be large enough that $E(u') < a$, where $u' = u - u \wedge b$. For f in Z, denote by f' its truncation at $-b$ and b. Let $Z' = \{f'; f \in Z\}$. We have $P^{2n*}(B_{n,n}(Z', A, \alpha, \beta)) = P^{2n}(A)$ for each n. Theorem 16(c) shows there exist two bounded measurable functions v, w on Ω, with $E(w - v) \geq 3a$ and $(P^{k+l})^*(M(k, l)) = 1$ for each k, l, where

$$M(k, l) = \big\{(s_1, \dots, s_k, t_1, \dots, t_l) \in \Omega^{k+l}; \exists f \in Z, \forall i \leq k,$$
$$f'(s_i) < v(s_i), \forall j \leq l, f'(t_j) \geq w(t_j)\big\}.$$

We can assume $v \leq u + 1$ and $w \geq -u - 1$. Let $n(p)$ be a sequence with $\lim n(p)/n(p+1) = 0$. Let

$$M = \big\{s \in \Omega^\infty; \forall p, (s_{n(2p)+1}, \dots, s_{n(2p+2)})$$
$$\in M(n(2p+1) - n(2p), n(2p+2) - n(2p+1))\big\},$$

so $P^{\infty *}(M) = 1$. Let

$$M' = \Big\{s \in M; \lim_n Q_n(s)(u) = E(u); \lim_n Q_n(s)(u') = E(u');$$
$$\lim_n Q_n(s)(v) = E(v); \lim_n Q_n(s)(w) = E(w)\Big\},$$

so $P^{\infty *}(M') = 1$. Fix $s \in M'$. For each p, let $f_p \in Z$ with $f_p'(s_i) < v(s_i)$ for $n(2p) < i \leq n(2p+1)$; $f_p'(s) > w(s_i)$ for $n(2p+1) < i \leq n(2p+2)$. We have

$$Q_{n(2p+1)}(s)(f_p) \leq \frac{1}{n(2p+1)} \sum_{1 \leq n(2p+1)} v(s_i) + \frac{2}{n(2p+1)} \sum_{i \leq n(2p)} (u(s_i) + 1)$$
$$+ \frac{1}{n(2p+1)} \sum_{i \leq n(2p+1)} u'(s_i),$$

$$Q_{n(2p+2)}(s)(f_p) \geq \frac{1}{n(2p+2)} \sum_{i \leq n(2p+2)} w(s_i) - \frac{2}{n(2p+2)} \sum_{i \leq n(2p+1)} (u(s_i) + 1)$$
$$- \frac{1}{n(2p+2)} \sum_{i \leq n(2p+2)} u'(s_i),$$

so

$$\begin{aligned} 2 \limsup_n \sup_{f \in Z} &|Q_n(s)(f) - b(s, f)| \\ &\geq \limsup_p \left(|Q_{n(2p+1)}(s)(f_p) - b(s, f_p)| + |Q_{n(2p+2)}(s)(f_p) - b(s, f_p)| \right) \\ &\geq \limsup_p |Q_{n(2p+2)}(s)(f_p) - Q_{n(2p+1)}(s)(f_p)| \\ &\geq E(w) - E(v) - 2a > 0. \end{aligned}$$

This contradiction shows that Z is stable. The result is proved. □

Proof of Theorem 7. (a) We actually show that for each $\varepsilon > 0$, $(1/n)\log N_n^\infty{}^*(Z, \varepsilon, s) \to 0$ a.s. and in L^1.

Step 1. We assume that Z is bounded, so we can suppose $Z \subset [-1/2, 1/2]^\Omega$. Since $(1/n)\log N_n^\infty(Z, \varepsilon, s) \leq \log(1 + 1/\varepsilon)$ it is enough to show that $(1/n)\log N_n^\infty{}^*(Z, \varepsilon, s) \to 0$ a.s. Let $\beta > 0$. Let $0 < \alpha < \varepsilon$ be such that

$$\alpha \log(1 + 2/\varepsilon) + \alpha|\log \alpha| + (1 - \alpha)|\log(1 - \alpha)| < \beta. \tag{6}$$

From Theorem 22, there is a finite subalgebra $\mathscr{A}$ of Σ such that

$$\limsup_n \left(\sup_{f \in Z} Q_n(s)(|f - E(f|\mathscr{A})|) \right)^* \leq \alpha\varepsilon/2 \quad \text{a.s.}$$

Let $Z_1 = \{E(f|\mathscr{A});\ f \in Z\}$; $Z_2 = \{f - E(f|\mathscr{A});\ f \in Z\}$. We have

$$N_n^\infty(Z, 2\varepsilon, s) \leq N_n^\infty(Z_1, \varepsilon, s) N_n^\infty(Z_2, \varepsilon, s).$$

Since $\mathscr{A}$ is finite, we have $(1/n)\log(N_n^\infty(Z_1, \varepsilon, s)) \to 0$ a.s. so it is enough to show that

$$\limsup_n \frac{1}{n} \log(N_n^\infty{}^*(Z_2, \varepsilon, s)) \leq \beta \quad \text{a.s.} \tag{7}$$

Note that $Z_2 \subset [-1, 1]^\Omega$. Let

$$A_n = \left\{ s \in \Omega^\infty,\ \sup_{f \in Z} Q_n(s)(|f - E(f|\mathscr{A})|) \leq \alpha\varepsilon \right\},$$

so $P^\infty(\liminf_n A_n) = 1$. For $s \in A_n$, for each f in Z_2, we have $Q_n(s)(\{|f| > \varepsilon\}) < \alpha$. Let $m = [\alpha n]$.

Consider the set of functions that are zero everywhere, except at most at m of the points $s_1, \dots, s_n$, where their values are of the type $k\varepsilon$ ($k \in \mathbb{Z}$) with $|k\varepsilon| \leq 1$. For $s \in A_n$ each function of Z_2 is at distance $\leq \varepsilon$ from this set; so

$$N_n^\infty(Z_2, \varepsilon, s) \leq \binom{n}{m}(1 + 2/\varepsilon)^m,$$

so

$$\log N_n^\infty{}^*(Z_2, \varepsilon, s) \leq \log\binom{n}{m} + n\alpha \log(1 + 2/\varepsilon).$$

Computation using Stirling's formula shows that

$$\lim_n \frac{1}{n} \log\binom{n}{m} = \alpha|\log \alpha| + (1 - \alpha)|\log(1 - \alpha)|,$$

so (7) follows from (6).

Step 2. Let $\varepsilon > 0$ and $\beta > 0$. Let u be such that $|f| \le u$ for $f \in Z$ and $\log(u+1) \in \mathscr{L}^1$. Let b be large enough that

$$P(\{u > b\})|\log \varepsilon| \le \beta/2, \qquad E(1_{\{u \ge b\}}\log u) \le \beta/2.$$

For $f \in Z$ let f' be the truncation of f at $-b$ and b. Let

$$Z_1 = \{f';\ f \in Z\};\ Z_2 = \{f - f';\ f \in Z\}.$$

We have

$$N_n^\infty(Z, 2\varepsilon, s) \le N_n^\infty(Z_1, \varepsilon, s) N_n^\infty(Z_2, \varepsilon, s).$$

Step 1 shows that $\limsup_n (1/n)\log(N_n^{\infty *}(Z_1, \varepsilon, s)) = 0$, so it is enough to prove that

$$\beta \vee \frac{1}{n}\log N_n^{\infty *}(Z_2, \varepsilon, s) - \beta \to 0 \quad \text{a.s. and in } L^1.$$

For $t \in \Omega$, we set $v(t) = 1$ for $u(t) \le b + \varepsilon$, $v(t) = 1 + (u(t) - b)/\varepsilon$ otherwise. We see that

$$N_n^{\infty *}(Z_2, \varepsilon, s) \le \prod_{i \le n} v(s_i).$$

Since for $u(t) \ge b + \varepsilon$ we have $v(t) \le u(t)/\varepsilon$, we get

$$E(\log v) \le E(1_{\{u \ge b\}}\log u) + |\log \varepsilon| P(\{u \ge b + \varepsilon\}) \le \beta$$

and the result follows from the law of large numbers.

(b) Suppose first that Z is bounded, and let $\varepsilon > 0$. From Theorem 22, there is a finite subalgebra $\mathscr{A}$ of Σ such that

$$\limsup_n \sup_{f \in Z} Q_n(s)(|f - E(f|\mathscr{A})|) \le \varepsilon/3 \quad \text{a.s.}$$

Let $Z_1 = \{E(f|\mathscr{A});\ f \in Z\}$. If

$$\sup_{f \in Z} Q_n(s)(|f - E(f|\mathscr{A})|) \le 2\varepsilon/3,$$

we have

$$N_n^1(Z, \varepsilon, s) \le N_n^1(Z_1, \varepsilon/3, s),$$

so we have

$$\limsup_n N_n^1(Z, \varepsilon, s) \le \limsup_n N_n^1(Z_1, \varepsilon/3, s) \quad \text{a.s.}$$

and the result follows since $\mathscr{A}$ is finite.

The case where Z is order bounded reduces to the case where Z is bounded using the usual truncation argument.

(c) If Z is not stable but satisfies condition (M), Proposition 4 shows that there is $A \in \Sigma$, with $P(A) > 0$ and $\alpha < \beta$ such that for each m, $P^m(B_m) = P(A)^m$, where

$$B_m = \{(s_1, \dots, s_m) \in A^m;\ \forall I \subset \{1, \dots, m\}, \exists f \in Z, \forall i \in I,\ f(s_i) < \alpha,$$
$$\forall i \notin I,\ f(s_i) > \beta\}.$$

For $s \in \Omega^\infty$, let $m_n(s) = \operatorname{card}\{i \le n;\ s_i \in A\}$. If $i(1), \dots, i(m_n(s))$ are the indices

$i \le n$ such that $s_i \in A$, let $s(n) = (s_{i(1)}, \ldots, s_{i(m_n(s))})$. Then it is clear that $P^\infty(C) = 1$, where

$$C = \left\{ s \in \Omega^\infty;\ \forall\ n,\ s(n) \in B_{m_n(s)} \right\}.$$

We now observe that in a set I of k elements, one can find subsets $D_1, \ldots, D_l$, with $l \ge e^{k/8}$ and $\operatorname{card}(D_i \vartriangle D_j) \ge k/4$ for $i \ne j$. (This follows, e.g., from a random choice and Hoeffding's inequality.) It follows that for s in C, if $m_n(s) \ge nP(A)/2$ and $\varepsilon = (\beta - \alpha)P(A)/16$, then $N_n^1(Z, \varepsilon, s) \ge e^{m_n(s)/8}$ so

$$\frac{1}{n}\log N_n^1(Z, \varepsilon, s) \ge \frac{m_n(s)}{8n}.$$

The law of large numbers shows that $\lim_n m_n(s)/n = P(A)$ a.e., so

$$\liminf \frac{1}{n}\log N_n^1(Z, \varepsilon, s) \ge P(A)/8 \quad \text{a.e.}$$

This concludes the proof. □

We now give an example to show that Theorem 7(c) fails when no measurability assumption is made.

EXAMPLE 23. Let (Ω, Σ, P) be $[0, 1]$ with Lebesgue's measure. There exists a set Z of $\{0, 1\}$-valued measurable functions that is not stable, but such that for each $\varepsilon > 0$, $\int (1/n)\log N_n^{\infty *}(Z, \varepsilon, s)\, dP^\infty(s) \to 0$; for each n, s, $N_n^1(Z, \varepsilon, s) \le 1/\varepsilon$.

PROOF. It is routine to construct a disjoint family $\mathscr{A}$ of finite sets such that if $Z = \{1_A;\ A \in \mathscr{A}\}$, then $G_n^*(s) = 1$ a.s. for each n. However, for each n, each ε, and each s, one has $N_n^\infty(Z, \varepsilon, s) \le n + 1$ and $N_n^1(Z, \varepsilon, s) \le 1/\varepsilon$. The result follows. □

This example also shows that condition (M) cannot be dropped from Theorem 5.

We conclude this section with two stability properties of Glivenko–Cantelli classes.

PROPOSITION 24. *Let $Z_1, \ldots, Z_n$ be uniformly bounded stable sets. Let θ: $\mathbb{R}^n \to \mathbb{R}$ be a continuous function. Then*

$$Z_\theta = \{\theta(f_1, \ldots, f_n);\ \forall\ i \le n,\ f_i \in Z_i\}$$

is stable.

PROOF. *Observation* 1: If Z_1 is stable, and $[a, b]$ contains the range of each function of Z_1, for each homeomorphism ζ of $[a, b]$ onto its image, $\zeta(Z_1) = \{\zeta \circ f;\ f \in Z_1\}$ is stable; this is obvious from the definition of stability.

Observation 2: If Z_1, Z_2 are stable and uniformly bounded, $Z_1 + Z_2 = \{f_1 + f_2;\ f_1 \in Z_1,\ f_2 \in Z_2\}$ is stable. This is also obvious since Z_1, Z_2 are Glivenko–Cantelli classes and so $Z_1 + Z_2$ is also a Glivenko–Cantelli class. (It is however not obvious from the definition of stability!)

The proof will follow from repeated application of these observations. We first show that

$$Z_1Z_2 = \{ f_1 f_2;\ f_1 \in Z_1,\ f_2 \in Z_2\}$$

is stable. Let a be large enough that $f + a \geq 1$ for $f \in Z_1 \cup Z_2$. Observation 1 shows that $\log(Z_1 + a)$ and $\log(Z_2 + a)$ are stable. Observation 2 shows that $\log(Z_1 + a) + \log(Z_2 + a)$ is stable. Observation 1 again shows that $Z_3 = \exp(\log(Z_1 + a) + \log(Z_2 + a))$ is stable. We have

$$Z_3 = \{(f_1 + a)(f_2 + a);\ f_1 \in Z_1,\ f_2 \in Z_2\}.$$

Observation 2 shows that

$$Z_4 = Z_3 - aZ_1 - aZ_2 - a^2$$

is stable. We have

$$Z_4 = \{(f_1 + a)(f_2 + a) - af_3 - af_4 - a^2;\ f_1, f_3 \in Z_1,\ f_2, f_4 \in Z_2\}.$$

Since this class contains Z_1Z_2, the claim is proved.

It follows that if $\mathscr{C}$ is the class of functions θ for which Z_θ is stable, $\mathscr{C}$ is an algebra. This algebra contains the coordinate functions. A routine argument shows that it is closed for uniform convergence on compact sets. The result follows. □

Our next result is a comparison principle for processes. (In a further work [10], we prove a similar principle for the Central Limit Theorem.)

PROPOSITION 25. *Let X be any set. Consider two bounded functions ϕ, ψ: $X \times \Omega \to \mathbb{R}$. Assume that $\phi(x, \cdot)$ is measurable on Ω for each x in X and*

$$\forall\, x, y \in X, \forall\, s \in \Omega, |\phi(x, s) - \phi(y, s)| \leq |\psi(x, s) - \psi(y, s)|.$$

If

$$Z_\psi = \{s \to \psi(x, s);\ s \in X\}$$

is a stable set, then

$$Z_\phi = \{s \to \phi(x, s);\ x \in X\}$$

is also a stable set.

REMARK. (1) This is of course a statement about the law of large numbers, since the hypothesis that Z_ψ is stable is equivalent to saying that the process ψ satisfies the law of large numbers in the following sense:

$$\limsup_n \sup_{x \in X} \left(\left| n^{-1} \sum_{i \leq n} \psi(x, s_i) - E(\psi(x, s)) \right| \right)^* = 0, \quad P^\infty \text{ a.s.}$$

(2) The hypothesis of boundedness can be relaxed with the usual truncation argument to $|\phi(x, s)| \leq u(s)$, $|\psi(x, s)| \leq u(s)$ for each x, where $u \in \mathscr{L}^1$.

PROOF. For $x \in X$, let ψ_x (resp. ϕ_x) be given by $\psi_x(s) = \psi(x, s)$ for $s \in \Omega$ [resp. $\phi_x(s) = \phi(x, s)$]. Let $\varepsilon > 0$. Apply Theorem 22 to Z_ψ, and let $\mathscr{A}$ be a finite algebra as in the condition (III) of this theorem, that is,

$$\limsup_n \Big(\sup_{x \in X} Q_n(s)\big(|\psi_x - E(\psi_x|\mathscr{A})|\big)\Big)^* \le \varepsilon, \quad P^\infty \text{ a.s.}$$

For $x, y \in X$, let

$$d(x, y) = \big\| E(\psi_x|\mathscr{A}) - E(\psi_y|\mathscr{A}) \big\|_1.$$

We note that since $\mathscr{A}$ is finite, for $y \in X$ we have by the usual law of large numbers

$$\limsup_n \sup_{x \in X;\, d(x, y) \le \varepsilon} Q_n(s)\big(\big|E(\psi_x|\mathscr{A}) - E(\psi_y|\mathscr{A})\big|\big) \le \varepsilon, \quad P^\infty \text{ a.s.}$$

Let F be a finite subset of X such that

$$\forall\, x \in X, \exists\, y \in F, \qquad d(x, y) < \varepsilon.$$

Given $y \in F$, we have

$$\limsup_n \Big(\sup_{d(x, y) \le \varepsilon} Q_n(s)\big(|\psi_x - \psi_y|\big)\Big)^* \le 3\varepsilon, \quad P^\infty \text{ a.s.}$$

Since $|\phi_x - \phi_y| \le |\psi_x - \psi_y|$ by hypothesis, we get

$$\limsup_n \Big(\sup_{d(x, y) \le \varepsilon} Q_n(s)\big(|\phi_x - \phi_y|\big)\Big)^* \le 3\varepsilon, \quad P^\infty \text{ a.s.}$$

Since

$$d(x, y) \le \varepsilon \Rightarrow \big|E(\phi_x) - E(\phi_y)\big| \le \varepsilon,$$

we get for each y in F

$$\limsup_n \Big(\sup_{d(x, y) < \varepsilon} \big|Q_n(s)(\phi_x) - E(\phi_x) - Q_n(s)(\phi_y) + E(\phi_y)\big|\Big)^* \le 4\varepsilon, \quad P^\infty \text{ a.s.,}$$

so since F is finite,

$$\limsup_n \Big(\sup_{x \in X} \big|Q_n(s)(\phi_x) - E(\phi_x)\big|\Big)^* \le 4\varepsilon, \quad P^\infty \text{ a.s.}$$

The result follows by Theorem 22 again.

4. The law of large numbers in a Banach space. Let E be a Banach space, and E_1^* be the unit ball of E^*. Consider a map $\phi\colon \Omega \to E$ (no measurability is assumed). We define the (extended) L^1-norm of ϕ by

$$\|\phi\|_1^* = \int^* \|\phi\|\, dP.$$

We define the Glivenko–Cantelli norm of ϕ by

$$\|\phi\|_{\mathrm{GC}} = \limsup_n \int^* g_n \, dP^\infty,$$

where

$$g_n(s) = \sup_{x^* \in E_1^*} \frac{1}{n} \sum_{i \le n} |x^*(\phi(s_i))|.$$

It is useful to note that, by subadditivity, $\|\phi\|_{GC} = \inf_n \int^* g_n \, dP^\infty$. Since $g_n(s) \leq (1/n)\Sigma_{i \leq n}\|\phi(s_i)\|$, we have $\int^* g_n \, dP^\infty \leq \|\phi\|_1^*$, so $\|\phi\|_{GC} \leq \|\phi\|_1^*$. We define the Pettis norm of ϕ by

$$\|\phi\|_P = \sup_{x^* \in E_1^*} \int^* |x^* \circ \phi| \, dP.$$

For $x^* \in E_1^*$, we have $g_n(s) \geq (1/n)\Sigma_{i \leq n}|x^* \circ \phi|(s_i)$, so

$$\int^* g_n \, dP^\infty \geq \int^* |x^* \circ \phi| \, dP, \quad \text{so} \quad \|\phi\|_P \leq \|\phi\|_{GC}.$$

($\|\cdot\|_1^*$, $\|\cdot\|_{GC}$ and $\|\cdot\|_P$ are actually only seminorms.)

Let

$$Z_\phi = \{x^* \circ \phi;\ x^* \in E_1^*\}.$$

We say that ϕ is *properly measurable* if Z_ϕ is stable. We note that $\|\phi\|_1^* < \infty$ if and only if there is $u \in \mathscr{L}^1$ with $\|\phi\| = \sup_{x^* \in E^*}|x^* \circ \phi| \leq u$ everywhere. In other words, $\|\phi\|_1^* < \infty$ if and only if Z_ϕ is order bounded.

The following characterizes the maps that satisfy the strong law of large numbers.

THEOREM 26. *Given a map ϕ: $\Omega \to E$, and $s \in \Omega^\infty$, set $S_n(s) = \Sigma_{i \leq n}\phi(s_i)$. The following are equivalent*:

(a) *For almost each $s \in \Omega^\infty$, the sequence $(1/n)S_n(s)$ converges in norm.*
(b) *There exists $a \in E$ such that $(1/n)S_n(s) \to a$ in norm a.s.*
(c) *ϕ is properly measurable, and $\|\phi\|_1^* < \infty$.*
(d) *For each $\varepsilon > 0$, there is a step function ψ: $\Omega \to E$ such that $\|\phi - \psi\|_{GC} \leq \varepsilon$.*

REMARK. When these conditions hold the one-dimensional law of large numbers shows that for x^* in E^*, $E(x^* \circ \phi) = x^*(a)$. So a is the integral of ϕ in a weak sense.

PROOF. (a) ⇒ (c). Let $\Omega_0 \subset \Omega^\infty$ with $P^\infty(\Omega_0) = 1$ be such that for $s \in \Omega_0$, the sequence $S_n(s)/n$ converges in norm to some $x(s)$ in E. Let $a(s) = \|x(s)\|$. For $x^* \in E_1^*$, let $b(x^*, s) = x^*(x(s))$, so we have $|b(x^*, s)| \leq a(s)$. Since $S_n(s)/n$ converges to $x(s)$, we have

$$\lim_n \sup_{x^* \in E_1^*} \left| \frac{1}{n} x^*(S_n) - x^*(x(s)) \right| = 0 \quad \text{a.s.}$$

The last assertion of Theorem 22 shows that $x^* \circ \phi$ is measurable for each $x^* \in E_1^*$. (In other words, ϕ is scalarly measurable.) So Z_ϕ is a set of measurable functions that satisfies condition (IV) and hence all the conditions of Theorem 22. In particular this means that (c) holds.

(c) ⇒ (d). Using Theorem 22, we see that there is a finite subalgebra $\mathscr{A}$ of Σ such that

$$\text{(8)} \quad \limsup_n \int \sup_{x^* \in E_1^*} \left[\frac{1}{n} \sum_{i \leq n} |x^* \circ \phi(s_i) - E(x^* \circ \phi|\mathscr{A})(s_i)| \right] dP^\infty(s) \leq \varepsilon.$$

Since ϕ is properly measurable and $\|\phi\|_1^* < \infty$, [9], Theorem (6-1-2), shows that ϕ is Pettis integrable; so in particular for each atom A of $\mathscr{A}$ there exists $x_A \in E$ with

$$\forall x^* \in E^*, \qquad x^*(x_A) = \int_A x^* \circ \phi \, dP.$$

Define ψ: $\Omega \to E$ by $\psi = \Sigma x_A 1_A$, so $E(x^* \circ \phi | A) = x^* \circ \psi$ for each $x^* \in E^*$. Now (8) means that $\|\phi - \psi\|_{\mathrm{GC}} \le \varepsilon$.

(d) $\Rightarrow$ (b). For each k, let ψ_k be a step function such that $\|\phi - \psi_k\|_{\mathrm{GC}} < 2^{-k}$. Then

$$\limsup_n \left\| \frac{1}{n} \sum_{i \le n} (\phi(s_i) - \psi_k(s_i)) \right\| \le 2^{-k}, \quad P^\infty \text{ a.s.}$$

This is a consequence of Theorem 19(a) (or, much more easily of a simple subadditivity argument). So, since ψ_k takes only finitely many values, we get

$$\limsup_n \left\| \frac{1}{n} \sum_{i \le n} \psi_k(s_i) - E(\psi_k) \right\| = 0, \quad P^\infty \text{ a.s.,}$$

so

$$\limsup_n \left\| \frac{1}{n} S_n(s) - E(\psi_k) \right\| \le 2^{-k}, \quad P^\infty \text{ a.s.}$$

This implies $\|E(\psi_k) - E(\psi_{k+1})\| \le 2^{-k+1}$, so the sequence $(E(\psi_k))$ converges in norm to some $a \in E$, and it is routine to check (b).

(b) $\Rightarrow$ (a) is obvious. □

COROLLARY 27. *If ϕ is properly measurable and $\|\phi\|_1^* < \infty$, then $\|\phi\|_{\mathrm{GC}} = \|\phi\|_P$.*

PROOF. This holds when ϕ is finitely valued. In general, Theorem 28(d) shows that ϕ is the limit in $\|\cdot\|_{\mathrm{GC}}$ (and hence also in $\|\cdot\|_P$) of a sequence of finitely valued functions. □

The notion of properly measurable function has other interesting aspects. In particular, it seems to be the right class for Pettis integration. This aspect is expanded on in [9].

Acknowledgments. The author is indebted to P. Gaenssler and J. Hoffmann-Jørgensen for reading early versions of this work and for many valuable suggestions. Thanks are also due to R. M. Dudley for stimulating correspondence. Special thanks are due to anonymous referees whose infinite patience and care made publication of this paper possible.

REFERENCES

[1] DUDLEY, R. M (1984). A course on empirical processes. *Ecole d'Été de Probabilités de Saint-Flour XII—1982. Lecture Notes in Math.* **1097** 1–142. Springer, Berlin.

[2] ELKER, J. Über ein gleichmässiges Gesetz der grossen Zahlen. Diploma Thesis, Ruhr Univ. of Bochum.

[3] GAENSSLER, P. and STUTE, W. (1979). Empirical processes: a survey of results for independent and identically distributed random variables. *Ann. Probab.* **7** 193–243.

[4] GINÉ, E. and ZINN, J. (1984). Some limit theorems for empirical measures (with discussion). *Ann. Probab.* **12** 929–989.

[5] HOFFMAN-JØRGENSEN, J. (1985). The law of large numbers for non measurable and non separable elements. *Astérisque* **131** 299–356.

[6] RANGA-RAO, R. (1962). Relations between weak and uniform convergence of measures with applications. *Ann. Math. Statist.* **33** 659–680.

[7] STEELE, J. M. (1978). Empirical discrepancies and subadditive processes. *Ann. Probab.* **6** 118–127.

[8] TALAGRAND, M. (1982). Closed convex hull of set of measurable functions, Riemann-measurable functions and measurability of translations. *Ann. Inst. Fourier (Grenoble)* **32** 39–69.

[9] TALAGRAND, M. (1984). Pettis integral and measure theory. *Mem. Amer. Math. Soc.* **51** (307).

[10] TALAGRAND, M. (1987). Measurability problems for empirical processes. *Ann. Probab.* **15** 204–212.

[11] VAPNIK, V. N. and ČERVONENKIS, A. JA. (1981). Necessary and sufficient conditions for the uniform convergence of means to their expectations. *Theory Probab. Appl.* **26** 532–553.

EQUIPE D'ANALYSE-TOUR 46
UNIVERSITÉ PARIS VI
4 PLACE JUSSIEU
75230 PARIS CEDEX 05
FRANCE

DEPARTMENT OF MATHEMATICS
THE OHIO STATE UNIVERSITY
231 WEST 18TH AVENUE
COLUMBUS, OHIO 43210