WEAK-STAR CONVERGENCE IN THE DUAL OF
THE CONTINUOUS FUNCTIONS ON THE n-CUBE, 1 ≤ n ≤ ∞

BY

RICHARD B. DARST AND ZORABI HONARGOHAR

ABSTRACT. Let n be a positive integer and let

J = \times_{i=1}^{n}[0,1] denote the n-cube. Let

C = C(J) denote the (sup norm) space of continuous (real-valued) functions defined
on J, and let \mathcal{M} denote the (variation norm) space of (real-valued) signed Borel
measures defined on the Borel subsets of J. Let \{\mu_i\} be a sequence of elements of
\mathcal{M}. Necessary and sufficient conditions are given in order that \lim_{i} f d\mu_i exists for
every f \in C. After considering a finite dimensional case, the infinite dimensional
case is entertained.

I. Introduction to the finite dimensional case. We begin with a brief unchronologi-
cal orientation. One of the Riesz representation theorems establishes an isometric
isomorphism between \mathcal{M} and the dual, C*, of C: \mu \in \mathcal{M} which corresponds to
L \in C* via the equation L(f) = \int_J f d\mu = \int f d\mu, f \in C.

For x = (x^1, \ldots, x^n), y \in \mathbb{R}^n, x < y means x^i < y^i, i \leq j \leq n. For \bar{0} = (0, \ldots, 0)
\leq x_1 \leq x_2 \leq 1 = (1, \ldots, 1), the closed subinterval \[x_1, x_2\] = \{x; x_1 \leq x \leq x_2\}. The
distribution function \Gamma of \mu \in \mathcal{M} is defined on J by \Gamma(x) = \mu(\bar{0}, x) for x > 0 and
\Gamma(x) = 0 otherwise. The variation norm, \|\Gamma\|, of \Gamma is equal to \|\mu\| and \int f d\mu = \int f d\Gamma,
where the latter integral is a Riemann-Stieltjes integral. Let \Gamma denote the space of
distribution functions.

Let \{\mu_i\} be a sequence in \mathcal{M} and suppose that \lim_{i} f d\mu_i exists, f \in C. Let L(f)
denote this limit. Then L \in C* and \mu_i \overset{w^*}{\rightharpoonup} \mu, i.e., \lim_{i} \int f d(\mu_i - \mu) = 0, f \in C. Thus, it
suffices to consider weak-star convergence to zero.

Let S denote the set of all proper subsets of \{1, 2, \ldots, n\}, and for \theta \in S, let
J_\theta = \{x \in J; x^j = 1, j \in \theta\}.

Let \nu denote Lebesgue measure on J, and let \nu_\theta denote m-dimensional Lebesgue
measure on J_\theta, where \theta \in S, |\theta| is the number of elements in \theta and
m = |\nu_\theta| = n - |\theta|.

Conditions for weak-star convergence to zero follow.

THEOREM 1. Let \{\mu_i\} be a sequence in \mathcal{M} and let \{\Gamma_i\} be the corresponding sequence
of distribution functions. Then, \mu_i \overset{w^*}{\rightharpoonup} 0 if and only if the following three conditions are
met:

(i) \|\mu_i\| \leq M for some M and all i;
(ii) \mu_i(J) \to 0, as i \to \infty;
(iii) \forall \theta \in S: \int_{J_\theta} |\Gamma_i| d\nu_\theta \to 0 as i \to \infty.
For instance, when \( n = 2 \), \( S = \{ \phi, \{1\}, \{2\} \} \) and (iii) is the union of three statements:
\[
\int_0^1 \int_0^1 |\Gamma(x, y)| \, dx \, dy \to 0, \quad \int_0^1 |\Gamma(1, y)| \, dy \to 0 \quad \text{and} \quad \int_0^1 |\Gamma(x, 1)| \, dx \to 0.
\]

Notice that the necessity of (i) is a consequence of the uniform boundedness theorem and (ii) merely says that \( \mu(J) = \int_1 \, d\mu_t \to 0 \), so (iii) is the crucial condition. G. Högnäs [4] considered the case \( n = 1 \); however, our approach is quite different. Our proof of Theorem 1 is based on two Helly theorems and a technical result (Theorem 2). Before establishing Theorem 1, we will go back to the one-dimensional case and recall several facts to motivate our strategy. Theorem 2 is a multidimensional version of some of these facts, so the one-dimensional version of Theorem 1 turns out to be a consequence of two Helly theorems and the Lebesgue Dominated Convergence Theorem. In this section we give a proof of Theorem 1. §I contains a discussion of \( n \)-dimensional Riemann-Stieltjes integration, including proofs of the Helly theorems, for the interested reader. We give a proof of Theorem 2 in §III. §IV is a brief discussion of the infinite dimensional case.

A subinterval \([x_1, x_2]\) of \( J \) has \( 2^n \) corners, namely
\[
(x_{i_1}^1, x_{i_2}^2, \ldots, x_{i_n}^n); \quad i_j = 1, 2 \quad \text{for each} \quad j = 1, 2, \ldots, n.
\]

We define
\[
\gamma_{i_1, i_2, \ldots, i_n} = \text{sign}\left([x_{i_1}^1, x_{i_2}^2, \ldots, x_{i_n}^n]\right) = \begin{cases} + & \text{if } \sum_{j=1}^n i_j \text{ is even}, \\ - & \text{if } \sum_{j=1}^n i_j \text{ is odd}. \end{cases}
\]

For a function \( g \) we define
\[
\Delta g(I) = \sum_{i_1, \ldots, i_n} \gamma_{i_1, i_2, \ldots, i_n} g\left(x_{i_1}^1, \ldots, x_{i_n}^n\right).
\]

In accordance with our usage of superscripts, a partition \( \sigma_j \) of \([0, 1]\), \( 1 \leq j \leq n \), will be given by
\[
\sigma_j: 0 = x_0^j < x_1^j < \cdots < x_{m_j}^j = 1.
\]

By a partition \( \sigma = \times_{j=1}^n \sigma_j \) of \( J \), we mean the set of subintervals
\[
I_{i_1, \ldots, i_n} = \times_{j=1}^n \left[x_{i_j}^{j-1}, x_{i_j}^j\right]; \quad 1 \leq i_j \leq m_j, \quad j = 1, 2, \ldots, n.
\]

In case all the partitions \( \sigma_j \) are disjoint, i.e.,
\[
\sigma_j = \left\{ [0, x_1^j], (x_1^j, x_2^j], \ldots, (x_{m_j-1}^j, x_{m_j}^j]\right\},
\]
\( \sigma \) is referred to as a disjoint partition.

A function \( g \) is said to be of bounded variation (on \( J \)) if and only if
\[
\|g\| = \sup_{\sigma} \sum_{I \in \sigma} |\Delta g(I)| < \infty,
\]

where \( \sigma \) ranges over all partitions of \( J \); \( \|g\| \) is called the total variation of \( g \) (on \( J \)).
A point \( x \in J \) is said to lie on coordinate planes if \( x^j = 0 \) for at least one \( j = 1, 2, \ldots, n \). Let \( B \) denote the set of all functions of bounded variation vanishing at all the points lying on the coordinate planes.

Using the linearity property of the integral, to every \( g \in B \) we can associate an \( L \in C^* \) by defining \( L(f) = \int f \, dg \), \( f \in C \), and hence a \( \mu = \mu_g \in \mathcal{M} \) such that

\[
\int f \, dg = L(f) = \int f \, d\mu = \int f \, d\Gamma_{\mu}, \quad f \in C.
\]

The relations \( \|L\| = \|\mu\| = \|\Gamma_{\mu}\| \leq \|g\| \) obtain and the equivalence classes \( \{h \in B; \int f \, dh = \int f \, dg, f \in C\} \) comprise a partition of \( B \). With this in mind, given a sequence \( \langle \mu_i \rangle \) in \( \mathcal{M} \), by a corresponding sequence \( \langle g_i \rangle \) in \( B \) we mean any sequence such that (1-1) holds for each pair \( \mu_i, g_i \), for \( l = 1, 2, \ldots \). Let \( \langle g_i \rangle \) be a sequence in \( B \) and \( g \in B \). Pointwise convergence on \( J \) is denoted \( g_i \to g \), and weak-star convergence by \( g_i \rightharpoonup g \), where the latter is defined as usual:

\[
\int f \, dg_i \to \int f \, dg, \quad f \in C.
\]

Put \( n = 1 \) and recall the following facts: (i) a function \( g \) in \( B \) is the difference of two nondecreasing functions \( p \) and \( q \) in \( B \); moreover, \( p \) and \( q \) can be so chosen that \( \|g\| = p(1) + q(1) \); (ii) a distribution function is right continuous on \((0, 1)\); (iii) if a uniformly bounded sequence of distribution functions converges pointwise to a function \( g \), then \( g \in B \), but \( g \) need not be in \( \Gamma \); (iv) a function of bounded variation has only a countable number of points of discontinuity; (v) a function of bounded variation has left side limits on \((0, 1)\) and right side limits on \([0, 1)\); (vi) a function in \( B \) corresponds to the zero functional \( \leftrightarrow \) it is zero at one and is zero a.e.; (vii) a function \( g \) in \( B \) is a distribution function \( \leftrightarrow \) it is right continuous on \((0, 1)\).

When \( n > 1 \), things are more complicated; however, basic facts tend to be quite similar. The case \( n = 2 \) is a nice case to consider in order to see what is happening in the sequel: the resulting spaces are flexible enough to display the types of things that can occur, it is easy to draw pictures and there is a rather complete treatment of two-dimensional Riemann-Stieltjes integration in [5]. To illustrate, suppose that \( n = 2 \) and \( L \in C^* \) is defined by \( L(f) = f(\frac{1}{2}, \frac{1}{2}) \). Then \( \mu \) corresponds to a unit mass at the point \((\frac{1}{2}, \frac{1}{2})\), but \( \Gamma \) is discontinuous on \( \{((\frac{1}{2}, y); \frac{1}{2} \leq y \leq 1\} \cup \{(x, \frac{1}{2}); \frac{1}{2} \leq x \leq 1\}; \) however, if we think of a point as a hyperplane in \( \mathbb{R}^1 \), then (iv) says that the discontinuities of elements of \( B \) lie on a countable set of hyperplanes and this is a valid statement for \( 1 \leq n < \infty \). Theorem 2 following the Helly theorems below is an \( n \)-dimensional version of (vi).

[\( H_1 \)] Let \( \langle g_l \rangle \) be a sequence in \( B \) with \( \|g_l\| \leq M, l = 1, 2, \ldots \). Then there exists a subsequence \( \langle g_{l_k} \rangle \) such that \( g_{l_k} \to g \) and \( \|g\| \leq M \).

[\( H_2 \)] Let \( \langle g_l \rangle \) be a sequence in \( B \) with \( \|g_l\| \leq M, l = 1, 2, \ldots \). If \( g_l \to g \), then \( g \in B \) and \( g_n \rightharpoonup g \).

For \( \theta \in S \) and \( h \) a function on \( J \), \( h_\theta \) denotes the restriction of \( h \) to \( J_\theta \).

**THEOREM 2.** Let \( g \in B \). Then the following statements are equivalent:

(i) \( \int f \, dg = 0, f \in C \).

(ii) \( g(1) = 0 \) and \( \forall \theta \in S, g_\theta \) vanishes at all of its points of continuity.

(iii) \( g(1) = 0 \) and \( \forall \theta \in S, g_\theta = 0, \text{ a.e. on } J_\theta \).
Proof of Theorem 1. Necessity. (i) By the Principle of Uniform Boundedness there exists an \( M \) such that \( \| \Gamma_i \| = \| \mu_i \| \leq M \).

(ii) Let \( f \equiv 1 \) on \( J \). Then

\[
\Gamma_i(\bar{t}) = \int d\Gamma_i = \int d\mu_i \to 0 \quad \text{as } l \to \infty.
\]

(iii) Since \( \| \Gamma_i \| \leq M \), by [H1] there exists a subsequence \( \langle \Gamma_{i_k} \rangle \) such that \( \Gamma_{i_k} \to g \in B \) with \( \| g \| \leq M \). By [H2] then \( \Gamma_{i_k} \rightharpoonup g \). Since \( \Gamma_{i_k} \to 0 \) this implies \( \int f dg = 0 \), \( f \in C \), so that by Theorem 2, for all \( \theta \in S \), \( g_{\theta} \) vanishes at all of its continuity points, which implies

\[
\int_{J_\theta} |g_{\theta}| \, d\nu_{\theta} = 0, \quad \theta \in S.
\]

Clearly, for every \( \theta \in S \), \( |(\Gamma_{i_k})_{\theta}| \to |g_{\theta}| \) and for \( k = 1, 2, \ldots \), we have \( \|(\Gamma_{i_k})_{\theta}\| \leq M \), \( \|g_{\theta}\| \leq M \). By Lebesgue’s Dominated Convergence Theorem we find \( \int_{J_\theta} |(\Gamma_{i_k})_{\theta}| \to 0 \), \( \theta \in S \).

If, for some \( \theta \in S \), we do not have \( \int_{J_\theta} |(\Gamma_i)_{\theta}| \to 0 \), then there exists a subsequence \( \langle \Gamma_{m_k} \rangle \) of \( \langle \Gamma_i \rangle \), and some \( \varepsilon > 0 \) such that

\[
\int_{J_\theta} |(\Gamma_{m_k})_{\theta}| > \varepsilon, \quad m = 1, 2, \ldots.
\]

However, since \( \Gamma_{m_k} \rightharpoonup 0 \) we can extract a subsequence \( \langle \Gamma_{m_k} \rangle \) as above such that \( \int_{J_\theta} |(\Gamma_{m_k})_{\theta}| \to 0 \), which is a contradiction.

Sufficiency. By [H1] there exists a subsequence \( \langle \Gamma_{i_k} \rangle \) such that \( \Gamma_{i_k} \to g \in B \). Therefore, \( g(1) = 0 \) and \( (\Gamma_{i_k})_{\theta} \to g_{\theta}, \theta \in S \). By Lebesgue’s Dominated Convergence Theorem, for every \( \theta \in S \),

\[
\int_{J_\theta} |(\Gamma_{i_k})_{\theta}| \to \int_{J_\theta} |g_{\theta}|,
\]

from which it follows that for all \( \theta \in S \), \( g_{\theta} = 0 \), a.e. on \( J_\theta \). By Theorem 2, we must have \( \int f dg = 0 \), \( f \in C \). By [H2], \( \Gamma_{i_k} \rightharpoonup g \), and hence \( \Gamma_{i_k} \rightharpoonup 0 \), so that \( \mu_{i_k} \rightharpoonup 0 \).

If \( \mu_i \rightharpoonup 0 \) is not true, then for some subsequence \( \langle \mu_{m_k} \rangle \) of \( \langle \mu_i \rangle \), some \( f \in C \) and some \( \varepsilon > 0 \), we must have

\[
\left| \int f d\mu_m \right| \geq \varepsilon, \quad m = 1, 2, \ldots.
\]

However, we can extract a subsequence \( \langle \mu_{m_k} \rangle \), as above, such that \( \mu_{m_k} \rightharpoonup 0 \). This is a contradiction and the proof is complete.

II. A discussion of \( n \)-dimensional Riemann-Stieltjes integration. For \( 0 < t \in J \), denote the closed subinterval \( \times_{n=1}^n [0, t_i] \) by \( I_t \).

Let \( g \in B \) and \( t > 0 \). Then, \( |g(t)| = |\Delta g(I_t)| \leq \| g \| \), so that

\[
\| g \|_\infty \leq \| g \|, \quad g \in B.
\]
Now let \( g, h \in B \) and let \( \sigma \) be any partition of \( J \). We have

\[
\sum_{I \in \sigma} |\Delta(g \pm h)(I)| = \sum_{I \in \sigma} |\Delta g(I) \pm \Delta h(I)| \\
\leq \sum_{I \in \sigma} |\Delta g(I)| + \sum_{I \in \sigma} |\Delta h(I)| < \|g\| + \|h\|,
\]

from which it follows that

\[
(2-1) \quad \|g \pm h\| \leq \|g\| + \|h\|.
\]

Hence, \( B \) is a normed linear space under \( \| \cdot \| \).

A function \( p \) on \( J \) is said to be positively monotone increasing if and only if for all subintervals \( I \) of \( J \) we have \( \Delta p(I) \geq 0 \). Such functions are called positively monotonely monotone in [5]. When \( p \) is positively monotone increasing and vanishes at all the points lying on coordinate planes we find \( p \geq 0 \) on \( J \) since for any \( t > 0 \) we have \( p(t) = \Delta p(tI) \). Moreover, in this case, we also have \( \|p\| = p(\bar{1}) \), so that \( p \in B \).

Given a subinterval \( K \) of \( J \) and a partition \( \tau \) of \( K \), we may extend \( \tau \) to a partition \( \sigma \) of all of \( J \) such that for all subintervals \( I \in \tau \), we have \( I \in \sigma \). Let \( g \in B \) and consider the restriction \( g|_K \). Then,

\[
\sum_{I \in \tau} |\Delta g(I)| \leq \sum_{I \in \sigma} |\Delta g(I)| < \|g\|,
\]

so that \( g|_K \) is of bounded variation on \( K \) and if we denote its total variation on \( K \) by \( \|g\|_K \), then

\[
(2-2) \quad \|g\|_K \leq \|g\|, \quad \text{\( K \) a subinterval of \( J \).}
\]

**DEFINITION (2-1).** Let \( g \in B \). We define the variation function of \( g \), denoted \( \Pi_g \) or \( \Pi \), as follows: \( \Pi(t) = \|g\|_{t^*} \), for \( t > 0 \) and \( \Pi(t) = 0 \) otherwise.

We have \( \Pi(\bar{1}) = \|g\| \), and that for every subinterval \( I \),

\[
(2-3) \quad \Delta \Pi(I) \geq \Delta g(I),
\]

so that \( \Pi \) is positively monotone increasing.

By (2-1), the difference of two positively monotone increasing functions in \( B \) is again in \( B \). For the converse we introduce

**DEFINITION (2-2).** Let \( g \in B \). The positive variation of \( g \), denoted \( \psi \), and the negative variation of \( g \), denoted \( \phi \), are defined as \( \psi = \frac{1}{2}(\Pi + g) \), \( \phi = \frac{1}{2}(\Pi - g) \); on \( J \).

It follows from (2-3) that \( \psi, \phi \) are positively monotone increasing lying in \( B \). Moreover, by their definition

\[
g = \psi - \phi, \quad g \in B,
\]

to which we shall refer as the Jordan decomposition of \( g \).

Let \( p, q \in B \) be positively monotone increasing such that \( g = p - q \) on \( J \). By (2-1), for all \( t > 0 \) in \( J \) we have

\[
\Pi(t) = \|g\|_{t^*} \leq \|p\|_{t^*} + \|q\|_{t^*} = p(t) + q(t).
\]

However, \( \Pi = \psi + \phi \) and so \( p \geq \psi, q \geq \phi \) on \( J \).
When \( n = 1 \), \( g \) is continuous at \( x \) \( \Leftrightarrow \Pi_q \) is continuous at \( x \). When \( n > 1 \), this is no longer the case. For examples, put \( n = 2, u_j = \frac{1}{4} + 4^{-j}, x_0 = (\frac{1}{2}, \frac{1}{2}), x_j = (u_j, u_j), y_j = (\frac{1}{2}, u_j) \) and consider two sequences of functionals, \( K_j(f) = f(x_0) - f(x_j) \) and \( L_j = f(x_0) - f(y_j) \). Compute their \( \mu \)'s and \( \Gamma \)'s, and notice that both sequences are weak-star convergent to zero.

Let \( P_\theta \) denote the orthogonal projection of \( J \) on \( J_\theta, \theta \in S \).

For \( x, y \in J_\theta, x <_\theta y \) means \( x_j < y_j, j \notin \theta \). We simply write \( x < y \) whenever it is clear that we mean \( x <_\theta y \). Let

\[
J(m) = \prod_{j=1}^{m} [0,1],
\]

so that \( J(m) = J \). Then \( J_\theta \) and \( J(m) \) are naturally isomorphic \((m = n - |\theta|)\) and \( < \) is preserved under the natural isomorphism. Moreover, the natural isomorphism of \( J_\theta \) and \( J(m) \) induces an isometry between \( C_\theta \) and \( C(m) \), the space of continuous functions on \( J_\theta \) and \( J(m) \), respectively. We mention \( J(m) \) to clarify statements made for \( J_\theta \).

Given a subinterval \( I \) of \( J \) such that \( I \cap J_\theta = I_\theta \neq \emptyset \), then \( I_\theta \) naturally corresponds to some subinterval in \( J(m) \). In the same vein, if \( \sigma \) is a (disjoint) partition of \( J \) and \( \sigma_\theta \) is defined as

\[
\sigma_\theta = \{I_\theta \mid I \in \sigma, I \cap J_\theta \neq \emptyset \},
\]

then \( \sigma_\theta \) is a (disjoint) partition of \( J_\theta \).

**Definition (2-3).** A subinterval in \( J_\theta \) is given by \( I_\theta = \times_{j=1}^{n} [x^\theta_j, x^\theta_j] \) where for all \( j \in \theta \) we have \( x^\theta_i = x^\theta_i = 1 \) so that \( [x^\theta_i, x^\theta_j] = \{1\} \). We may extend \( I_\theta \) to a subinterval \( I \) of \( J \) by replacing each \( \{1\} \) by the linear subinterval \([0,1]\): \( I = P_\theta^{-1}(I_\theta) \). We refer to \( I \) as the standard extension of \( I_\theta \).

Next, let \( h \) be any function on \( J_\theta \). We may extend \( h \) to a function \( f \) on \( J \) by defining \( f(x) = h(P_\theta(x)), x \in J \). Then \( f_\theta = f|J_\theta = h \), and we call \( f \) the standard extension of \( h \).

Let \( I_\theta \) be a subinterval of \( J_\theta \) and let \( I \) be its standard extension. Then \( I_\theta = I \cap J_\theta \) and the corners of \( I \) lie either in \( J_\theta \) or else in coordinate planes.

Notice that if \( g \in B, \theta \in S \), then \( \|g_\theta\| \leq \|g\| \). Also, observe that if \( p \in B \) is positively monotone increasing on \( J \), then \( p_\theta \) on \( J_\theta \); moreover, \( \|p_\theta\| = \|p\| = p(1) \).

Let \( g \in B \) and let \( I = \times_{j=1}^{n} [x^I_j, x^I_j] \) be a subinterval of \( J, n > 1 \). Fix some \( j \) and consider the sets

\[
I_1 = \times_{k=1}^{n} [x^I_k, x^I_k] \quad \text{with} \quad [x^I_j, x^I_j] = \{x^I_j\},
\]

\[
I_2 = \times_{k=1}^{n} [x^I_k, x^I_k] \quad \text{with} \quad [x^I_j, x^I_j] = \{x^I_j\}.
\]

Then \( I_1 \) is a subinterval in a hyperplane (a subset of \( J \) obtained by fixing a single coordinate) say \( H_1 \) and \( I_2 \) is a subinterval in a hyperplane \( H_2 \). Let \( g_1 = g|H_1, g_2 = g|H_2 \). Then

\[
\Delta g(I) = \Delta g_2(I_2) - \Delta g_1(I_1).
\]
We observe that in computing \( \gamma_{i_1, \ldots, i_{n-1}} \) for vertices of \( I_1 \), we ignore the coordinate \( x_1^k \), thinking of \( H_1 \) as \( J_{(n-1)} \). When vertices of \( I_1 \) are thought of as vertices of \( I \), then \( x_1^k \) has to be considered as a coordinate and in this case \( \gamma_{i_1, \ldots, i_n} \) will have opposite sign to \( \gamma_{i_1, \ldots, i_{n-1}} \) since in the sum \( \sum_{j=1}^n i_j \), we have \( i_k = 1 \).

For a set \( A \subset J_\theta \), we denote its interior by \( A^0 \) and its closure by \( \overline{A} \), both with respect to relative topology on \( \mathbb{R}_\theta^n \).

Given \( t \in J^0 \), there are \( 2^n \) subintervals having a corner at \( t \) such that their union is all of \( J \). By a quadrant with respect to \( t \) we mean any one of these subintervals, containing only that portion of their boundary which is common to the boundary of \( J \). Two quadrants of \( t \) will be given special name and symbol. The quadrant which has a corner at 0 will be called the left quadrant, denoted \( I_t^- \), and the one with a corner at 1 will be referred to as the right quadrant, denoted \( I_t^+ \).

For a subinterval \( I = [x_1, x_2] \) put

\[
\|I\| = \max\{x_2 - x_j : j = 1, 2, \ldots, n\}.
\]

**Definition (2-4).** Let \( f \) be a function and \( x \in J \). We shall say that the left limit of \( f \) at \( x \), denoted \( f(x - 0) \), exists if and only if for every \( \varepsilon > 0 \) there exists a \( \delta > 0 \) such that for every subinterval \( K \subset I_x^- \) with \( \|K\| < \delta \) and having a corner at \( x \), we have

\[
\forall y \in K^0 : |f(y) - f(x - 0)| < \varepsilon.
\]

The \( 2^n \) quadrantal limits at a point \( x \in J^0 \) are defined similarly and the right limit is denoted \( f(x + 0) \).

If in the preceding definition we replace \( K^0 \) by \( K \), the closed subinterval, then we talk about the strong quadrantal limit.

**Definition (2-5).** A function \( f \) is said to be left continuous at \( x \in J^0 \) if and only if the strong left limit of \( f \) at \( x \) exists and equals \( f(x) \). The function is said to be continuous at \( x \) if and only if all the strong quadrantal limits at \( x \) exist and coincide with \( f(x) \).

**Definition (2-6).** Let \( f \) be a function.

(a) The oscillation of \( f \) on a subinterval \( I \), denoted \( O(f; I) \), is defined as

\[
O(f; I) = \sup \{|f(x) - f(y)| : x, y \in I\}.
\]

(b) The oscillation of \( f \) at a point \( x \in J^0 \), denoted \( O(f; x) \), is defined as

\[
O(f; x) = \inf \{O(f; I) : x \in I^0\}.
\]

**Remark (2-1).** Let \( f \) be a function and \( x \in J^0 \). Suppose \( f(x - 0) \) exists. Then, for every \( \varepsilon > 0 \), we can find a subinterval \( K \subset I_x^- \) with a corner at \( x \) such that for every subinterval \( I \subset K^0 \), we have \( O(f; I) < \varepsilon \).

We note that, in the usual way, \( f \) will have a limit at a point \( x \) if and only if all the strong quadrantal limits of \( f \) at \( x \) exist and coincide. In this language, \( f \) is continuous at \( x \) if and only if its limit at \( x \) exists and coincides with \( f(x) \). Furthermore, \( f \) will be continuous at \( x \in J^0 \) if and only if for every \( \varepsilon > 0 \) there exists a subinterval \( K \) centered at \( x \) such that \(|f(y) - f(x)| < \varepsilon, y \in K^0\).

Finally, a function \( f \) is continuous at \( x \in J^0 \) if and only if \( O(f; x) = 0 \).
Let $p$ be positively monotone increasing and $K \subset I$ subintervals of $J$. It is clear that $K$ may be extended to a partition $\tau$ of $I$ with $K \in \tau$. If we denote the subintervals in $\tau$ by $H$, then

$$\Delta p(I) = \Delta p(K) + \sum_{H \in \tau} \Delta p(H \neq K)$$

so that $\Delta p(K) \leq \Delta p(I)$.

**Proposition (2-1).** Let $p$ be positively monotone increasing and $x \in J$. Then all the quadrantal limits of $p$ at $x$ exist.

**Proof.** We take $x \in J^0$ and prove the existence of $p(x - 0)$. Clearly, $y < x$ implies $y \in I_x^-$. Let

$$\alpha = \inf_{y < x} \Delta p([y, x]).$$

Consider any sequence $y_l \to x$, $y_l < x$ for $l = 1, 2, \ldots$. Given $\varepsilon > 0$ choose $K \subset I_x^-$ such that $K$ has a corner at $x$ and $\Delta p(K) - \alpha < \varepsilon$. Since $K$ is fixed and $y_l \to x$, for some $l_0$ we must have $l > l_0 \Rightarrow [y_l, x] \subset K$, and therefore $\Delta p([y_l, x]) \leq \Delta p(K)$. Hence

$$l > l_0 \Rightarrow |\Delta p([y_l, x]) - \alpha| < \varepsilon,$$

which means $\lim_{l \to \infty} \Delta p([y_l, x]) = \alpha$. For the sequence $(y_l)$, fix all the coordinates of each of its terms except the $j$th coordinate. Then, as $l \to \infty$ we have $y_l \to x^j$ and $y_l < x^j$, $l = 1, 2, \ldots$. We know that

$$\lim_{l \to \infty} p(y_l, y_2, \ldots, y_l, \ldots, y^n)$$

exists for each $j = 1, 2, \ldots, n$, and is independent of the way $y_l$ approaches $x^j$ (the function is simply monotone increasing in the $j$th coordinate). It follows from the definition of $\Delta p$ that $\lim_{y_l \to x} p(y_l)$ exists and is independent of the way $y_l$ approaches $x$. But this is equivalent to the existence of $p(x - 0)$.

**Proposition (2-2).** If $p$ is positively monotone increasing then the points of discontinuity of $p$ lie on a countable number of hyperplanes.

**Proof.** For each $x \in J^0$, all the quadrantal limits of $p$ at $x$ exist. Given $\varepsilon > 0$, we can choose $2^n$ subintervals according to Remark (2-1), one in each quadrant and with a corner at $x$. We then take $K$ centered at $x$ and contained in the union of the above $2^n$ subintervals. Then, for any $y \in K$, $y$ lying in a quadrant, we have

$$O(p; y) < \varepsilon.$$

We now cover $J$ with a finite number of such subintervals, say $K_r$, $r = 1, 2, \ldots, m$, centered at $x_r \in J$ (with obvious interpretation of $K_r$ being centered at a boundary point $x_r$). It follows that (2-5) is satisfied by all the points which do not lie on the hyperplanes passing through $x_r$. Hence, the set of points $z \in J$ such that $O(p; z) \geq 1/l$, $l = 1, 2, \ldots$, is contained in the union of a finite number of hyperplanes. This means that the set of points in $J$ at which the oscillation of $p$ exceeds zero is contained in the union of a countable number of hyperplanes. See Remark (2-1).
**Proposition (2-3).** Let \( \langle p_i \rangle \) be a sequence of positively monotone increasing functions in \( B \). If the sequence is pointwise bounded on \( J \), then there exists a subsequence of it which converges to a positively monotone increasing function, pointwise on \( J \).

**Proof.** The proof is by induction on \( n \), the dimension. For \( n = 1 \), this is Lemma 2 on p. 221 of [7]. Let \( 1 < n < \infty \) and suppose that the proposition is true for \( 1, 2, \ldots, n - 1 \). Now we proceed as follows.

Let \( E \) be the countable dense subset of \( J \) consisting of points all of whose coordinates are rational. Extract a subsequence \( \langle p_{i_k} \rangle \) so that it converges on \( E \) and put \( \lim p_{i_k} = p \) on \( E \). Clearly, \( p \) is positively monotone increasing on \( E \). For any point \( t \in J - E \) define

\[
p(t) = \sup_{x \in I_t \cap E} p(x),
\]

We assert that \( p \) is positively monotone increasing. Let \( I = \times_{j=1}^n [c^j, d^j] \) be any subinterval, \( c < d \) points in \( J \). Choose \( 2^n \) sequences in \( E \), each converging to a corner of \( I \) and lying in the left quadrant of the corner to which they converge. Consider the corner \( c \) of \( I \) and let \( \langle x_i \rangle \) be the sequence in \( E \) that converges to \( c \). Given \( \varepsilon > 0 \), take \( y \in E \cap I_c \) such that

\[
(2-6) \quad p(c) - p(y) < \varepsilon
\]

and choose \( l_0 \) such that \( l > l_0 \Rightarrow y < x_i < c \). This means, for \( l > l_0 \) we have \( I_y \subset I_{x_i} \) and hence \( p(c) \) is supremum,

\[
l > l_0 \Rightarrow p(y) \leq p(x_i) \leq p(c),
\]

which together with (2-6) implies \( \lim_{l \to \infty} p(x_i) = p(c) \). Clearly, a similar result holds for all the corners of \( I \). Let \( I_l \) be the subinterval having as its corners the points of the \( 2^n \) sequences (converging to the corners of \( I \)) for \( l = 1, 2, \ldots \). It follows that

\[
\Delta p(I) = \lim_{l \to \infty} \Delta p(I_l) > 0
\]

and proves the assertion.

Let \( x_0 \) be a point of continuity of \( p \). We assert that \( p_{i_k}(x_0) \to p(x_0) \). Given \( \eta > 0 \), choose \( K \) centered at \( x_0 \) such that

\[
(2-7) \quad \forall I \subset K^0: O(p; I) < \eta/2.
\]

Let \( I = [x, y] \) be centered at \( x_0 \), \( I \subset K^0 \), and \( x, y \in E \). Since \( x < x_0 < y \), we find

\[
(2-8) \quad p_{i_k}(x) \leq p(x_0) \leq p_{i_k}(y), \quad k = 1, 2, \ldots
\]

(This follows from the fact that \( I_x \subset I_{x_0} \subset I_y \).) Choose \( k_0 \) such that

\[
k > k_0 \Rightarrow |p_{i_k}(x) - p(x)| < \eta/2 \quad \text{and} \quad |p_{i_k}(y) - p(y)| < \eta/2.
\]

From (2-7) we have \( |p(x) - p(x_0)| < \eta/2 \) and \( |p(y) - p(x_0)| < \eta/2 \), which combined with preceding inequalities yields

\[
k > k_0 \Rightarrow |p_{i_k}(x) - p(x_0)| < \eta \quad \text{and} \quad |p_{i_k}(y) - p(y_0)| < \eta.
\]

The first inequality above and (2-8) give

\[
k > k_0 \Rightarrow p_{i_k}(x_0) - p(x_0) \geq p_{i_k}(x) - p(x_0) > -\eta,
\]
and similarly the second inequality gives
\[ k > k_0 \Rightarrow p_i(x_0) - p(x_0) \leq p_i(y) - p(x_0) < \eta, \]
and the assertion follows.

Finally, the points of discontinuity of \( p \) lie on a countable number of hyperplanes \( H_1, H_2, \ldots \). Since \( p_k \mid H_i \) are positively monotone increasing for \( k = 1, 2, \ldots \), we invoke the induction hypothesis to extract a subsequence of \( \langle p_{m_i} \rangle \) which converges pointwise on \( H_1 \). Then, we extract a subsequence \( \langle p_{m_{i,j}} \rangle \) of the sequence \( \langle p_{m_i} \rangle \) which converges pointwise on \( H_2 \), and continue the process. The diagonal subsequence will then converge pointwise on \( J \) to a positively monotone increasing function.

One can replace the class of positively monotone increasing functions by functions in \( B \) in these propositions via an application of the Jordan decomposition. The consequent modification of Proposition (2-3) is \( [H_1] \).

When \( n = 1 \) and \( \sigma : 0 = x_0 < x_1 < \cdots < x_m = 1 \) is a partition of \( J \) we define \( \| \sigma \| = \max \{ x_{k-1} - x_k : k = 1, 2, \ldots, m \} \), and for \( 1 \leq n < \infty \), we let \( \| \sigma \| = \max \{ \| \sigma_j \| : j = 1, 2, \ldots, n \} \). We say the partition \( \sigma \) is finer than \( \tau \), denoted \( \tau \preceq \sigma \), if and only if \( \tau_j \leq \sigma_j \) for \( j = 1, 2, \ldots, n \).

Let \( f, g \) be two functions and \( \sigma \) a partition. We define the sum of \( f \) with respect to \( g \) for \( \sigma \) by
\[ S(f, g, \sigma) = \sum_{I \in \sigma} f(t_I) \Delta g(I), \]
where \( t_I \) is a point in \( I \). The (Stieltjes) integral of \( f \) with respect to \( g \), over \( J \), is denoted
\[ \int_J f dg; \]
we have not written \( J \) when the integral was understood to be over all of \( J \). The following two definitions of integral will be considered.

**The Refinement Definition.** We shall say that the refinement integral of \( f \) with respect to \( g \) exists if there exists a (real) number denoted by \( \int_J f dg \) such that for every \( \varepsilon > 0 \), there exists a partition \( \sigma \) with the property that for all \( \tau \preceq \sigma \), and independent of the choice of the \( t_I \), we have
\[ |S(f, g, \tau) - \int_J f dg| < \varepsilon. \]

**The Norm Definition.** We shall say that the norm integral of \( f \) with respect to \( g \) exists if there exists a number, also denoted by \( \int_J f dg \), such that for every \( \varepsilon > 0 \), we can find a \( \delta > 0 \), with the property that for all \( \tau \), with \( \| \tau \| < \delta \), (2-10) is satisfied independent of the choice of \( t_I \).

In (2-9), \( f \) is called the integrand and \( g \) the integrator. In consideration of integral, the integrand will always lie in \( C \) and the integrator will always lie in \( B \) in which case the integral exists in both senses defined above and the values coincide. This is a consequence of Theorem 6.8, p. 108 of [10] for \( n = 1 \), and Theorem 9.3, p. 129 of [5] for \( n \geq 2 \).
The integral is linear with respect to both the integrand and the integrator. The proof given for \( n = 1 \) in Theorems 9-2 and 9-3, p. 193 of [1], is valid for \( 1 \leq n < \infty \).

We have \( | \int f \, dg | \leq \| f \|_\infty \cdot \| g \| \) since for any partition \( \sigma \),

\[
|S(f, g, \sigma)| = \left| \sum_{I \in \sigma} f(t_I) \cdot \Delta g(I) \right| \leq \| f \|_\infty \cdot \sum_{I \in \sigma} |\Delta g(I)| \leq \| f \|_\infty \cdot \| g \|.
\]

From (2-2) it follows that if \( f \in C \) and \( g \in B \) and \( \sigma \) is a partition, then \( \forall I \in \sigma: \int_I f \, dg \) exists. In fact, by Theorem 8.4, p. 126 of [5], we have

\[
\int f \, dg = \sum_{I \in \sigma} \int_I f \, dg.
\]

**Remark (2-2).** Let \( \theta \in S \), let \( f_\theta \) be a continuous function on \( J_\theta \) and let \( f \) be its standard extension. Then, for every \( g \in B \), \( \int f \, dg = \int_{J_\theta} f_\theta \, dg_\theta \).

**Proof of [H2].** [H,1] implies that \( \| g \| \leq M \), so that \( g \in B \). Let \( f \) be any function in \( C \). By the uniform continuity of \( f \) on \( J \), Theorem 7.3 on p. 180 of [6], for any \( \varepsilon > 0 \) we can find a \( \delta > 0 \) such that for every partition \( \sigma \) with \( \| \sigma \| < \delta \), we have

\[
\forall I \in \sigma, \forall t', t'' \in I: |f(t') - f(t'')| < \frac{\varepsilon}{4M}.
\]

Let \( t_I \) be any point in \( I \in \sigma \). We have

\[
\int f \, dg = \sum_{I \in \sigma} \int_I f \, dg = \sum_{I \in \sigma} \int_I [f - f(t_I) + g(t_I)] \, dg = \sum_{I \in \sigma} \int_I [f - f(t_I)] \, dg + \sum_{I \in \sigma} f(t_I) \int_I dg = \sum_{I \in \sigma} \int_I [f - f(t_I)] \, dg + \sum_{I \in \sigma} f(t_I) \cdot \Delta g(I).
\]

Keeping the partition \( \sigma \) fixed, a similar computation gives

\[
\int f \, dg_I = \sum_{I \in \sigma} \int_I [f - f(t_I)] \, dg_I + \sum_{I \in \sigma} f(t_I) \cdot \Delta g(I), \quad l = 1, 2, \ldots.
\]

Let each \( \sigma_j \) have \( m_j \) linear subintervals, \( j = 1, 2, \ldots, n \). Then \( \sigma \) has \( m = m_1 \cdot m_2 \cdot \ldots \cdot m_n \) subintervals. From \( g_I \to g \) it follows that \( \Delta g(I) \to \Delta g(I) \) for each \( I \in \sigma \).

So, we can choose \( l_0 \) such that

\[
l > l_0 = \forall I \in \sigma: |\Delta g_I(I) - \Delta g(I)| < \frac{\varepsilon}{2\| f \|_\infty \cdot m}.
\]

Hence, for \( l > l_0 \) we find

\[
\left| \sum_{I \in \sigma} f(t_I) \Delta g(I) - \sum_{I \in \sigma} f(t_I) \Delta g(I) \right| < \| f \|_\infty \cdot \sum_{I \in \sigma} |\Delta g(I) - \Delta g(I)| < \frac{\varepsilon}{2}.
\]

Next, we take care of the sum involving integral. Thus,

\[
\left| \sum_{I \in \sigma} \int_I [f - f(t_I)] \, dg_I - \sum_{I \in \sigma} \int_I [f - f(t_I)] \, dg \right| = \left| \sum_{I \in \sigma} \int_I [f - f(t_I)] \, d(g_I - g) \right| \leq \frac{\varepsilon}{4M} \cdot \| g_n - g \| \leq \frac{\varepsilon}{2}.
\]
Therefore, by the triangle inequality
\[ l > l_0 \Rightarrow \left| \int f \, dg_l - \int f \, dg \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \]
Since \( f \in C \) was arbitrary, we have shown \( g_l \to g \).

**Proposition (2-4).** Let \( \Gamma \) be the distribution function of \( \mu \in \mathcal{M} \). Then \( \Gamma_{\theta} \) is right continuous on \( J_0^0 \) for all \( \theta \in S \).

**Proof.** It suffices to prove the assertion for a positive measure \( \mu \in \mathcal{M} \). Furthermore, for notational convenience we assume \( \theta \) is the empty set, but the argument will work for all \( \theta \in S \).

Consider any sequence of points \( \langle x_l \rangle \) in \( I_x^+ \) such that \( x_l \to x \) and \( x_{l+1} < x_l \), \( l = 1, 2, \ldots \). Clearly,
\[ I_{x_{l+1}} \subset I_{x_l}, \quad l = 1, 2, \ldots, \]
and
\[ I_x = \bigcup_{l=1}^{\infty} I_{x_l}. \]  
(2-11)
Proposition 14, p. 61 of [8] and (2-11) give \( \mu(I_x) = \lim_{l \to \infty} \mu(I_{x_l}) \), so that \( \Gamma(x) = \lim_{l \to \infty} \Gamma(x_l) \). If \( \Gamma \) is not right continuous, we must have \( \Gamma(x_l) > \Gamma(x + 0) > \Gamma(x) \) which contradicts the fact just established.

**III. A proof of Theorem 2.**

**Lemma (3-1).** Let \( E \in J^0 \) with \( \nu(E) = 0 \). Then, for every \( \delta > 0 \), there exists a partition \( \sigma \) of \( J \) such that \( \| \sigma \| < \delta \) and no points of \( \sigma \) (i.e., corners of subintervals in \( \sigma \)) are in \( E \).

**Proof.** Let \( n = 1, \tau: 0 = x_0 < \cdots < x_i = 1 \) a partition of \( J = [0, 1] \) such that every subinterval in \( \tau \) has length equal to \( \delta_1 < \delta \). Suppose some \( x_i \in E \). Choose the points \( x_i' < x_i < x_i'' \) such that
\[ x_i', x_i'' \notin E \quad \text{and} \quad x_i - x_i' < \frac{\delta_1}{2}, \quad x_i'' - x_i < \frac{\delta_1}{2}. \]
Then replace \( x_i \) by \( x_i', x_i'' \). Continuing in this way we end up with a partition \( \sigma \) as desired.

Let the statement be true for \( n - 1, n > 1 \) and let \( J \) be \( n \)-dimensional. Let \( H_x \) be a hyperplane obtained by fixing the \( n \)th coordinate \( x^n \) and put \( E_x = H_x \cap E \), i.e., \( E_x \) is a cross-section of \( E \). Then \( E_x \) are measurable a.e. on \( 0 < x^n < 1 \) and the function \( \xi(x^n) = \nu_{(n-1)}(E_x) \) is nonnegative on its domain with
\[ \int_0^1 \xi \, d\nu_{(1)} = \nu(E) = 0. \]
It follows that \( \xi = 0 \), a.e. and hence \( \nu_{(n-1)}(E_x) = 0 \), a.e. on \( 0 < x^n < 1 \). Choose a partition \( \sigma_n \): \( 0 = x_0 < \cdots < x_m^n = 1 \) so that \( \| \sigma_n \| < \delta \) and \( \nu_{(n-1)}(E_{x_i^n}) = 0 \) for \( i = 1, 2, \ldots, m_n - 1 \). Set
\[ F = \bigcup_{i=1}^{m_n-1} E_{x_i^n}, \]
and we have $v_{(n+1)}(F) = 0$. Let $\theta = \{n\} \in S$. By the induction hypothesis there exists a partition $\tau$ of $J_\theta$ with norm less than $\delta$ which contains no points of $P_\theta(F)$; the intersection of $P_\theta^{-1}$ (points in $\tau$) with hyperplanes $H_{x_i}$, $i = 0, 1, \ldots, m_n$, generates the points of a partition with required properties.

**Proof of Theorem 2.** (i) $\Rightarrow$ (ii). Let $f \equiv 1$ on $J$. Then

$$g(1) = \int dg = 0.$$ 

Recalling Remark (2-2), it suffices to consider the case where $\theta$ is the empty set.

Given $x \in J^0$, let $\langle x_i \rangle$ be a sequence in $I_x^+$, such that $x_i \to x$ and $x_i+1 < x_i < 1$, $l = 1, 2, \ldots$. For each $l$, consider the closed sets

$$A_k^l = \prod_{j=1}^n \left[ \alpha_j, 1 \right], \quad k = 1, 2, \ldots, n,$$

with $\alpha_k = x_{k}^i$ ($k$ th coordinate of $x_i$), and $\alpha_j = 0$ for $j \neq k$. Let

$$A_l = \bigcup_{k=1}^n A_k^l, \quad l = 1, 2, \ldots.$$

For each $l$ then we have two closed sets, $I_x$ and $A_l$. By Urysohn Lemma, p. 207 of [6] for each $l$ there exists an $f_l \in C$ such that

$$f_l(I_x) = 1, \quad f_l(A_l) = 0, \quad \|f_l\|_\infty \leq 1, \quad l = 1, 2, \ldots.$$

Now let $p \in B$ be any positively monotone increasing function which is continuous at the point $x$. Let $B_l$ be the closure of $J - (I_x \cup A_l)$ for $l = 1, 2, \ldots$. Then

$$\int f_l dp = \int_{I_x} f_l dp + \int_{B_l} f_l dp + \int_{A_l} f_l dp = p(x) + \int_{B_l} f_l dp, \quad l = 1, 2, \ldots.$$

Since

$$\int_{B_l} f_l dp \leq \|p\|_{B_l} = p(x_l) - p(x), \quad l = 1, 2, \ldots,$$

we obtain

$$p(x) \leq \int f_l dp \leq p(x_l), \quad l = 1, 2, \ldots.$$ 

Since $p$ is continuous at $x$, we find

$$\int f_l dp \to p(x) \quad \text{as } l \to \infty.$$

Now we write $g = \psi - \phi$, the Jordan decomposition. Let $x \in J$ be a point such that $\psi, \phi$ are both continuous at $x$. Construct the sequence $\langle f_l \rangle \subset C$ as above. Then

$$0 = \int f_l dg = \int f_l d\psi - \int f_l d\phi = \psi(x) - \phi(x) = g(x),$$

so that $g(x) = 0$. Suppose there is a point $y \in J^0$ with $g$ continuous at $y$ but $\psi$ and $\phi$ both discontinuous there. Then, for every $\epsilon > 0$, we can find a subinterval $K$, centered at $y$, such that

$$|g(z') - g(z'')| < \epsilon, \quad z', z'' \in K.$$
By Proposition (2-2), for some \( z \in K, \psi \) and \( \phi \) are both continuous at \( z \) so that \( g(z) = 0 \). It follows that, for every \( \varepsilon > 0 \), \( |g(z) - g(y)| = |g(y)| < \varepsilon \), and hence \( g(y) = 0 \). This completes the proof of (i) \( \Rightarrow \) (ii).

(ii) \( \Rightarrow \) (iii). This is an immediate consequence of Proposition (2-2).

(iii) \( \Rightarrow \) (i). Let \( n = 1 \). Given \( \varepsilon > 0 \) for every \( f \in C \) there is a \( \delta > 0 \) such that for all partitions \( \sigma \) with \( ||\sigma|| < \delta \), we have \( |S(f, g, \sigma) - \int f \, dg| < \varepsilon \). By Lemma (3-1) we may choose \( \sigma \) such that \( g \) vanishes at all the points of \( \sigma \). Hence \( S(f, g, \sigma) = 0 \) and it follows that \( |\int f \, dg| < \varepsilon \). Assume now the statement holds for all \( k < n, n > 1 \). Let \( \theta_j(j), 1 \leq j \leq n \). Let \( h_i \) denote the standard extension of \( f_{\theta_i} = f|J_{\theta_i} \). By the induction hypothesis and Remark (2-2), we have

\[
\int_{J_{\theta_i}} h_i \, dg = \int_{J_{\theta_i}} (h_i)_{\theta_i} \, dg_{\theta_i} = 0.
\]

Let \( f_1 = f - h_i \); then \( (f_1)_{\theta_i} = 0 \) and \( \int f_1 \, dg = \int f \, dg \). Iterating this process, let \( h_2 \) denote the standard extension of \( (f_1)_{\theta_i} \) and set \( f_2 = f_1 - h_2 \), so that \( (f_2)_{\theta_i} = 0 \), \( j = 1, 2 \), and \( \int f_2 \, dg = \int f \, dg \). After \( n \) iterations we end up with \( f_n \in C \) such that \( (f_n)_{\theta_i} = 0, 1 \leq j \leq n \), and \( \int f_n \, dg = \int f \, dg \). The proof will be complete when we show the left side of the preceding equation vanishes. Let \( \varepsilon > 0 \) be given. Since \( f_n \) is uniformly continuous, we can find a \( \delta > 0 \) such that for every subinterval \( I \) with \( ||I|| < \delta \), we have

\[
\int_I f_n \, dg = \int_I (f_n)_{\theta_i} \, dg_{\theta_i} \leq \frac{\varepsilon}{n \cdot ||g||}.
\]

By Lemma (3-1) we can choose a partition \( \sigma \) with \( ||\sigma|| < \delta \) and such that \( g \) vanishes at all the points of \( \sigma \) in \( J^0 \). Consider any one of the linear partitions comprising \( \sigma \), say \( \sigma_j \): \( 0 = x^1_0 < x^1_1 < \cdots < x^1_m = 1 \). Let \( I_j = \times_{j=1}^n [z^i_j, z^i_{j+1}] \) where \( [z^i_j, z^i_{j+1}] = [x^i_{m_j}, 1] \) and \( [z^i_j, z^i_{j+1}] = [0, 1] \) for \( 2 \leq j \leq n \). Then, for every \( y \in I_1 \) we have \( |f_n(y)| < \varepsilon/n \cdot ||g|| \). Choose the subintervals \( I_2, \ldots, I_n \) similar to \( I_1 \). Then

\[
\left| \int f_n \, dg \right| \leq \sum_{j=1}^n \left| \int_{I_j} f_n \, dg \right| \leq n \cdot \frac{\varepsilon}{n \cdot ||g||} = \varepsilon.
\]

**Corollary (3-1).** The integral of every \( f \in C \) with respect to some \( g \in B \) vanishes if and only if \( g(1) = 0 \) and for every \( \theta \in S \) all the quadrantals limits of \( g_{\theta} \) vanish.

**Proof.** Clearly, the same argument works for all \( \theta \in S \) and for convenience we take \( \theta \) to be the empty set.

Suppose \( \int f \, dg = 0, f \in C \). Then \( g(1) = 0 \). Without loss of generality, let \( 0 < x \) be any point in \( J \) and we shall only show \( g(x - 0) = 0 \). Given \( \varepsilon > 0 \), choose \( K \subset I_x^c \) such that \( K \) has a corner at \( x \) and

\[
|g(y) - g(x - 0)| < \varepsilon, \quad y \in K^0.
\]

Choose \( z \in K^0 \) with \( g \) continuous there, so that \( g(z) = 0 \). This gives

\[
\forall \varepsilon > 0: |g(x - 0)| < \varepsilon \Rightarrow g(x - 0) = 0.
\]

The converse is obvious.
Let \( \hat{\Gamma} = \{ g \in B \mid g_\theta \text{ is right continuous on } J_\theta^0, \theta \in S \} \). By Proposition (2-4), \( \hat{\Gamma} \) contains all the distribution functions of elements in \( \hat{\mathcal{M}} \). Conversely, every element of \( \hat{\Gamma} \) is a distribution function because by the preceding corollary we have

**Corollary (3-2).** Let \( \Gamma_1, \Gamma_2 \in \hat{\Gamma} \). Then, for every \( f \in C \),

\[
\int f \, d\Gamma_1 = \int f \, d\Gamma_2 \iff \Gamma_1 \equiv \Gamma_2 \quad \text{on } J.
\]

So, given a sequence \( \langle \mu_i \rangle \) in \( \hat{\mathcal{M}} \) the corresponding sequence \( \langle \Gamma_i \rangle \) in \( \Gamma \) is unique, but infinitely many sequences \( \langle g_i \rangle \) in \( B \) correspond to \( \langle \mu_i \rangle \) such that \( \mu_i \rightharpoonup 0 \) if and only if \( g_i \rightharpoonup 0 \) where we assume each \( \mu_i \) corresponds to \( g_i, l = 1, 2, \ldots \). However, \( \| \mu_i \| < M \) does not even imply that the sequence \( \langle g_i \rangle \) is bounded. For instance, let \( J = [0, 1] \), and for the sequence \( \| g_i \| = 0 \) choose \( \langle g_i \rangle \) as follows:

\[
g_i(x) = 0, \quad x \neq \frac{1}{i}, \quad g_i\left(\frac{1}{i}\right) = l.
\]

We have the following generalization of Theorem 1.

**Theorem 3.** Let \( \| \mu_i \| \leq M \). Then \( \mu_i \rightharpoonup 0 \) if and only if for every corresponding sequence \( \langle g_i \rangle \) in \( B \) we have

(i) \( g_i(1) \to 0 \), as \( l \to \infty \);

(ii) \( \forall \theta \in S: \int_{J_\theta} |g_i| \, d\nu_\theta \to 0 \), as \( l \to \infty \).

**Proof.** Let \( \langle \Gamma_i \rangle \) be the corresponding sequence of distribution functions. Then

\[
(3-1) \quad \int f \, d(g_i - \Gamma_i) = 0, \quad f \in C,
\]

so that by Theorem 2, \( \forall \theta \in S: g_i = \Gamma_i \), a.e. on \( J_\theta \), \( l = 1, 2, \ldots \). Hence,

\[
\forall \theta \in S: \int_{J_\theta} |g_i| = \int_{J_\theta} |\Gamma_i|, \quad l = 1, 2, \ldots
\]

Moreover, by letting \( f \equiv 1 \) on \( J \) in (3-1), we find \( g_i(1) \to 0 \) if and only if \( \Gamma_i(1) \to 0 \). Therefore, the theorem follows by Theorem 1.

The preceding theorem may be rephrased as: a sequence \( \langle g_i \rangle \) in \( B \) converges weak-star to zero if and only if \( \langle g_i \rangle \) satisfies conditions (i) and (ii) in the theorem and \( \| \mu_i \| \leq M \) where \( \langle L_i \rangle \) is the corresponding sequence in \( C^* \).

**IV. The infinite dimensional case.** Let \( \Lambda \) be an finite index set and let \( J = \times_{\alpha \in \Lambda} [0, 1]_\alpha \). Put the product topology on \( J \) and obtain a compact Hausdorff space. Let \( \mathcal{F} \) denote the set of finite subsets of \( \Lambda \). For \( F \in \mathcal{F} \) and \( x = \{ x_\alpha \} \in J \), let \( P_F(x) = \{ y_\alpha \} \), where \( y_\alpha = x_\alpha, \alpha \in F \), and \( y_\alpha = 1, \alpha \notin F \). For \( F \in \mathcal{F} \), let \( C_F \) denote the subspace of \( C(J) \) comprised of the functions \( f \in C(J) \) with the property that \( f(x) = f(y) \) whenever \( P_F(x) = P_F(y) \). Let \( C_{\mathcal{F}} = \bigcup_{F \in \mathcal{F}} C_F \). Then \( C_{\mathcal{F}} \) is a subalgebra of \( C(J) \) that separates points and contains the constant functions, so \( C_{\mathcal{F}} \) is dense in \( C(J) \). The Riesz representation theorem tells us that the dual of \( C(J) \) is isomorphic and isometric to the space \( \mathcal{B}(J) \) of real-valued, regular Borel measures on \( J \). A bounded sequence, \( \langle \mu_n \rangle \), in \( \mathcal{B}(J) \) converges weakly to zero if and only if \( \int f \, d\mu_n \to 0 \) for each \( f \in C_{\mathcal{F}} \). For \( \phi \neq F \in \mathcal{F} \), let \( J_F = \{(x_{\alpha_1}, x_{\alpha_2}, \ldots, x_{\alpha_n}) \}; \ F = \{ \alpha_1, \alpha_2, \ldots, \alpha_n \} \).
For \( x \in J_F \), let \( \Pi_F(x) = \{ y_\alpha \} \), where \( y_\alpha = x_\alpha, \alpha \in F \), and \( y_\alpha = 1, \alpha \notin F \). For \( \mu \in \mathcal{B}(J) \), define \( \mu^F \) on the Borel subsets \( E \) of \( J_F \) by \( \mu^F(E) = \mu(P_F^{-1}(\Pi_F(E))) \). Notice that \( J_F \) is isomorphic to the \( \text{card}(F) \)-cube. Define the finite dimensional distribution \( \Gamma^F \) on \( J_F \) by \( \Gamma^F(x) = \mu^F(\{ y \in J_F; y \leq x \}) \), \( x > 0, \Gamma^F(x) = 0, x \geq 0 \). Since \( \int f d\mu = \int_J \Gamma_F \) \( f \) \( d\mu^F \) when \( f \in C_F \) and \( f_F(x) = f(\Pi_F(x)) \), we have the following characterization of weak-star convergence to zero.

**Theorem 4.** Let \( \langle \mu_n \rangle \) be a bounded sequence of regular Borel measures on \( J \). Then \( \mu_n \xrightarrow{w^*} 0 \) if and only if

1. \( \mu_n(J) \to 0 \) and
2. for \( \phi \neq F \in \mathcal{F}, \int_J (\Gamma_n)^F | dm_{\text{card}(F)} \to 0 \), where \( m_{\text{card}(F)} \) denotes \( \text{card}(F) \) dimensional Lebesgue measure on \( J_F \).

**REFERENCES**


**DEPARTMENT OF MATHEMATICS, COLORADO STATE UNIVERSITY, FORT COLLINS, COLORADO 80523**

**SCHOOL OF SCIENCES, ACADEMY OF MARINE SCIENCES, I.R., NAVY OF IRAN, P.O. BOX 94-65, NOWSHahr, IRAN**