which can be as small as we please. Thus the sufficient condition

\[
\lim_{\text{mesh}(\Gamma)\to 0} \sum_{j=1}^{k} \omega_j(f;\Gamma) \cdot V[g; t_{j-1}, t_j] = 0
\]

for the integral to exist in sense (II) is satisfied for \( g(t) = t \).

So there are really three theorems about Riemann integration here—a necessary and a sufficient condition for integrability in sense (I) and the fact that sense (I) integrability implies sense (II) integrability—and we would like to see whether and how each extends to Riemann-Stieltjes integration, at least for integrators of bounded variation. Since each theorem involves Lebesgue measure in a highly nontrivial way—we needed to be able to integrate the functions \( \hat{f} \) and \( \check{f} \), which may themselves be discontinuous on a set of positive measure and hence not be Riemann-integrable—the first thing we need is really a countably additive set function that “is to Stieltjes integration \( dg \) as Lebesgue measure is to Riemann integration \( dt \).”

5. Measures from Functionals.

There is no reason to make the construction of Lebesgue-Stieltjes measures any less general than it can be, notwithstanding the fact that we want them to be concretely related to increasing integrators \( g \) on intervals \([a, b]\) of the real line. So the construction given below will be one that works for any positive linear functional \( G \) on the vector lattice \( \mathcal{K}(X) \) of continuous real-valued functions of compact support on a locally compact Hausdorff space \( X \) which will sometimes be assumed \( \sigma \)-compact. {Any unrecognized terms above will be defined below; the prime examples of \( \sigma \)-compact locally compact Hausdorff spaces are the spaces \( \mathbb{R}^n \), and those are what the reader should think of when working through this material.}

Recall that a linear functional on a vector space is simply a linear mapping of the space into its scalar field.

**Definition:** Let \( X \) be a topological space, \( Y \) be a vector space, and \( f : X \to Y \) be a function. Then the support of \( f \) is the closure of the set of points in \( X \) where \( f \) differs from zero:

\[
\text{Supp}(f) = \overline{\{x \in X : f(x) \neq 0\}}.
\]

Logically enough, \( f \) has compact support if \( \text{Supp}(f) \) is compact. It is obvious that

\[
\text{Supp}(f + h) \subseteq \text{Supp}(f) \cup \text{Supp}(h)
\]

and therefore the \( Y \)-valued functions of compact support form a vector space (under the pointwise operations). The same observation applies to the continuous \( Y \)-valued functions of compact support; this space is called \( \mathcal{K}(X,Y) \). If \( Y = \mathbb{R} \) one simply calls this space \( \mathcal{K}(X) \), the space of real-valued continuous functions of compact support on \( X \). Evidently \( \mathcal{K}(X) \) is an algebra and a vector lattice as well as a vector space, and is closed under multiplication by continuous \( \mathbb{R} \)-valued functions (even if they’re not of compact support). If \( 0 \leq f \in \mathcal{K}(X) \) then \( f \wedge 1 \in \mathcal{K}(X) \); this elementary property will be quite useful to us in what follows.

Of course the only element of \( \mathcal{K}(X) \) may be the identically-zero function, but for locally compact Hausdorff spaces \( X \) this doesn’t happen. {If \( X \) is itself compact, then \( \mathcal{K}(X) \) is the same space as \( \mathcal{C}(X) \).}

The basic existence lemma is the following:

**Lemma:** If \( X \) is locally compact Hausdorff, \( K \subseteq X \) is compact, and \( K \subseteq U \) where \( U \subseteq X \) is open, then there exists an \( f \in \mathcal{K}(X) \) satisfying

\[
0 \leq f \leq 1, \quad f[K] = \{1\}, \quad \text{and} \quad \text{Supp}(f) \subseteq U.
\]

**Proof.** Using the local compactness of \( X \) and simple finite-covering arguments, it is easy to find open sets \( V \) and \( W \) with compact closures, such that

\[
K \subseteq V \subseteq \overline{V} \subseteq W \subseteq \overline{W} \subseteq U.
\]
Since $\overline{W}$ is a compact Hausdorff space it is normal, so Uryson’s lemma gives a function $f : \overline{W} \rightarrow [0, 1]$ with $f[K] = \{1\}$ and $f[\overline{W} \setminus V] = \{0\}$. It is easy to check that extending $f$ to $X \setminus \overline{W}$ by setting it equal to zero there results in a continuous function on $X$.

The reader who prefers to think of metric spaces (or $\mathbb{R}^n$) can simply consider the function
\[
f(t) = d(t, X \setminus V) / [d(t, X \setminus V) + d(t, K)]
\]
which clearly satisfies our requirements (with $\text{Supp}(f) \subseteq \overline{V}$, which is known to be compact).

**Definition:** A linear functional $G : \mathcal{K}(X) \Rightarrow \mathbb{R}$ is **positive** (or nonnegative) if
\[
0 \leq f \in \mathcal{K}(X) \Rightarrow 0 \leq G(f) \in \mathbb{R}.
\]
The basic example, and the one for whose sake we’re doing all this, is $G(f) = \int_a^b f(t) \, dg(t)$ where $X = [a, b]$ (so $\mathcal{K}(X) = \mathcal{C}[a, b]$) and $g(t)$ is an increasing function on $[a, b]$. This functional is positive because all the Riemann-Stieltjes sums whose limit defines the integral are nonnegative.

To extract a measure from $G$, we begin by trying to get an analogue of the length-of-intervals set function with which the construction of Lebesgue measure begins. It will be handy to have a couple of directed sets all defined and ready to use.

**Definitions:** For open $U \subseteq X$,
\[
\mathcal{L}(U) = \{ f \in \mathcal{K}(X) : 0 \leq f \leq 1, \text{Supp}(f) \subseteq U \}.
\]
For compact $K \subseteq X$,
\[
\mathcal{U}(K) = \{ f \in \mathcal{K}(X) : 0 \leq f \text{ and } f \geq \chi_K \}.
\]
Both of these sets are directed in the pointwise order of continuous functions, but $\mathcal{L}(U)$ is directed upward and $\mathcal{U}(K)$ is directed downward. If $f_1$ and $f_2$ belong to $\mathcal{L}(U)$, then so does $f_3 = f_1 \lor f_2$, so there is an $f_3 \in \mathcal{L}(U)$ bigger than both of them. Similarly, if $f_1$ and $f_2$ belong to $\mathcal{U}(K)$, then $f_3 = f_1 \land f_2$ belongs to $\mathcal{U}(K)$ and is smaller than both of them. It follows that the correspondence
\[
f \in \mathcal{L}(U) \mapsto G(f) \in \mathbb{R}^+
\]
is an increasing net and the correspondence
\[
f \in \mathcal{U}(K) \mapsto G(f) \in \mathbb{R}^+
\]
is a decreasing net; we have
\[
\sup\{G(f) : f \in \mathcal{L}(U)\} = \lim_{f \in \mathcal{L}(U)} G(f) \in \mathbb{R}^+ \cup \{\infty\} \quad \text{and}
\]
\[
\inf\{G(f) : f \in (K)\} = \lim_{f \in \mathcal{U}(K)} G(f) \in \mathbb{R}^+.
\]

**Definition:** For a given positive linear functional $G : \mathcal{K}(X) \rightarrow \mathbb{R}$, define the $\mathbb{R} \cup \{\infty\}$-valued set function $\nu_G$ on open sets by
\[
\nu_G(U) = \sup\{G(f) : f \in \mathcal{L}(U)\}
\]
and on compact sets by
\[
\nu_G(K) = \inf\{G(f) : f \in \mathcal{U}(K)\} \quad \text{(where its value is necessarily } < \infty).\]
There is no ambiguity here: if $K = U$ is both compact and open then $\mathcal{U}(K)$ has a largest element and $\mathcal{U}(K)$ a smallest element: $\chi_K$ is both (it is continuous because $K$ is both open and closed), and the common value given by the definition is $\nu_G(K) = G(\chi_K)$. Just to be on the safe side, we shall explicitly put $\nu_G(\emptyset) = 0$.
It is obvious that $\nu_G$ is monotone (increasing) on open sets and monotone (increasing) on compact sets (we shall see momentarily that it is monotone on both kinds of sets).

In situations where only one linear functional $G$ is in the context, we shall write simply $\nu$ instead of $\nu_G$.

We now start grinding out a sequence of elementary propositions whose joint result will be to let us extend $\nu_G$ from the open and compact sets to an “outer measure” like Lebesgue $m^*$, and then to restrict the extension to the class of “sets that split outer measure additively,” the “measurable sets for this outer measure.”

**Proposition:** If $K$ is compact, $U$ is open, and $K \subseteq U$, then $\nu(K) \leq \nu(U)$.

*Proof.* As we saw on pp. 00–00 above, there is an $f \in \mathcal{X}(X)$ with the properties that $0 \leq f \leq 1$, $\text{Supp}(f) \subseteq U$, and $f[K] = \{1\}$. So $\nu(K) \leq G(f) \leq \nu(U)$ in view of the definition of $\nu$ on compact and on open sets respectively.

**Proposition:** For any open $U \subseteq X$ one has

$$\nu(U) = \sup \{ \nu(K) : K \text{ compact}, K \subseteq U \}.$$

*Proof.* We just saw that the inequality $\geq$ holds in place of the $=$ of the assertion. To get the inequality in the opposite direction, take any real $a < \nu(U)$; then by definition of $\nu(U)$ there exists $f \in \mathcal{L}(U)$ with $G(f) > a$. Take $K_0 = \text{Supp}(f)$. If $h \in \mathcal{U}(K_0)$ then $h \geq f$, so $G(h) \geq G(f) > a$; taking the infimum on $h$, we get $\nu(K_0) > a$. This implies $a < \sup \{ \nu(K) : K \text{ compact}, K \subseteq U \}$ and since $a < \nu(U)$ was arbitrary, we have $\nu(U) \leq \sup \{ \nu(K) : K \text{ compact}, K \subseteq U \}$.

**Proposition:** For any compact $K \subseteq X$, one has

$$\nu(K) = \inf \{ \nu(U) : U \text{ open}, U \supseteq K \}.$$

*Proof.* Here the inequality $\leq$ in place of the asserted equality is immediate. For the reverse inequality, take any $\epsilon > 0$ and find $f \in \mathcal{U}(K)$ with $G(f) < \nu(K) + \epsilon$. The set

$$U(\epsilon) = \{ t \in X : (1 + \epsilon) \cdot f(t) > 1 \}$$

is an open neighborhood of $K$. If $h \in \mathcal{L}(U(\epsilon))$ then

$$(1 + \epsilon) \cdot f \geq h$$

$$(1 + \epsilon) \cdot G(f) \geq G(h)$$

$$(1 + \epsilon) \cdot (\nu(K) + \epsilon) \geq G(h) \quad \text{and since } h \in \mathcal{L}(U(\epsilon)) \text{ was arbitrary}$$

$$(1 + \epsilon) \cdot (\nu(K) + \epsilon) \geq \nu(U) \geq \inf \{ \nu(U) : U \text{ open}, U \supseteq K \}.$$

The l. h. s. of this inequality can be as close to $\nu(K)$ as we wish, from above; so $\nu(K) \geq \inf \{ \nu(U) : U \text{ open}, U \supseteq K \}$.

**Proposition:** $\nu$ is finitely additive on disjoint open sets: if $U_1$ and $U_2$ are open and $U_1 \cap U_2 = \emptyset$, then $\nu(U_1 \cup U_2) = \nu(U_1) + \nu(U_2)$.

*Proof.* A moment’s reflection will convince the reader that if $U_1$ and $U_2$ are disjoint, then $\mathcal{L}(U_1 \cup U_2) = \mathcal{L}(U_1) + \mathcal{L}(U_2)$ (the set of all sums $f_1 + f_2$ where each term belongs to the corresponding $\mathcal{L}(U_j)$). Containment in the $\supseteq$ direction is obvious. Conversely, if $f \in \mathcal{L}(U_1 \cup U_2)$, then $\text{Supp}(f) \cap U_1$ is a compact subset of $U_1$, because it equals $\text{Supp}(f) \setminus U_2$. If we define $f_1$ to equal $f$ on $U_1$ and 0 elsewhere, then $f_1$ is continuous since $f$ already was zero on $U_1 \setminus (\text{Supp}(f) \setminus U_2)$. Defining $f_2$ similarly ($= f$ on $U_2$ and 0 elsewhere), we have $f = f_1 + f_2$, and this shows (since $f \in \mathcal{L}(U_1 \cup U_2)$ was arbitrary) that $\mathcal{L}(U_1 \cup U_2) \subseteq \mathcal{L}(U_1) + \mathcal{L}(U_2)$. But now

$$\nu(U_1 \cup U_2) = \lim \{ G(f) : f \in \mathcal{L}(U_1 \cup U_2) \} = \lim \{ G(f_1 + f_2) : f_1 \in \mathcal{L}(U_1) \}$$

$$= \lim \{ G(f_1) : f_1 \in \mathcal{L}(U_1) \} + \lim \{ G(f_2) : f_2 \in \mathcal{L}(U_2) \}$$

$$= \nu(U_1) + \nu(U_2)$$
by the continuity of addition (even in $\mathbb{R} \cup \{\infty\}$).

**Proposition:** $\nu$ is countably subadditive on open sets: if $U = \bigcup_{j=1}^{\infty} U_j$ where all the $\{U_j\}_{j=1}^{\infty}$ are open, then $\nu(U) \leq \sum_{j=1}^{\infty} \nu(U_j)$.

**Proof.** Consider any finite subfamily $\{U_1, \ldots, U_k\}$ of these open sets. If $f_j \in \mathcal{L}(U_j)$ for $1 \leq j \leq k$, then $(f_1 + \cdots + f_k) \wedge 1 \in \mathcal{L}(U)$. Moreover, the family of functions that can be defined in this way is directed upward; if another such function is $(g_1 + \cdots + g_k) \wedge 1$, then

$$(f_1 \vee g_1) + \cdots + (f_k \vee g_k) \wedge 1$$

is such a function and majorizes both the given functions. The family of all such functions is thus upward-directed.

Let $K \subseteq U$ be compact; then $K \subseteq (U_1 \cup \cdots \cup U_k)$ for some finite subfamily of the $\{U_j\}_{j=1}^{\infty}$. If $x \in K$ belongs to a particular $U_j$, then by the lemma on pp. 00–00 there is an $f_j \in \mathcal{L}(U_j)$ with $f_j(x) = 1$. The restrictions of the family of functions

$$\{(f_1 + \cdots + f_k) \wedge 1 : f_j \in \mathcal{L}(U_j), j = 1, \ldots, k\}$$

to $K$ thus form an increasing net (“indexed by itself”) converging pointwise to 1 on $K$. By the Dini monotone convergence theorem (see pp. 0–0 of the lemmas), the convergence is uniform on $K$, and therefore for any $\epsilon > 0$ we can find $f_j \in \mathcal{L}(U_j)$, $1 \leq j \leq k$, for which

$$((f_1 + \cdots + f_k) \wedge 1)(t) > \frac{1}{1+\epsilon} \quad \text{for every } t \in K.$$  

We then have

$$(1+\epsilon) \cdot (f_1 + \cdots + f_k) \in \mathcal{L}(K)$$

and so

$$\nu(K) \leq (1+\epsilon) \cdot G(f_1 + \cdots + f_k) \leq (1+\epsilon) \cdot (\nu(U_1) + \cdots + \nu(U_k)).$$

Since $K$ did not depend upon the choice of $\epsilon > 0$, 

$$\nu(K) \leq \nu(U_1) + \cdots + \nu(U_k) \leq \sum_{j=1}^{\infty} \nu(U_j)$$

and since the r. h. s. of this does not depend on $K$, taking the sup on $K$ gives the desired

$$\nu(U) \leq \sum_{j=1}^{\infty} \nu(U_j).$$

**Corollary:** $\nu$ is countably additive on disjoint open sets.

**Proof.** If the $\{U_j\}_{j=1}^{\infty}$ are disjoint, then the finite additivity and monotonicity of $\nu$ give us

$$\nu(U) \geq \nu \left( \bigcup_{j=1}^{k} U_j \right) = \sum_{j=1}^{k} \nu(U_j)$$

for every $k$. Consequently

$$\nu(U) \geq \sum_{j=1}^{\infty} \nu(U_j)$$

and the known countable subadditivity gives the reverse inequality.
At this point, to see what we have constructed so far, it may be instructive to look at what the definitions mean for Riemann-Stieltjes integrals on a finite interval.

**Proposition:** If \( X = [a, b] \) and \( G(f) = \int_a^b f(t) \, dg(t) \), where \( g \) is an increasing real-valued function on \([a, b]\), then

\[
\begin{align*}
  &\text{For } a < c < b, \quad \nu((c, d)) = g(d) - g(c) ; \\
  &\text{for } a < d < b, \quad \nu([a, d)) = g(d) - g(a) ; \\
  &\text{for } a < c < b, \quad \nu((c, b]) = g(b) - g(c) .
\end{align*}
\]

where \( g(d^-) \) and \( g(c^+) \) are the limits from the left and right respectively of \( g(t) \) at \( d \) and \( c \) respectively.

**Proof.** If \( f \in L((c, d)) \), then since \( \text{Supp}(f) \subseteq (c, d) \) there is a closed interval \([p, q] \subseteq (c, d)\) that contains \( \text{Supp}(f) \) in its interior. Therefore

\[
G(f) = \int_a^b f(t) \, dg(t) \leq \|f\|_\infty \cdot V[g; p, q] = g(q) - g(p) \leq g(d) - g(c^+)
\]

and taking the supremum gives

\[
\nu((c, d)) \leq g(d^-) - g(c^+) .
\]

On the other hand, for \( c < p < q < d \) with \( p \) and \( q \) arbitrarily close to \( c \) and to \( d \) from the right and left respectively, one can find functions \( f \) satisfying \( 0 \leq f \leq 1, \text{Supp}(f) \subseteq (c, d) \), and \( f(t) = 1 \) for \( p \leq t \leq q \). Suppose that \( \Gamma \) is a partition of \([a, b]\) containing the points \( p \) and \( q \), say

\[
\Gamma : a = t_0 < \cdots < t_r = p < \cdots < t_s = q < \cdots < t_k = b .
\]

Then any Riemann-Stieltjes sum \( R(f, g; \Gamma, \tau) \) will have only nonnegative terms and will include among them a block

\[
\sum_{j=r+1}^{s} 1 \cdot [g(t_j - g(t_{j-1})] = g(q) - g(p) .
\]

So \( (f; g; \Gamma, \tau) \geq g(q) - g(p) \); in the limit \( \int_a^b f(t) \, dg(t) \geq g(q) - g(p) \). By definition of \( \nu \), this gives \( \nu((c, d)) \geq g(q) - g(p) \), and the r. h. s. of that can be as close to \( g(d^-) - g(c^+) \) as we wish. The half-open interval cases are proved similarly. Since any open set in \([a, b]\) is a union of countably many open intervals and \( \nu \) is countably additive, we now know (in principle) how to compute \( \nu(U) \) for any open \( U \).

If you believe that \( g(t) \) is correctly interpreted as “the amount of mass contained in the interval \([a, t]_c \),” then the calculations above should give the amount of mass in the interval \((c, d)\). Computing the amount of mass in a compact set is harder—their structure is more varied. However, for very simple compact sets the answer is easy:

**Corollary:** If \( X = [a, b] \) and \( G(f) = \int_a^b f(t) \, dg(t) \), where \( g \) is an increasing real-valued function on \([a, b]\), then

\[
\begin{align*}
  &\text{For } a < c < b, \quad \nu([c]) = g(c^+) - g(c^-) , \\
  &\nu([a]) = g(a^+) - g(a) , \quad \text{and} \quad \nu([b]) = g(b) - g(b^-) .
\end{align*}
\]

**Proof.** For \( a < c < b \), we have

\[
\nu([c]) = \inf \{ \nu(U) : c \in U , \ U \text{ open} \} = \inf \{ \nu((c - \epsilon, c + \epsilon) : \epsilon > 0 \} \text{ by monotonicity} = g(c^+) - g(c^-) \text{ by the proposition} .
\]

The other cases are proved similarly.
What we now do is start reading through Royden’s Ch. 3, starting at §2, p. 56, and imitating the
construction of Lebesgue measure. We begin with

**Definition:** The outer $\nu_G$ measure of sets in $X$ is defined by

$$
\nu^*_G(A) = \inf \left\{ \sum_{j} \nu_G(U_j) : \text{all } U_j \text{ open, } A \subseteq \bigcup_{j} U_j \right\}.
$$

As usual, we shall drop the subscript “$G$” when only one linear functional occurs in the context.

$\nu^*$ is well-defined, since $X$ is open. It is immediate that $\nu^*(\emptyset) = 0$ and that $A \subseteq B \Rightarrow \nu^*(A) \leq \nu^*(B)$.

The first thing we should check is that we have not introduced any surprises.

**Proposition:** For all $A \subseteq X$, $\nu^*(A) = \inf \{ \nu(U) : U \text{ open, } A \subseteq U \}$.

*Proof.* Since the r. h. s. of this uses only a subclass of the class of countable covers, clearly $\nu^*(A) \leq \inf \{ \nu(U) : U \text{ open, } A \subseteq U \}$. But by countable subadditivity of $\nu$ on open sets, for any covering of $A$ by a countable family of open sets we have $\nu \left( \bigcup_{j} U_j \right) \leq \sum_{j} \nu(U_j)$ and thus

$$
\inf \{ \nu(U) : U \text{ open, } A \subseteq U \} \leq \inf \left\{ \sum_{j} \nu(U_j) : \text{all } U_j \text{ open, } A \subseteq \bigcup_{j} U_j \right\} = \nu^*(A).
$$

**Corollary:** For open $U$ and compact $K$, $\nu(U) = \nu^*(U)$ and $\nu(K) = \nu^*(K)$.

**Proposition:** $\nu^*$ is countably subadditive on $2^X$.

*Proof.* Just like Royden’s Ch. 3, §2, Prop. 2, p. 57; or you can simplify it even further using the countable subadditivity of $\nu$ on open sets.

The following definition can hardly come as a surprise.

**Definition:** A set $E$ is said to be $\nu$-measurable if for each $A \subseteq X$ we have

$$
\nu^*(A) = \nu^*(A \cap E) + \nu^*(A \setminus E).
$$

It suffices to check that $\nu^*(A) \geq \nu^*(A \cap E) + \nu^*(A \setminus E)$ for every $A$ for which $\nu^*(A) < \infty$. Less will suffice, as we shall see shortly.

As with Lebesgue measure, $\nu^*$-null sets—sets $E$ with $\nu^*(E) = 0$—are automatically measurable. Just as with Royden’s propositions and theorems 7 through 10, Ch. 3, §2, the $\nu$-measurable sets form a $\sigma$-algebra. (Cf. also Ch. 12, §1, p. 288 ff.). However, since we’re working in a fairly general class of topological spaces $X$, there will be a replacement for Royden’s Lemma 11 (p. 60) that makes Borel sets measurable.

**Lemma** {cf. Royden, Ch. 13, §3, p. 346}: In order that $E \subseteq X$ be $\nu$-measurable it is necessary and sufficient that the relation

$$
\nu^*(U) \geq \nu^*(U \cap E) + \nu^*(U \setminus E)
$$

hold for all open sets $U \subseteq X$ with $\nu^*(U) < \infty$.

*Proof.* Of course the condition is necessary. To see that it is sufficient, suppose $\nu^*(A) < \infty$, let $\epsilon > 0$ be given, and find open $U \subseteq A$ with $\nu(U) \leq \nu^*(A) + \epsilon$. We have $A \cap E \subseteq U \cap E$, $A \setminus E \subseteq U \setminus E$, and thus

$$
\nu^*(A \cap E) + \nu^*(A \setminus E) \leq \nu^*(U \cap E) + \nu^*(U \setminus E) \leq \nu^*(U) \leq \nu^*(A) + \epsilon
$$

under the assumption that “$E$ splits open sets additively.” Since $\epsilon > 0$ is in our hands, this gives

$$
\nu^*(A \cap E) + \nu^*(A \setminus E) \leq \nu^*(A)
$$
Thus \( K \) at least for Borel sets. This would give us the analogues of Royden’s Ch. 3 approximators, produces no surprises. We would like to show that this property persists for measurable, or according to the plan above has “inner- and outer-regularity” on open and compact sets: the measure of their subsets are desirable immediately to restrict \( \nu \) to the \( \sigma \)-algebra of Borel sets. For example, if the class of \( \nu \)-measurable sets will in general be larger than that of Borel sets, it can be undesirable to use all the sets of the former class. For example, if \( X = [0, 1] \) and \( g = \chi_{[0.1/2]} \), then the calculation on pp. 00–00 above shows us that \([0,1/2)\) and \((1/2,1]\) are null sets for \( \nu_G \)—and therefore all their subsets are \( \nu_G \)-measurable—so every subset of \([0,1]\) is measurable: \( \nu_G \) assigns the mass 0 to every subset that does not contain \( 1/2 \) and 1 to every subset that does contain it. This is not necessarily a good arrangement—e.g., there are many sets for which it is impossible to compare their \( \nu_G \)- and Lebesgue measures. So in this context it is desirable immediately to restrict \( \nu_G \) to the \( \sigma \)-algebra of Borel sets, and to think of it only as a Borel measure.

The three little propositions on pp. 00–00 above show that any \( \nu_G \) (henceforth just called \( \nu \)) constructed according to the plan above has “inner- and outer-regularity” on open and compact sets: the measure of the open sets from the inside using compact approximators, or of compact sets from the outside using open approximators, produces no surprises. We would like to show that this property persists for measurable, or at least for Borel sets. This would give us the analogues of Royden’s Ch. 3 §3 Prop. 15, p. 63 for measurable sets, and as a result analogues of the “measurable functions are almost continuous” results of Ch. 3 §5, especially Prop. 22 and Prob. 23, pp. 69–71. Unfortunately these are not true in great generality; one has to be a bit careful, though not for the kind of spaces \( X \) in which most of analysis is done.

Here, first, is a positive result. Remember that \( \nu \) has been constructed from a positive linear functional \( G \) as described above.

**Proposition:** If \( E \subseteq X \) is a \( \nu \)-measurable set of finite measure, then given any \( \epsilon > 0 \) there exist an open \( U \) and a compact \( K \) with \( K \subseteq E \subseteq U \) and \( \nu(U \setminus K) < \epsilon \).

**Proof.** By definition of \( \nu(E) \) there is some open \( U \supseteq E \) with \( \nu(U) < \nu(E) + \epsilon/2 \), and again there is an open \( V \supseteq (U \setminus E) \) with \( \nu(V) < \epsilon/2 \) (since \( \nu(U \setminus E) = \nu(U) - \nu(E) < \epsilon/2 \)). By “inner regularity” of \( U \) we can now find compact \( F \subseteq U \) for which \( \nu(U) < \nu(F) + \epsilon/2 \). Evidently \( F \setminus V \) is compact (as a closed subset of a compact set) and \( (F \setminus V) \subseteq (F \cap U) \cup (F \cap E) = (F \cap E) \subseteq E \). We then have

\[
\nu(U \setminus (F \setminus V)) = \nu(U \cap (F \cap V)) = (U \cap (F \cap V) = (U \setminus F) \cup (U \cap V) \\

\nu(U \setminus (F \setminus V)) \leq \nu(U \setminus F) + \nu(U \cap V) \\
\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon .
\]

Thus \( K = F \setminus V \) has the properties we want.
Corollary: If $E$ is a countable union of $\nu$-measurable sets of finite measure, then for any $\epsilon > 0$ there exist open $U$ and $\sigma$-compact $A$ with $A \subseteq E \subseteq U$ and such that $\nu(U \setminus A) < \epsilon$.

Proof. Write $E$ as a disjoint union (“disjointizing” in the usual way if necessary) $E = \bigcup_j E_j$ where each $E_j$ has finite measure, and find open $U_j$ and compact $K_j$ with $K_j \subseteq E_j \subseteq U_j$ and $\nu(U_j \setminus K_j) < \frac{\epsilon}{2^j}$. Then $A = \bigcup_j K_j \subseteq E \subseteq \bigcup_j U_j = U$ and we have

$$U \setminus A = \left( \bigcup_j U_j \right) \setminus \left( \bigcup_j K_j \right) \subseteq \bigcup_j (U_j \setminus K_j)$$

$$\nu(U \setminus A) \leq \sum_j \nu(U_j \setminus K_j) < \epsilon.$$ 

Corollary: If $E$ is a countable union of $\nu$-measurable sets of finite measure, then there exist a $G_\delta$-set $H \subseteq E$ and a $\sigma$-compact set $B \subseteq E$ with $\nu(H \setminus B) = 0$.

Proof. Let $\epsilon = \frac{1}{n}$ in the corollary above, find a sequence of open $U_n$’s and a sequence of $K_\sigma$-sets (countable unions of compact sets) $A_n$, let $H$ be the intersection of the $U_n$’s and $B$ be the union of the $A_n$’s.

So if $X$ itself is a countable union of sets (that condition is independent of $G$) or a countable union of sets of finite $\nu$-measure (that condition does depend on $G$) then all the mass of $\nu$ is concentrated on a countable family of compact sets and the situation is very much like that of Lebesgue measure on the line. (For the situation that got us interested in this in the first place—the measure obtained from the Stieltjes integral on $[a, b] \subseteq \mathbb{R}$—$X$ is compact and there aren’t even any sets of infinite measure). Here, however, is an example of what can happen: suppose $X$ is the disjoint union of uncountably many copies of the unit interval with Lebesgue measure (the reader should be able to come up with a functional on $\mathcal{K}(X)$ that produces this measure). Let $E$ be the set consisting of the point “$1/2$” in each of those intervals. Any open set containing $E$ has to contain uncountably many open intervals and therefore have measure $\infty$, so $\nu(E) = \infty$. However, any compact subset of $E$ is finite (the copies of $[0, 1]$ are open sets, so a compact subset of $X$ is contained in only finitely many of them) so its measure is $0 = \sup \{\nu(K) : K \subseteq E \} = 0$. Now it turns out that this example is not very pathological and that there are ways—for this space $X$—to finesse the corollary above in situations in which we think we need it, but don’t have it. Nonetheless, things will not be as nice in general as they are in the case of Lebesgue measure.

In any event, we can define $\nu$-measurable real-valued functions on $X$ in the same way as Lebesgue-measurable real-valued functions are defined on $\mathbb{R}$ {functions with their values in $\mathbb{R}^n$, if we should want to use them, would provide no problems to speak of}. We cannot prove a proposition like Royden’s Ch. 3, §5, Prop. 22 (p. 69) in general, because there is no way to say “step function” in a general $X$; however, if we got $\nu_G$ on $[a, b]$ from a monotone-increasing Stieltjes integrator $g$, then we could prove that proposition with Lebesgue measure replaced by $\nu_G$. The definition of the Lebesgue integral for bounded functions on a set of finite measure of Royden, Ch. 4, §2 uses only basic notions of measurability and uses simple functions, not step functions, except for Prop. 4 (p. 81) which says that the Lebesgue integral (as so far defined for bounded measurable functions on a measurable set of finite measure) agrees with the Riemann integral.

There is an analogue in the present situation, and we need to examine it and prove it. We got $\nu_G$ out of a linear functional $G : \mathcal{K}(X) \to \mathbb{R}$ in the hope that we would be able to write

$$G(f) = \int f(t) \, d\nu_G(t) \text{ for } f \in \mathcal{K}(X),$$

and that now needs to be checked. There is a “long way” to do this which I shall give as an exercise; here is the fast and brutal proof of Royden’s Ch. 13, §4, Theorem 23, pp. 353–354; only the notation is slightly
altered to conform to the practice I started in these notes (and avoid Royden’s use of the letter “O” for open sets—the confusion with “0” is too easy to make).

**Theorem:** If $\nu = \nu_G$ is the measure constructed from $0 \leq G : \mathcal{X}(X) \to \mathbb{R}$ as above, then $G(f) = \int f(t) \, d\nu(t)$ for all $f \in \mathcal{X}(X)$.

**Proof.** Because $G$ is linear it suffices to prove the assertion under the further assumption that $0 \leq f \leq 1$. Fix $n \in \mathbb{N}$ and for $j = 0, \ldots, n$ put

$$f_j = \left( \frac{j-1}{n} \right) \wedge \frac{1}{n} \vee 0 .$$

Each of these is a continuous function and satisfies $0 \leq f_j \leq \frac{1}{n}$; if

$$U_0 = \text{some open set with } \text{Supp}(f) \subseteq U_0 \text{ and } \overline{U_0} \text{ compact,}$$

$$U_j = \left\{ t \in U_0 : f(t) > \frac{j-1}{n} \right\} \text{ for } 1 \leq j \leq n+1 \text{ (so } U_{n+1} = \emptyset) ,$$

$$f_j(t) = \frac{1}{n} \text{ if } t \in U_{j-1} ,$$

$$f_j(t) = f(t) - \frac{j-1}{n} \text{ if } t \in U_j \setminus U_{j-1} ,$$

$$f_j(t) = 0 \text{ if } t \in X \setminus U_j , \text{ so } \text{Supp}(f_j) \subseteq \overline{U_j} \subseteq U_{j-1} .$$

By looking at the cases $t \in U_j \setminus U_{j+1}$ we can easily see that $f = \sum_{j=1}^{n} f_j$. For each $j$ we have $n \cdot f_j \in \mathcal{L}(U_{j-1})$ and $n \cdot f_j \in \int(\overline{U_{j+1}})$; this gives

$$\frac{1}{n} \cdot \nu(U_{j+1}) \leq G(f_j) \leq \frac{1}{n} \cdot \nu(U_{j-1}) .$$

Similarly

$$\frac{1}{n} \cdot \nu(U_{j+1}) \leq \int f_j(t) \, d\nu(t) \leq \frac{1}{n} \cdot \nu(U_j) .$$

This gives for each $j = 1, \ldots, n$

$$\frac{1}{n} \cdot [\nu(U_{j+1}) - \nu(U_j)] \leq G(f_j) - \int f_j(t) \, d\nu \leq [\nu(U_{j-1}) - \nu(U_{j+1})] .$$

Adding these up gives a telescoping series on the left, the sums of the differences in the middle, and on the right two telescoping series ($U_{n+1} = \emptyset$):

$$-\frac{1}{n} \cdot \nu(U_1) \leq \sum_{j=1}^{n} G(f_j) - \sum_{j=1}^{n} \int f_j(t) \, d\nu(t) \leq \frac{1}{n} \cdot [\nu(U_0) + \nu(U_1) - \nu(U_n)] .$$

Adding up the $G$’s and the integrals gives us $G(f)$ and $\int f(t) \, d\nu(t)$ respectively, and an obvious overestimate of the sums on the ends gives

$$|G(f) - \int f(t) \, d\nu(t)| \leq \frac{2}{n} \cdot \nu(U_0) .$$

Since $n \to \infty$ is in our hands,

$$G(f) = \int f(t) \, d\nu(t)$$

as desired.

We can now get a sufficient condition for these integrals to exist in sense (II) that parallels the theorem on p. 30 above.

**Theorem:** A function \( f : [a, b] \to \mathbb{R} \) is Riemann-Stieltjes integrable in sense (II) with respect to an increasing integrator \( g \) if it is continuous \( \nu_G \)-almost everywhere.

**Proof.** We are assuming that the set of discontinuities of \( f \) has \( \nu_G \)-measure zero. For any \( \epsilon > 0 \), then, the set

\[
E(\epsilon) = \{ t \in [a, b] : \omega(f; t) \geq \epsilon \}
\]

has \( \nu_G \)-measure zero, and it is a closed and therefore compact set; hence there is a finite family \( \{ U_r : 1 \leq r \leq s \} \) of open intervals with \( \nu_G(U_r) < \epsilon \) that covers it. On the other hand, if \( x \in [a, b] \setminus E(\epsilon) \), then there is an open interval \( V(x) \in \mathfrak{R}(x) \) on which the oscillation of \( f \) is \( \epsilon \). The union of the families of \( \{ U_r \} \) and \( \{ V_x \} \) is then an open covering of \( [a, b] \). Let \( \delta > 0 \) be a Lebesgue number of this covering. If \( \Gamma : a = t_0 < \cdots < t_k = b \) is a partition of \( [a, b] \) with \( \text{mesh}(\Gamma) < \delta \), then each cell of the partition is a subset either of some \( U_r \) or some \( V_x \). At this point we must be a little bit more careful than we were on p. 30 above. Consider a particular \( U_r \) that contains some elements of the partition \( \Gamma \). We can group these elements into connected components of their union, i.e., families containing all the cells for, say, \( t_p < \cdots < t_q \). For the sum of terms \( \omega_j(f; \Gamma) \cdot [g(t_j) - g(t_{j-1})] \) in such a family we have

\[
\sum \omega_j(f; \Gamma) \cdot [g(t_j) - g(t_{j-1})] \leq 2 \cdot \| f \|_\infty \cdot [g(t_q) - g(t_p)].
\]

The various such intervals \( [t_p, t_q] \) in \( U_r \) are disjoint, so there exist disjoint open intervals containing them; the sum of the \( \nu_G \)-measures of those disjoint open intervals must be \( \leq \nu_G(U_r) \). Since each of those open intervals contains \( [t_p, t_q] \) its measure is \( \geq g(t_q) - g(t_p) \), so the sum of the \( [g(t_q) - g(t_p)] \)'s from all the groups in \( U_r \) is \( \leq \nu_G(U_r) \), and so finally

\[
\sum \omega_j(f; \Gamma) \cdot [g(t_j) - g(t_{j-1})] : [t_{j-1}, t_j] \subseteq U_r \text{ for some } r \leq 2 \cdot \| f \|_\infty \cdot \epsilon.
\]

On the other hand, each of the remaining cells of the partition is contained in some \( V_x \), so the oscillation of \( f \) on such a cell is \( \epsilon \), and thus

\[
\sum \omega_j(f; \Gamma) \cdot [g(t_j) - g(t_{j-1})] < \epsilon \cdot [g(b) - g(a)]
\]

where the sum is taken over all the remaining cells. We thus have

\[
\sum_{j=1}^{k} \omega_j(f; \Gamma) \cdot [g(t_j) - g(t_{j-1})] < 2 \cdot [\| f \|_\infty + g(b) - g(a)] \cdot \epsilon
\]

which can be as small as we please. Thus the sufficient condition

\[
\lim_{\text{mesh}(\Gamma) \to 0} \sum_{j=1}^{k} \omega_j(f; \Gamma) \cdot V[g; t_{j-1}, t_j] = 0
\]

for the existence of integral in sense (II) is satisfied.

Unfortunately, we need to do some rather dreary uniqueness-checking. We shall do this for Stieltjes integrals here and for the representing measures of positive linear functionals in §7.

First of all, suppose that \( g_1 \) and \( g_2 \) are two increasing functions on \( [a, b] \) for which the two linear functionals \( f \mapsto \int f(t) \, dg_j(t) \) \( (j = 1, 2) \) agree. Then the two measures \( \nu_1 \) and \( \nu_2 \) that would be constructed
from them by the method of §5 above are the same measure. On pp. 00–00 above, we observed that then we had
\[ \nu([a, x)) = g(x^-) - g(a) \]
\[ \nu([x)) = g(x^+) - g(x^-) \]
\[ \nu([a, b]) = g(b) - g(a) \]
for any \( x \in [a, b] \). If the two functions defined the same linear functional on the space \( C([a, b]) \), then the l. h. sides of these two equations are the same for \( g_1 \) and \( g_2 \) and therefore so are the r. h. sides: we thus have \( g_1 \) and \( g_2 \) are continuous (or discontinuous) at the same points;
\[ g_1(x) - g_1(a) = g_2(x) - g_2(a) \text{ at every point of continuity (of both);} \]
\[ g_1(x^+) - g_1(x^-) = g_2(x^+) - g_2(x^-) \text{ at every point of discontinuity (of both).} \]
In other words, up to a constant the two functions are equal at every point of continuity, and where their left- and right-hand limits differ their jumps are equal. The converse of this is almost obvious: if this condition holds then the measures corresponding to the two functions will have
\[ \nu_1((a, c)) = g_1(c^-) - g_1(a) = g_2(c^-) - g_2(a) = \nu_2((a, c)) \]
\[ \nu_1((c, d)) = g_1(d^-) - g_1(c^+) = g_2(d^-) - g_2(c^+) = \nu_2((c, d)) \]
so they will assign the same measure to every (relatively) open set in \([a, b]\)—and then they must be the same measure. On top of that, since (as we saw on pp. 00–00 above) integrating against the measure gives back the linear functional, we have
\[ \int_a^b f(t) \, dg_1(t) = \int_a^b f(t) \, dg_2(t) \]
for every \( f \in C([a, b]). \) {Of course that could have been demonstrated by elementary arguments involving the Riemann-Stieltjes sums, but this approach is faster.}

Penultimatey, we should check that if \( \mu \) is a nonnegative Borel measure on \([a, b]\) with \( \mu([a, b]) < \infty \) and we define a linear functional on \( C([a, b]) \) by
\[ f \mapsto \int f(t) \, d\mu(t) \]
then (1) this linear functional can also be given by Riemann-Stieltjes integration against an appropriate integrator and (2) the measure obtained from that functional by the process of §5 above will again be \( \mu \). It’s easier to check (2) first, and the method is the same as that of pp. 00–00 above. If \( 0 \leq f \leq 1 \) and \( \text{Supp}(f)(c, d) \), then
\[ \int f(t) \, d\mu(t) \leq \chi(c,d)(t) \, d\mu(t) = \mu((c, d)) \]
so \( \nu((c, d)) \leq \mu((c, d)) \). On the other hand, whatever the functional is, for \( a \leq c < p < q < d \leq b \) with \( p \) and \( q \) arbitrarily close to \( c \) and \( d \), one can find functions \( f \in C([a, b]) \) with \( 0 \leq f \leq 1 \), \( \text{Supp}(f) \subseteq (c, d) \), and \( f(t) = 1 \) for \( p \leq t \leq q \). If \( \nu \) is the measure constructed from the functional as in §5, we shall then have
\[ \mu((p, q)) = \int \chi_{(p,q)}(t) \, d\mu(t) \leq \int f(t) \, d\mu(t) \leq \nu((c, d)) \]
and by letting \( p \) decrease to \( c \) through a countable set and \( q \) increase to \( d \) through a countable set, we find that \( \mu((c, d)) \leq \nu((c, d)) \) and thus that
\[ \mu((c, d)) = \nu((c, d)) \]
for every open interval. The same argument works for every half-open interval of the form \([a, c]\) or \((c, b]\). So \( \mu \) and \( \nu \) assign the same mass to every open interval, and therefore every open set, in \([a, b]\). By complementation,
the same is true for closed sets. It is straightforward to verify that on a metric space \( X \), every Borel measure with total mass \(< \infty\) has the property that the relations

\[
\mu(E) = \inf \{ \mu(U) : U \text{ open}, U \supseteq E \}
\]

\[
\mu(E) = \sup \{ \mu(F) : F \text{ closed}, F \subseteq E \}
\]

hold for every Borel set \( E \). \{This is Royden’s Ch. 15, Prop. 11, p. 407. He leaves the proof to the reader, and I will too—but it’s quite straightforward.\} Of course in \([a, b]\) “closed” means the same as “compact.” So for every Borel set \( E \) we have

\[
\mu(E) = \inf \{ \mu(U) : U \text{ open}, U \supseteq E \} = \inf \{ \nu(U) : U \text{ open}, U \supseteq E \} = \nu(E)
\]

and the functional has uniquely determined the measure.

Last thing on the agenda, then, is to show that given a nonnegative Borel measure on \([a, b]\) we can find a Stieltjes integrator that gives the same linear functional on \( \mathcal{C}([a, b]) \). There is a natural choice:

**Definition:** If \( \mu \) is a nonnegative finite (\( \mu(\mathbb{R}) < \infty \)) Borel measure on (some Borel subset of) \( \mathbb{R} \), its **cumulative function on** \( \mathbb{R} \) is the function defined by

\[
F(\mu; x) = \mu((\infty, x])
\]

Similarly, its **cumulative function on** \([a, b]\) is the function defined by

\[
F(\mu; x) = \mu([a, x]) : F(\mu; a) = 0
\]

(Yes, it is notationally the same, but which one is meant will always be [made] clear in the context.)

The right-closed interval is used so that the following will hold:

**Proposition:** \( F(; x) \) is continuous from the right, except that in the case of the cumulative function on \([a, b]\), it will have a jump of \( \mu(\{a}\) at \( a \).

**Proof.** It suffices to check sequential continuity, and if \( \{x_j\}_{j=1}^{\infty} \) is a sequence decreasing to \( x \), then

\[
\lim_{j \to \infty} F(\mu; x_j) = \lim_{j \to \infty} \mu((\infty, x_j]) = \\
\mu \left( \bigcap_{j=1}^{\infty} (\infty, x_j) \right) = \mu(\infty, x] = F(\mu; x).
\]

**Proposition:** If two finite Borel measures on \( \mathbb{R} \) or \([a, b]\) have the same cumulative function, then they are the same measure.

**Proof.** The \( \mu \)-measure of any interval \((c, d]\) can be computed as

\[
\mu(c, d] = F(\mu; d) - F(\mu; c)
\]

For any sequence \( \{d_j\}_{j=1}^{\infty} \) increasing to \( d \)

\[
\mu(c, d) = \lim_{j \to \infty} \mu(c, d_j] = F(\mu; d^-) - F(\mu; c) \text{ also.}
\]

It follows that \( F(\mu; x) \) determines the value of \( \mu \) on any open set, hence (because \( \mu(\mathbb{R}) < \infty \)) on any closed set, hence (by the propostion of Royden, p. 407 cited above) on any Borel set.

The same proof works for the cumulative function on \([a, b]\) with only minor attention to the l. h. endpoint.

Wouldn’t it be surprising, then, if it weren’t true that
**Proposition:** If \( \mu \) is a Borel measure on \([a, b]\), then for all \( f \in C([a, b])\)

\[
\int_a^b f(t) \, d\mu(t) = \int_a^b f(t) \, dF(\mu; t)
\]

where the integral on the right is Riemann-Stieltjes and \( F(\mu; x) \) is the cumulative function of \( \mu \) on \([a, b]\).

**Proof.** For any partition \( \Gamma : a = t_0 < \cdots < t_k = b \) of \([a, b]\), we have

\[
\int_{t_0}^{t_1} \chi_{[t_0, t_1]}(t) \, d\mu(t) = \mu([t_0, t_1]) = F(\mu; t_1) - F(\mu; t_0)
\]

\[
\int_{t_{j-1}}^{t_j} \chi_{(t_{j-1}, t_j]}(t) \, d\mu(t) = \mu((t_{j-1}, t_j]) = F(\mu; t_j) - F(\mu; t_{j-1}) \quad \text{for } 2 \leq j \leq k
\]

and if \( M_j(f; \Gamma) = \sup\{ f(t) : t_{j-1} \leq t \leq t_j \} \) as usual, then for the function

\[
h = M_1 \chi_{[a, t_1]} + \sum_{j=2}^k M_j \chi_{(t_{j-1}, t_j]}
\]

we have \( f \leq h \) and therefore

\[
\int_a^b f(t) \, d\mu(t) \leq \int_a^b h(t) \, d\mu(t) = \sum_{j=1}^k M_j \cdot [F(\mu; t_j) - F(\mu; t_{j-1})] = U(f; \Gamma)
\]

so \( \int_a^b f(t) \, d\mu(t) \leq \int_a^b f(t) \, dF(\mu; t) \) (Riemann-Stieltjes).

Dually we have \( \int_a^b f(t) \, dF(\mu; t) \leq \int_a^b f(t) \, d\mu(t) \). Since the Riemann-Stieltjes integral exists (even in sense (II)), the two inequalities are equalities and the Riemann-Stieltjes integral equals the integral \( d\mu \).

From what we saw above, it is clear that given an increasing Riemann-Stieltjes integrator \( g(t) \) on \([a, b]\), there is always a “right-continuous version” of the same integrator that gives the same integral against any continuous function on \([a, b]\) and differs by a constant from the cumulative function of the unique Borel measure that gives the same linear functional on the Banach space \( C([a, b]) \). {The right-continuous version must keep the jump \( g(a^+) - g(a) \) at the l. h. endpoint, of course.} It is possible to prove the following proposition about right-continuous integrators; we leave the proof as an exercise for the reader.

**Proposition:** Let \( g(t) \) be an increasing function on \([a, b]\) which is continuous from the right except perhaps at \( a \). Let \( \nu \) be the Borel measure on \([a, b]\) to which the functional on \( C([a, b]) \) defined by integrating continuous functions against \( g \) corresponds, so that \( g(x) = F(\nu; x) \) up to a constant. Let \( A \) be the (countable) set \( \{ x \in [a, b] : \nu(\{ x \}) > 0 \} \) and for Borel sets \( E \subseteq [a, b] \) put

\[
\nu_d(E) = \nu(E \cap A), \quad \nu_c(E) = \nu(E \setminus A).
\]

(A is for “atoms,” “d” is for discrete and “c” is for continuous.) Then in order that the Riemann-Stieltjes integral \( \int_a^b f(t) \, dg(t) \) exist in sense (I), it is sufficient that \( f \) be continuous from the right at each point of \( A \) and that the set of points of discontinuity of \( f \) be a null set for the measure \( \nu_c \).

Similarly, even without replacing \( g \) by its right-continuous version, we know that if \( f \) is continuous at each point of \( A \) and that the set of points of discontinuity of \( f \) is a null set for the measure \( \nu_c \), then the Riemann-Stieltjes integral will exist in sense (II); that statement is equivalent to the proposition on pp. 00–00 above (why?).
7. Regularity, Uniqueness, and the “Bourbaki Integral.”

In §5 above, we gave a construction that enabled one to extract a measure from a given nonnegative linear functional on $\mathcal{K}(T)$, where $T$ was a locally compact Hausdorff space; the functional then appeared as the integral with respect to the measure. It is important also to solve a converse problem: if a linear functional on $\mathcal{C}(T)$ is given by

$$F(f) = \int f(t) \, d\mu(t) \quad (\$)$$

where $\mu$ is a measure whose domain includes at least the Borel sets, can there be any measures $\rho$ different from $\mu$ for which (§) also holds—with $\mu$ replaced by $\rho$? There’s a rather delicate analysis of this question in Ch. 13 of Royden’s book, but in a first course we are best off concentrating on giving relatively simple additional conditions which will suffice to make the “representing measure of a linear functional” unique. The simplest conditions involve regularity of some kind.

**Definition**: Let $T$ be a locally compact Hausdorff space and let $\mu$ be a(n abstract) measure whose domain $\mathcal{M}$ is a $\sigma$-algebra containing the Borel sets. A set $E \in \mathcal{M}$ is said to be **outer** or **inner regular** respectively if it satisfies

$$\mu(E) = \inf \{ \mu(U) : U \text{ open}, U \supseteq E \} \quad \text{or} \quad \mu(E) = \sup \{ \mu(K) : K \text{ compact}, K \subseteq E \}$$

respectively. If $E$ has both these properties, $E$ is said to be **regular**. If every $E \in \mathcal{M}$ is outer regular, inner regular or regular respectively, then the pair $(\mu, \mathcal{M})$ is said to be **outer regular, inner regular or regular** respectively. For this definition cf. Royden, Ch. 13, §2, p. 337 ff.

**Proposition**: Suppose $T$ is a locally compact Hausdorff space and $\mu$ is an outer-regular “abstract measure” defined on a $\sigma$-algebra $\mathcal{M}$ containing the Borel sets. Suppose $\mu(K) < \infty$ for each compact $K \subseteq T$. Define a linear functional on $\mathcal{K}(T)$ by setting $G(f) = \int f(t) \, d\mu(t)$; let $\nu = \nu_G$. If every open set is inner regular for $\mu$, then every set in $\mathcal{M}$ is $\nu$-measurable and $\mu$ is equal to the restriction of $\nu^*$ to $\mathcal{M}$.

**Proof**. The fact that compact sets have finite $\mu$-measure insures that every continuous function of finite support is a bounded $\mu$-integrable function. For every open set $U$ and compact $K \subseteq U$ we can find $f \in \mathcal{L}(U)$ for which $\chi_K \leq f \leq \chi_U$, and so we have $\mu(K) \leq \int f(t) \, d\mu(t) \leq \mu(U)$ for each $f \in \mathcal{L}(U)$. Taking the supremum first with respect to $f$ and then with respect to $K$ we get $\mu(U) \leq \nu(U) \leq \mu(U)$, so $\nu$ and $\mu$ agree on open sets. Since we assumed that $\mu$ was outer regular, we then have for any $E \in \mathcal{M}$

$$\mu(E) = \inf \{ \mu(U) : U \text{ open}, U \subseteq E \} = \inf \{ \nu(U) : U \text{ open}, U \supseteq E \} = \nu^*(E) \, .$$

If $U$ is an open set of finite $\nu$-measure, then for any $E \in \mathcal{M}$ we have

$$\nu^*(U) = \mu(U) = \mu(U \cap E) + \mu(U \setminus E) \quad \text{because} \quad E \in \mathcal{M} \quad \text{and} \quad \mu \text{ is a measure}$$

$$= \nu^*(U \cap E) + \nu^*(U \setminus E) \quad \text{because} \quad \mu = \nu^* \text{ on } \mathcal{M}$$

so $E$ is (Carathéodory) measurable for $\nu$; we already knew that $\mu(E) = \nu^*(E) = \nu(E)$ and that concludes the proof.

It is now clear that, given a positive linear functional $G$ on $\mathcal{K}(T)$, the measure $\nu_G$ constructed from it in §5 above has maximal domain among the outer-regular measures that “represent $G$” in the sense that $\int f(t) \, d\mu(t) = G(f)$ for each $f \in \mathcal{K}(T)$ and for which each open set is inner regular; this is stronger than saying that the measure with those properties is uniquely determined by $G$.

If the space $T$ is compact, the two-way regularity hypotheses are superfluous:

**Proposition**: If $T$ is a compact Hausdorff space, then a measure $\mu$ defined on a $\sigma$-algebra containing the Borel sets of $T$ and such that $\mu(T) < \infty$ is outer regular if and only if it is inner regular.
Proof. The complement of an open set is a compact set, and $T$ itself has finite measure, so if we know that $E$ is inner regular then

$$
\mu(T \setminus E) = \mu(T) - \mu(E) = \mu(T) - \sup\{\mu(K) : K \text{ compact, } K \subseteq E\} \\
= \inf\{\mu(T) - \mu(K) : K \text{ compact, } K \subseteq E\} \\
= \inf\{\mu(T \setminus K) : K \text{ compact, } K \subseteq E\} \\
= \inf\{\mu(U) : U \text{ open, } U \supseteq T \setminus E\}
$$

and $T \setminus E$ is outer regular. Thus if $\mu$ is inner regular then it is outer regular; the converse argument is pretty much the same.

**Proposition:** If $T$ is a $\sigma$-compact metrizable space, then every Borel measure $\mu$ for which $\mu(K) < \infty$ for every compact set $K$ is both inner and outer regular.

Proof. Because $T$ is both $\sigma$-compact and locally compact, one can find a sequence $\{V_j\}_{j=1}^\infty$ of open sets with compact closure, such that

$$V_1 \subseteq \bigcap_{j=1} \subseteq \cdots \subseteq V_j \subseteq V_{j+1} \subseteq \cdots \text{ and } \bigcup_{j=1}^\infty V_j = T.$$

If for each $j$ we define a Borel measure $\mu_j$ by $\mu_j(E) = \mu(E \cap V_j)$, then it will be a finite measure and so

$$\mu_j(E) = \inf\{\mu_j(U) : U \text{ open, } U \supseteq E\} \quad \text{and} \quad \mu_j(E) = \sup\{\mu_j(F) : F \text{ closed, } F \subseteq E\}$$

will hold for every Borel set $E$. Consequently, given a Borel set $E$, if we put $E_k = E \cap \bigcap_{j=1}^k$ and $E_j = \bigcap_{j=1}^k \bigcap_{j=k+1}^\infty$ for $k \geq 2$, then given any $\epsilon > 0$ we can find open $U_j \subseteq V_{j+1}$ with $E_j \subseteq U_j$ and closed $F_j \subseteq E_j - F_j$ must actually be compact—with

$$\mu_j(U_j \setminus E_j) = \mu(U_j \setminus E_j) < \frac{\epsilon}{2^j+2}$$

$$\mu_{j+1}(E_j \setminus F_j) = \mu(E_j \setminus F_j) < \frac{\epsilon}{2^{j+2}}.$$ 

Setting $U = \bigcup_{j} U_j$ and $F = \bigcup_{j} F_j$, we then have $F \subseteq E \subseteq U$ and $\mu(U \setminus F) < \epsilon$. If $\mu(E) < \infty$ then also $\mu(U) < \infty$ and $\mu(U) - \mu(E) < \epsilon$, so

$$\mu(E) = \inf\{\mu(U) : U \text{ open, } U \supseteq E\}$$

holds in the case $\mu(E) < \infty$; the case $\mu(E) = \infty$ is trivial. Since the $\{E_j\}_{j=1}^\infty$ and $\{F_j\}_{j=1}^\infty$ are disjoint sequences we have

$$\mu\left(\bigcup_{j=1}^k E_j\right) \leq \mu\left(\bigcup_{j=1}^k F_j\right) + \frac{\epsilon}{2},$$

for every $k \in \mathbb{N}$. Since $\mu$ is countably additive, if $\mu(E) = \infty$ then the l. h. side can be made arbitrarily large for large enough $k$, and therefore so can $\mu\left(\bigcup_{j=1}^k F_j\right)$; since these unions are compact, we have

$$\mu(E) = \sup\{\mu(K) : K \text{ compact, } K \subseteq E\}$$

in this case. If $\mu(E) < \infty$, then for sufficiently large $k$ we have

$$\mu(E) < \mu\left(\bigcup_{j=1}^k E_j\right) + \frac{\epsilon}{2}$$
and therefore
\[ \mu(E) - \epsilon < \mu \left( \bigcup_{j=1}^{k} F_j \right) \]
so again
\[ \mu(E) = \sup \{ \mu(K) : K \text{ compact}, K \subseteq E \} . \]
Thus in this case—which is the one most frequently encountered—no Borel measure that assigns finite measure to compact sets can behave in unexpected ways.

**Corollary:** Under the hypotheses of the preceding proposition, given a Borel set \( E \) and an \( \epsilon > 0 \) one can write \( E = \bigcup_j E_j \) where the \( \{E_j\}_{j=1}^{\infty} \) are relatively compact, and find compact sets \( \{F_j\}_{j=1}^{\infty} \) and relatively compact open sets \( \{U_j\}_{j=1}^{\infty} \) with \( F_j \subseteq E_j \subseteq U_j \), such that \( \sum_{j=1}^{\infty} \mu(U_j \setminus F_j) < \epsilon \). In particular, every Borel set is contained in a \( G_{\delta} \)-set \( W \) and contains a \( K_{\sigma} \)-set \( F \) with \( \mu(U \setminus F) = 0 \).

**Proof.** The first assertion was proved in the course of proving the proposition, and the second follows in the same way as, e.g., Royden’s Ch. 3, Prop. 15, p. 63.

Now the “Bourbaki Integral.” Given a locally compact Hausdorff space \( T \) and a positive linear functional \( G : \mathcal{K}(T) \to \mathbb{R} \), there is a way of getting a theory of integration out of the functional by extending the functional rather than by getting a measure from it. This approach is exposed, e.g., in N. Bourbaki, "L’Intégration," Chs. I–IV, Hermann (Paris), 1952. Briefly, one gets the integral by extending \( G \) to a functional \( G^* \) defined on nonnegative \( \mathbb{R}^+ \cup \{+\infty\} \)-valued lower-semicontinuous functions \( g \) by setting
\[ G^*(g) = \sup \{ G(f) : 0 \leq f \leq g, \ f \in \mathcal{K}(T) \} \in \mathbb{R}^+ \cup \{+\infty\} \]
and then extending \( G^* \) to all \( \mathbb{R}^+ \cup \{+\infty\} \)-valued functions by setting
\[ G^*(h) = \inf \{ G^*(g) : 0 \leq h \leq g, \ g \ \text{l. s. c. and} \ \mathbb{R}^+ \cup \{+\infty\} \text{-valued} \} . \]
This is an “upper integral” having properties as a functional similar to those of outer measure as a set function. A suitable class of functions is then selected out as “integrable”, and the integrals of these functions are again given by
\[ G(f) = \inf \{ G^*(g) : 0 \leq f \leq g, \ g \ \text{l. s. c. and} \ \mathbb{R}^+ \cup \{+\infty\} \text{-valued} \} . \]

We should make contact with this approach because it is a very useful one in the study of boundary value problems of partial differential equations—its classical application occurs in the study of regularity of boundary points for the (Dirichlet problem of the) Laplace equation. The basic lemma is

**Proposition:** Let \( T \) be a locally compact Hausdorff space, \( G \) a positive linear functional on \( \mathcal{K}(T) \), and \( \nu \) the representing measure for \( G \) constructed from it as in §5 above. Then for any l. s. c. \( \mathbb{R}^+ \cup \{+\infty\} \)-valued function \( g \)
\[ \int g(t) \, d\nu(t) = \sup \{ G(f) : 0 \leq f \leq g, \ f \in \mathcal{K}(T) \} . \]

**Proof.** Given some l. s. c. \( \mathbb{R}^+ \cup \{+\infty\} \)-valued function \( g \), define
\[ \mathcal{L}(g) = \{ f \in \mathcal{K}(T) : 0 \leq f \leq g \} \]
in analogy with the family \( \mathcal{L}(U) \) for open \( U \). Just as in that case, \( \mathcal{L}(g) \) is directed upward, so \( \sup \{ G(f) : 0 \leq f \leq g, \ f \in \mathcal{K}(T) \} = \lim \{ G(f) : f \in \mathcal{L}(g) \} \), and that’s useful. It helps to get two rather silly cases out of the way first. If \( g \geq 0 \) is a l. s. c. function on \( T \) which equals \( +\infty \) on a set of positive \( \nu \)-measure, then for any \( n \in \mathbb{N} \) the set \( U_n = \{ t \in T : g(t) > n \} \) is open, and if \( f \in \mathcal{L}(U_n) \) then \( n \cdot f \leq g \) so \( G^*(g) \geq n \cdot G(f) \). Since \( G(f) \) can be as close to \( \nu(U_n) \) as we wish and \( \nu(U_n) \geq \nu(\{ t \in T : g(t) = +\infty \}) \), we see that there exist \( f \in \mathcal{L}(g) \) with \( G(f) \) as large as we wish, so \( G^*(g) = +\infty = \int g \, d\nu \) in this case. Similarly, if for some \( \epsilon > 0 \) we have \( \nu(\{ t \in T : g(t) > \epsilon \}) = +\infty \), then for any \( f \in \mathcal{L}(\{ t \in T : g(t) > \epsilon \}) \) we have \( \epsilon \cdot f \in \mathcal{L}(g) \), therefore
\[ \epsilon \cdot G(f) \leq G^*(g), \] and since \( G(f) \) can be as large as we please, again we have \( G^*(g) = +\infty = \int g \, d\nu. \) So the only case that really needs consideration is that in which \( g \) is finite \( \nu \)-a. e. and for which each of the open sets \( \{ t \in T : g(t) > \epsilon \} \) has finite \( \nu \)-measure. And then we can continue to simplify: if for \( n \in \mathbb{N} \) we put

\[ U_n = \left\{ t \in T : g(t) > \frac{1}{n} \right\} \quad \text{and} \quad g_n = (g \wedge n) \cdot \chi_{U_n}, \]

then each \( g_n \) is l. s. c. and the sequence \( \{g_n\}_{n=1}^{\infty} \) increases to \( g \) pointwise, so \( \int g \, d\nu = \lim_{n \to \infty} \int g_n \, d\nu \) by the monotone convergence theorem. Thus, since \( \bigcup_n \mathcal{L}(g_n) = \mathcal{L}(g) \), if we can establish the proposition for each \( g_n \) then we shall have established it in general.

So consider a l. s. c. function \( g : T \to (m, M) \subseteq \mathbb{R}^+ \) (we can assume that the endpoint values are not attained, with no loss of generality), and such that \( \nu(\{ t \in T : g(t) > m \}) < \infty. \) Partition the interval \( m = y_0 < \cdots < y_k = M, \) and let \( E_j = \{ t \in T : y_{j-1} < g(t) \leq y_j \}. \) These are evidently Borel sets of finite measure, and for a suitable choice of the \( \{y_j\}_{j=1}^{k} \) the simple function defined by \( s = \sum_{j=1}^{k} y_j - 1 \chi_{E_j} \leq g \) can have \( \int s \, d\nu \) as close to \( \int g \, d\nu \) as we desire. Given \( \epsilon > 0 \) we can find compact \( K_j \subseteq E_j \) for which the usual \( \nu(E_j \setminus K_j) < \epsilon \) holds. Each \( K_j \) is compact and contained in the open set \( \{ t \in T : g(t) > y_j \}. \) Since a disjoint finite family of compact sets has a disjoint family of corresponding open neighborhoods, we can find a disjoint family of open sets \( U_j \supseteq K_j \) with \( g(t) > y_j \) in each \( U_j, j = 1, \ldots, k. \) For each \( j \) we can find a continuous \( f_j, 0 \leq f_j \leq 1, \) with \( \text{Supp}(f_j) \subseteq U_j \) and \( f_j = 1 \) on \( K_j. \) Then

\[ f = \sum_{j=1}^{k} y_{j-1} f_j \leq g \quad \text{while} \quad G(f) = \sum_{j=1}^{k} y_{j-1} G(f_j) \geq \sum_{j=1}^{k} y_{j-1} \nu(K_j) \geq \int s \, d\nu - kM \cdot \epsilon \]

and since \( \epsilon > 0 \) is in our hands, we have demonstrated that \( G^*(g) = \sup \{ G(f) : 0 \leq f \leq g \text{ and } f \in \mathcal{K}(T) \} \) in the only case that remained.

**Proposition:** Let \( T \) be a locally compact Hausdorff space, \( G \) a positive linear functional on \( \mathcal{K}(T) \), and \( \nu \) the representing measure for \( G \) constructed from it as in §5 above. Then for any \( \nu \)-measurable function \( f : T \to \mathbb{R}^+ \cup \{+\infty\}, \)

\[ \int f(t) \, d\nu(t) = \inf \{ G^*(g) : 0 \leq f \leq g, \; g : T \to \mathbb{R}^+ \cup \{+\infty\} \text{ l. s. c.} \}. \]

**Proof.** This will look rather familiar. The only interesting case occurs when \( f \) is integrable, since \( g = +\infty \) identically will majorize \( f \) (so the infimum is not being taken over an empty class) and any l. s. c. \( g \) that majorizes \( f \) will have \( G^*(g) = \int g \, d\nu = +\infty \) by monotonicity of the integral.

If for each \( n \in \mathbb{N} \) we set

\[ E_n = \left\{ t \in T : f(t) > \frac{1}{n} \right\} \quad \text{and} \quad f_n = (f \wedge n) \cdot \chi_{E_n}, \]

then \( f_1 \leq f_2 \leq \cdots, \) each \( f_n \) is bounded and is nonzero only on \( E_n \) (where its values are at least \( \frac{1}{n} \)), and \( f_n(t) \to f(t) \) at each \( t \in T \) as \( n \to \infty. \) If we set \( h_1 = f_1 \) and \( h_n = f_n - f_{n-1} \) for \( n \geq 2, \) then \( f(t) = \sum_{n=1}^{\infty} h_n(t) \) at each point \( t \in T. \) Each term \( h_n \) is a nonnegative bounded measurable function which is zero everywhere outside a set of finite measure (the same \( E_n \)). By the monotone convergence theorem, we have

\[ \int f \, d\nu = \sum_{n=1}^{\infty} \int h_n \, d\nu. \]

It is now clear that the proposition will be established if we can establish it in the case in which \( f \) is a bounded measurable function which is identically zero outside a set of finite measure. For if we know that
case, then given any $\epsilon > 0$, for each $n \in \mathbb{N}$ we can find some l. s. c. $g_n \geq h_n$ with $G^*(g_n) \leq \int h_n \, d\nu + \frac{\epsilon}{2^n}$. If we put $g = \sum_{n=1}^{\infty} g_n$, then $g \geq f$, $g$ will be l. s. c., and

$$G^*(g) = \int g \, d\nu = \sum_{n=1}^{\infty} \int g_n \, d\nu \leq \sum_{n=1}^{\infty} \int h_n \, d\nu + \epsilon = \int f \, d\nu + \epsilon$$

so we’re done.

The method for handling the special case is the usual. If $f[T] \subseteq [0, M]$ and $0 = y_0 < \cdots < y_k = M$ is a partition of $[0, M]$, then put

$$E_j = \{ t \in T : y_{j-1} < f(t) \leq y_j \} .$$

The simple function $s = \sum_{j=1}^{k} y_j \chi_{E_j}$ majorizes $f$, and for a suitable choice of the partition we can have $\int s \, d\nu \leq \int f \, d\nu + \frac{\epsilon}{2}$. Each $E_j$ has finite $\nu$-measure, so there is an open $U_j \supseteq E_j$ for which the relation $\nu(U_j) < \nu(E_j) + \frac{\epsilon}{2(M+1)k}$ holds. The simple function $g = \sum_{j=1}^{k} y_j \chi_{U_j} \geq s \geq f$ is l. s. c. (as a positive linear combination of characteristic functions of open sets, which are themselves l. s. c.), and

$$G^*(g) = \sum_{j=1}^{k} y_j \nu(U_j) \leq \int s \, d\nu + \frac{\epsilon}{2} .$$

This establishes the special case, which establishes the proposition.

**Corollary:** $\mathcal{K}(T)$ is dense in $L^1(\nu)$: for every integrable $h$ there exists for every $\epsilon > 0$ some $f \in \mathcal{K}(T)$ with $\|h - f\|_1 < \epsilon$.

**Proof.** Write $h = h^+ - h^-$. One can find l. s. c. $g_1 \geq h^+$ for which

$$\int g_1 \, d\nu \leq h^+ \, d\nu - \frac{\epsilon}{4} \implies \|g_1 - h^+\|_1 < \frac{\epsilon}{4}$$

and then find $f_1 \in \mathcal{K}(T)$ for which

$$\int g_1 \, d\nu - \frac{\epsilon}{4} \leq \int f_1 \, d\nu \implies \|f_1 - g_1\|_1 < \frac{\epsilon}{4} \implies \|h^+ - f_1\|_1 < \frac{\epsilon}{2} .$$

One can similarly find $f_2$ with $\|h^- - f_2\|_1 < \frac{\epsilon}{2}$, and then

$$\|(f_1 - f_2) - h\|_1 < \epsilon$$

as desired.

One can show that $\mathcal{K}(T)$ is dense in $L^p(\nu)$, $1 < p < \infty$, in a similar manner; the details are left to the reader.