1 Borel measures

In order to understand the uniqueness theroem we need a better understanding of $h^1(\mathbb{D})$ and its boundary behavior, as well as $H^1(\mathbb{D})$. We recall that the boundary function of an element $U \in h^2(\mathbb{D})$ can be obtained from the Riesz representation theorem for L^2, which states that scalar products are the only continuous linear functionals on L^2.

To analyze $h^1(\mathbb{D})$, we need a description of the linear functionals acting on $C([-1/2, 1/2])$, i.e., the space of continuous functions on $[-1/2, 1/2]$). (Note that this is a closed space, since continuous functions on compact intervals are already uniformly continuous.) Since this is rather crucial for our purpose, we will first do a review of measures of finite total variation on an interval, and then give a proof of the Riesz representation theorem for functionals on continuous functions defined on a compact interval.

We recall the following description of positive Borel measures on $\mathbb{R}$. A positive Borel measure is a function defined on the Borel sigma algebra $\mathcal{B}$ with values in $[0, \infty]$ that satisfies $\mu(\emptyset) = 0$ and is countably additive. For our purpose the description of Borel measures on the real line given in Theorem 1.16 of Folland's Real Analysis is important. It essentially states that the Borel measures are in 1-1 correspondence to the inreasing, right continuous functions on $\mathbb{R}$ in the following sense: If F is such a function, then μ defined on half open intervals by

$$\mu((a, b]) = F(b) - F(a)$$

extends to a Borel measure on $\mathcal{B}$, and in the other direction, if μ is a Borel measure on $\mathbb{R}$, then F defined by

$$F(x) = \begin{cases} \mu((0, x]) \text{ if } x > 0, \\ 0 \text{ if } x = 0 \\ -\mu((x, 0]) \text{ if } x < 0 \end{cases}$$

is a right continuous, increasing function on $\mathbb{R}$. (Consider the example of Lebesgue measure, and the example of the Dirac measure to visualize the connection.) The measure μ_F is also called the Lebesgue-Stieltjes measure of F, and F is called the distribution function of μ_F. Some texts use the notation dF to mean μ_F.

Theorem 1. *Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and define $G(x) = F(x+)$.*

1. *The set of discontinuities of F is countable.*
2. *F and G are differentiable a.e., and $F' = G'$ a.e.*

Proof. Consider F on $[a, b]$. Let S_m be the set of points where the jump of F exceeds $1/m$. Assume that $x_1 < x_2 < ... < x_n$ are in S_m (S_m might have more points.) We have

$$\frac{n}{m} \leq \sum_j (F(x_j+) - F(x_j-)) \leq F(b) - F(a).$$

This gives an upper bound for n, $n \leq m(F(b) - F(a)$. Since $\cup S_m$ is the set of points where right and left limit of F do not agree, we get that there are countably many points with this property in $[a, b]$ and hence in $\mathbb{R}$.

We note that G is increasing and right continuous, and agrees with F except possibly at the discontinuities. Recall that the measure μ_G is defined by

$$\mu_G([x, y)) = G(y) - G(x).$$

The Lebesgue differentiation theorem implies that $\mu_G((x, x + r))/r \to f(x)$ for a.e. x, where $f \in L^1_{loc}$. This implies that the derivative of G exists a.e. (and equals f).

Define $H = G - F$. Consider H on $[-N, N]$. From the first part H has only countably many points x_j where it is non-zero, and since F is increasing,

$$H(x_j) = F(x_j+) - F(x_j) > 0$$

at every such point. Can show as in part one that

$$\sum_{|x_j|<N} H(x_j) \leq F(N+) - F(-N) < \infty.$$

Define

$$\mu = \sum_j H(x_j)\delta_{x_j}.$$

Then μ is finite on compact sets, hence regular by Theorem 1.16 and 1.18. Clearly, μ is singular with respect to m. Since

$$|H(x + h) - H(x)| \leq H(x + h) + H(x) \leq \mu((x - 2|h|, x + 2|h|)),$$

the difference quotient of H goes to zero a.e. by the previous theorem. $\square$

2 Complex measures and total variation

In order to formulate the Riesz representation theorem we require compex valued measures, that is, set functions μ defined on the Borel sigma algebra over $\mathbb{R}$ that satisfy $\mu(\emptyset) = 0$, that are countably additive, but in distiction to positive measures assume values in $\mathbb{C}$. (We will always assume that complex measures are finite.)

The distribution functions of positive measures are increasing. For complex valued measures it turns out that the distribution functions are functions of so called bounded variation. We consider first measures with values in $\mathbb{R}$. If $\mu : \mathcal{B} \to \mathbb{R}$, then the Jordan decomposition theorem (Folland, page 87) implies that μ can be decomposed as

$$\mu = \mu_+ - \mu_-,$$

where μ_+ and μ_- are unique positive Borel measures. The total variation measure $|\mu|$ is defined to be

$$|\mu| = \mu_+ + \mu_-.$$

This is a direct generalization of the corresponding decomposition for L^1-functions. In particular, the total variation measure $|\mu|$ is a generalization of the absolute value function. However, it should be pointed out that in general $|\mu(A)| \leq |\mu|(A)$ for a Borel set A. An important special case is the measure μ defined by

$$\mu(A) = \int_A f(x)dx,$$

where f is a real valued L^1 function. Then $\mu_+(A) = \int_A f_+ dx$ (and similarly for μ_-), and we have

$$|\mu|(A) = \int_A |f|dx.$$

Here the inequality above is clear, since $|\mu(A)| = |\int_A f dx|$.

For complex valued Borel measures one needs to use a polar representation to define the total variation of μ. One can show that there exists g with $|g| = 1$ μ-a.e. and a positive Borel measure, such that

$$\int_A d\mu = \int_A g d\nu.$$

Then ν is called the total variation measure of μ, and written as $\nu = |\mu|$. (It is instructive to write out what g and ν are in the above example $\mu(A) = \int_A f dx$.)

3 The Radon-Nikodym decomposition

Let ν be a signed (complex) measure, and let μ be a positive measure. We say that ν is absolutely continuous with respect to μ ($\nu \ll \mu$), if the implication $\mu(E) = 0$ implies $\nu(E) = 0$ holds. (One can use $|\nu|$ instead of ν in this defintion, see exercise 8 on page 92 of Folland.)

A straightforward way to generate a measure ν that is absolutely continuous with respect to a given measure ν is as follows. Let $f \in L^1(\mu)$ and define

$$\nu(A) = \int_A f(x) d\mu(x).$$

Then $\nu \ll \mu$. (The proof that $\nu \ll \mu$ is evident.)

The Radon-Nikodym decomposition states that if ν is a complex measure (in particular, this means that $|\nu|$ is finite measure) and μ is a positive σ-finite measure, then there exist $f \in L^1(\mu)$ and a complex measure ρ such that

$$d\nu = f d\mu + d\rho$$

and $\mu \perp \rho$. (Recall that this means that the measure space can be written as a disjoint union where μ is the zero measure on one set, while ρ is the zero measure on the other set.)

For Borel measure on the real line (and on n-dimensional space) one can further decompose the measure ρ (Folland page 106). One can write any complex Borel measure μ on $\mathbb{R}^n$ as

$$\mu = \mu_d + \mu_{ac} + \mu_s.$$

Here μ_d is a countable sum $\sum_j c_j \delta_{x_j}$, μ_{ac} is of the form $f dm$ with integrable f (m is Lebesgue measure), and μ_s is a singular measure. On $\mathbb{R}$ a singular measure can be characterized by its distribution function F; this function is continuous, monotonically increasing, and $F' = 0$ almost everywhere. In particular $d\mu_s = dF \neq F' dx$. (The Cantor function is an example of such an F.)

In fact, all three pieces of this decomposition can be characterized in terms of their distribution functions; discrete measures correspond to piecewise constant functions, absolutely continous measures have absolutely continuous distribution functions (see Section 3.5 on page 100 in Folland), and singular measures have continuous distribution functions with $F' = 0$ a.e.

4 Total variation

Since we know how to calculate integrals for monotonic functions, we also know how to calculate integrals for linear combinations of monotonic functions F. It turns out that the class of complex valued distribution functions for which we can define the Riemann Stieltjes integral is equivalent

We need the total variation of a function F. (The connection with the above total variation for measures lies in the fact that if F is the distribution function of μ, then the two notions of total variation turn out to be the same.) Define

$$T_F(x) = \sup\{\sum_{j=1}^{n} |F(x_j) - F(x_{j-1})| : n \in \mathbb{N}, -\infty < x_0 < ... < x_n = x\}.$$

T_F is called the total variation of F. We note that

$$T_F(b) - T_F(a) = \sup\{\sum_{j=1}^{n} |F(x_j) - F(x_{j-1})| : n \in \mathbb{N}, a = x_0 < ... < x_n = b\}$$

This quantity is called the total variation of F on $[a, b]$. We let $BV([a, b])$ be the set of functions that have finite total variation on $[a, b]$. Note that if F is a distribution function of a measure μ, then the sums in the above definition become

$$\sum_{j=1}^{n} |\mu((x_{j-1}, x_k])|,$$

and one can show that the supremum over this sums is just $|\mu|((a, b])$.

1. For a bounded, increasing function on $[a, b]$ we have $T_F(b) - T_F(a) = F(b) - F(a)$.

2. BV is a vector space.

3. If F is differentiable and F' is bounded, then F is in BV by the mean value theorem.

Theorem 2. *F in BV real-valued, then $T_F + F$ and $T_F - F$ are both increasing functions.*

Proof. Let $x < y$ and $\varepsilon > 0$. Choose $x_0 < ... < x_n = x$ such that $\sum |F(x_j) - F(x_{j-1})| \geq T_F(x) - \varepsilon$. We obtain

$$\begin{aligned} T_F(y) - F(y) \geq & \sum_{j=1}^{n} |F(x_j) - F(x_{j-1})| + |F(y) - F(x)| \\ & - (F(y) - F(x)) - F(x) \\ \geq & T_F(x) - \varepsilon - F(x) \end{aligned}$$

which implies the claim. □

Hence, real-valued functions in BV are the difference of two bounded increasing functions (namely $1/2(T_F \pm F)$. Since for increasing functions the one-sided limits always exist, we obtain that for $F \in BV$ the values $F(x+)$ and $F(x-)$ always exist.

By the previous theorem, F has countably many discontinuities, and its derivative exists almost everywhere and equals a.e. the derivative of $x \mapsto F(x+)$.

We define NBV to be the space of functions in BV that are right continuous and have limit zero at $-\infty$. We have $F \in BV$ implies that G defined by

$$G(x) = F(x+) - F(-\infty)$$

is in NBV. It is left as an exercise to show that

Lemma 1. *$F \in BV$ implies $T_F(-\infty) = 0$. If F is right continuous, then T_F is also right continuous.*

Theorem 3. *If μ is a complex Borel measure on $\mathbb{R}$, then F defined by $F(x) = \mu((-\infty, x])$ is in NBV. Conversely, if F is in NBV, then there exists unique complex Borel measure μ_F such that*

$$F(x) = \mu_F((-\infty, x]),$$

and $|\mu_F| = \mu_{T_F}$.

Proof. Decompose the complex measure μ as

$$\mu = (\mu_1^+ - \mu_1^- + i(\mu_2^+ - \mu_2^-)$$

with positive finite Borel measures. Define

$$F_j^{\pm}(x) = \mu_j^{\pm}((-\infty, x]).$$

By definition, each $F_j^\pm$ is increasing, right continuous, has limit zero at $-\infty$, and the limit at infinity is $\mu_j^\pm(\mathbb{R})$ which is finite. Hence $F = F_1^+ - F_1^- + i(F_2^+ - F_2^-)$ is in NBV.

For the converse, decompose complex F in NBV as above with increasing $F_j^\pm$. By Theorem 1.16 of Folland, each $F_j^\pm$ defines a positive finite Borel measure $\mu_j^\pm$, which in the obvious way can be used to define μ_F. Exercise 28 on page 107 in Folland shows that $|\mu_F| = \mu_{T_F}$. $\square$

Proposition 1. *1. $F \in NBV$ implies $F' \in L^1(\mathbb{R})$.*

2. $\mu_F \perp m$ (here m denotes Lebesgue measure) if and only if $F' = 0$ a.e.

3. $\mu_F \ll m$ if and only if $F(x) = \int_{-\infty}^x F'(t)dt$.

Proof. (Theorem numbers refer to theorems in Folland's book.) We write

$$F'(x) = \lim_{r \to 0} \frac{\mu_F(E_r)}{m(E_r)}$$

with $E_r = (x, x+r]$ or $E_r = (x-r, x]$. F' exists a.e. by Theorem 3.23. It is in L^1_{loc} by Theorem 3.22, and for the associated measure μ_F we have $d\mu_F = F'dm + d\lambda$. We note that $\lambda \perp m$ implies $d|\mu_F| = |F'|dm + d|\lambda|$ (see remark in the proof of Theorem 3.22). It follows that

$$\int |F'|dm \leq |\mu_F|(\mathbb{R}) = \mu_{T_F}(\mathbb{R}) = T_F(\infty) < \infty,$$

and hence F' must be integrable. The second statement follows from Theorem 3.22. If $\mu_F \ll m$, then $\lambda = 0$, i.e,

$$\mu_F((-\infty, x]) = \int_{(-\infty,x]} F'(u)du,$$

and the left side equals $F(x)$ by definition of μ_F. $\square$

Intuitively, absolutely continuous functions are those functions for which the fundamental theorem of calculus is valid, since the above statement implies that

$$F(x) = \int_{(0,x]} f(u)du.$$

The Lebesgue differentiation theorem provides the other direction, i.e., $F' = f$ a.e.

Corollary 1. *1. If $f \in L^1(\mathbb{R})$, then $F(x) = \int_{(-\infty,x]} f dm$ is in NBV and is absolutely continuous, and $f = F'$ a.e.*

2. If $F \in NBV$ is absolutely continuous, then $F' \in L^1(\mathbb{R})$ and $F(x) = \int_{(-\infty,x]} F'(x)dx$.

Proof. For simplicity assume f real valued. Since

$$|F(x_j) - F(x_{j-1})| \leq \int_{x_{j-1}}^{x_j} |f(u)|du,$$

F has finite total variation. From the definition of F it follows that F is right continuous. From the previous proposition it follows that F is absolutely continuous. Uniqueness in the Lebesgue-Radon-Nikodym theorem implies that $F' = f$ a.e. Second statement follows analogously. □

Cantor's function is continuous, monotonically increasing, satisfies $f(0) = 0$, $f(1) = 1$ and is a.e. constant. Hence, it is not equal to the integral of its derivative. This function is continuous, but not absolutely continuous.

5 Riemann-Stieltjes integral

We start by defining the Riemann-Stieltjes integral. It is worth pointing out that if f is continuous, then the Riemann-Stieltjes integral with respect to a function of bounded variation and the Lebesgue integral with respect to the corresponding Borel measure are the same.

Let $[a, b]$ be a finite interval. For a partition $P = \{a = x_0 < x_1 < ... < x_n = b$ we denote by $\|P\|$ the maximum distance between two consecutive partition elements. Let F be right continuous and have finite total variation on $[a, b]$. (This means that F is a finite complex linear combination of monotonic functions.)We define

$$\int_{[a,b]} f(x)dF(x) = \lim_{\|P\|\to 0} \sum_{j=1}^{n} f(x_j^*)(F(x_j) - F(x_{j-1}))$$

if this limit exists.

Evidently, if $F(x) = x$, this is just the Riemann integral of $f(x)$. Without proof, if $f(x)$ is continuous, then this integral exists. We note that if F is differentiable, then by the mean value therorem the exists y between x_{j-1} and x_j such that

$$F(x_j) - F(x_{j-1}) = F'(y)(x_j - x_{j-1}).$$

This leads to the question under which circumstances the formula

$$\int_{[a,b]} f(x)dF(x) = \int_{[a,b]} f(x)F'(x)dx$$

is true. The problem is that the value y above depends on the end points, but the limit defining the Riemann integral takes arbitrary points x_j^*. Hence we need to know under which conditions the two limits exist and are the same.

It turns out that continuity of F and even uniform continuity of F are not sufficient for this identity to hold. The correct condition is that F be absolutely continuous. We note that a function F is called *absolutely continuous* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every finite collection of disjoint intervals $(a_1, b_1), ..., (a_N, b_N)$ we have

$$\sum_{j=1}^{N} (b_j - a_j) < \delta \text{ implies } \sum_{j=1}^{N} |F(b_j) - F(a_j)| < \varepsilon.$$

Clearly, absolute continuity implies continuity by taking $N = 1$, but the converse is not true. Next, we require extension to complex valued distribution functions. It is worth mentioning that the decomposition of Borel measures into three measures $\mu = \mu_{ac} + \mu_d + \mu_s$ leads to the same decomposition for integrals; for a distribution function F we get a decomposition

$$\int_{[a,b]} f(x)dF(x) = \int_{[a,b]} f(x)F_1'(x)dx + \sum_k c_k f(x_k) + \int_{[a,b]} f(x)dF_2(x),$$

where F_1 is absolutely continuous, and F_2 is singular. (The middle term has at most countably many terms and corresponds to integrations with respect to Dirac measures.)

A useful formula is the following integration by parts formula. Let f be continuous and F have total variation. Then

$$\int_{[a,b]} f(x)dF(x) + \int_{[a,b]} F(x)df(x) = F(b)f(b) - F(a)f(a).$$

To see that this is true, note that

$$\sum_{j=1}^{n} f(x_j)(F(x_j) - F(x_{j-1}) + \sum_{j=1}^{n} F(x_{j-1})(f(x_j) - f(x_{j-1}))$$
$$= F(b)f(b) - F(a)f(a).$$

and as $n \to \infty$, the first two sums converge to the respective Riemann-Stieltjes integrals.

6 Riesz representation

We consider first the argument using the Riemann-Stieltjes integral notion. The setup is the following. We are given the space $C[a,b]$ of continuous functions on $[a,b]$ (with norm $\|f\|_\infty$), and we are given a functional $T : C[a,b] \to \mathbb{C}$ that is linear and bounded (or equivalently linear and continuous). The statement is that there exists right continuous F of finite total variation so that

$$Tf = \int_{[a,b]} f(x)dF(x)$$

for all $f \in C[a,b]$. First, we note that if this is true, then this gives an equivalence, because every functional that is given by an integration against a BV function is linear (property of integrals), and bounded, since

$$|Tf| \leq \|f\|_\infty \int_{[a,b]} dT_f,$$

where T_F is the total variation function of F. (Use the definition of the Riemann-Stieltjes integral to see this!) Since F is assumed to have bounded variation, we see that T is even Lipschitz continuous on $C[a,b]$.

If T is such a functional, what can we say about F in terms of T? We can obtain a formula relating F and T by plugging in special functions f, namely spline functions. Define

$$f_\xi(x) = \begin{cases} x \text{ if } x < \xi, \\ \xi \text{ if } x \geq \xi \end{cases}$$

Then

$$Tf_\xi = \int_{[a,b]} f_\xi(x)dF(x) = \int_{[a,\xi)} xdF(x) + \xi(F(b) - F(\xi)).$$

An integration by parts gives

$$\int_{[a,\xi)} xdF(x) = \xi F(\xi) - aF(a) - \int_{[a,\xi)} F(x)dx$$

and hence

$$Tf_\xi = \xi F(b) - aF(a) - \int_{[a,\xi)} F(x)dx.$$

Differentiate the whole mess with respect to ξ: We get

$$\frac{\partial}{\partial \xi}[Tf_\xi] = F(b) - F(\xi),$$

i.e.,

$$F(\xi) = F(b) - \frac{\partial}{\partial \xi} Tf_\xi.$$

The Riesz representation theorem states that there are no other functionals, i.e., if T is a bounded, linear functional on $C[a,b]$, then there exists $F \in BV[a,b]$ such that

$$Tf = \int_{[a,b]} f(x) dF(x)$$

for all $f \in C[a,b]$.The proof strategy is clear: Define a function F by

$$F(\xi) = C - \frac{\partial}{\partial \xi} Tf_\xi,$$

with the functions f_ξ defined above, and show that (a) F has finite total variation on $[a,b]$, and (b) show that for a suitable choice of C the formula

$$Tf = \int_{[a,b]} f(x) dF(x)$$

is true for all $f \in C[a,b]$.

Technical problem: under the above assumptions we don't know that Tf_ξ as a function of ξ is differentiable. Hence, we have to work with Tf_ξ directly rather than with its derivative. Thus, for the proof of the Riesz representation theorem we require a criterion that tells us when a given function is the antiderivative of a function of bounded variation. The criterion essentially states that the sum of the difference quotients of Tf_ξ has to be uniformly bounded.

Theorem 4. *Let $A : [a,b] \to \mathbb{C}$. Then there exists a function of bounded variation α with*

$$A(x) = \int_0^x \alpha(u) du + A(0),$$

if and only if there exists $C > 0$ so that for all partitions of $[a,b]$

$$\sum_{k=1}^{n-1} \left| \frac{A(x_{k+1}) - A(x_k)}{x_{k+1} - x_k} - \frac{A(x_k) - A(x_{k-1})}{x_k - x_{k-1}} \right| \leq C.$$

Sketch of proof. If A is such an antiderivative, then

$$A(y) - A(x) = \int_x^y \alpha(u) du.$$

Using this in the two fractions, it follows that the value of this sum cannot exceed the total variation of α on $[a, b]$.

Reverse direction: Assume that the sums in the statement of the theorem are bounded. It follows then that the right difference quotients of A are bounded, and hence

$$\alpha(x) = \limsup_{h \to 0, h > 0} \frac{A(x+h) - A(x)}{h}$$

exist. One can show that α has bounded variation, and that A is an antiderivative of α. $\square$

The Riesz-representation theorem using the Riemann-Stieltjes integral is as follows. For simplicity we consider an interval of the form $[0, b]$, and note that the general case can be done with a translation.

Theorem 5. *Let $T : C[a, b] \to \mathbb{C}$ be a continuous linear functional. Then there exists a right continuous function F of bounded variation such that*

$$Tf = \int_{[a,b]} f(x) dF(x)$$

for all $f \in C[a, b]$.

Sketch of proof. For simplicity we consider $a = 0$ and note that the general case can be obtained with a translation. The proof consists in combining the ideas and theorems above. Define a function A by

$$A(\xi) = -Tf_\xi.$$

One needs to show that this function is the antiderivative of a function α of bounded variation using the previous theorem. Once this is done, the derivative of A with respect to ξ is defined a.e., and its integral is A. Define

$$F(x) = A'(x) - A'(b).$$

Since $F(b) = 0$, $f_\xi(0) = 0$, and $A(0) = Tf_0 = T0 = 0$, we get for f_ξ that

$$\begin{aligned}
\int_{[0,b]} f_\xi dF &= F(b) f_\xi(b) - F(0) f_\xi(0) - \int_{[0,b]} F(x) df_\xi(x) \\
&= -\int_0^\xi A'(x) dx \\
&= -(A(\xi) - A(a)) \\
&= Tf_\xi.
\end{aligned}$$

We note that every 'hat' function is a linear combination of two f_ξ, and hence any piecewise linear continuous function with $f(a) = f(b) = 0$ is a finite linear combination of the form

$$f = \sum_j c_j f_{\xi_j}.$$

The final step in the proof is to show that continuous piecewise linear functions are dense in $C[a,b]$, and to use this to finish the proof. To see this, let $f \in C[a,b]$, and let $\varepsilon > 0$. Let $\delta > 0$ be such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$. Let n be so large that $h = n^{-1}(b-a)$ satisfies

$$h < \delta.$$

Define the points x_j of a partition of $[a,b]$ by $x_j = a + jh$. Let ψ_n be the piecewise linear continuous function obtained by connecting the points $(x_j, f(x_j))$ by line segments. Then for $x \in [x_{j-1}, x_j]$ there exists ξ_x such that $\psi_n(x) = f(\xi_x)$ (Intermediate Value Theorem). It follows that

$$|\psi_n(x) - f(x)| = |f(\xi_x) - f(x)| < \varepsilon.$$

Note that n depends on ε, but not on x. Hence

$$\lim_{n\to\infty} \|f - \psi_n\|_\infty = 0.$$

We have for all n that

$$T\psi_n = \int_{[a,b]} \psi_n(x) dF(x).$$

Now we can finish the proof. Let $f \in C[a,b]$. Let $\varepsilon > 0$. We obtain that there exists n_0 so that for $n \geq n_0$

$$|T(\psi_n - f)| \leq \|T\| \|\psi_n - f\|_\infty < \varepsilon$$

and

$$\left| \int_{[a,b]} (f - \psi_n) dF \right| \leq V(F) \|\psi_n - f\|_\infty < \varepsilon.$$

Combining these, we obtain

$$\left| Tf - \int_{[a,b]} f(x) dF(x) \right| < 2\varepsilon,$$

and since ε was arbitrary, we obtain the claim. □

The Riesz representation theorem when formulated via Borel measures is the following.

Theorem 6. *Let $T : C([-1/2, 1/2]) \to \mathbb{C}$ be a continuous linear functional. Then there exists a (complex valued) measure of finite total variation on $[-1/2, 1/2]$ such that for all $f \in C([-1/2, 1/2])$*

$$T(f) = \int_{[-1/2,1/2]} f(t) d\mu(t),$$

and $\|T\| = |\mu|$. (Here $\|T\| = \sup\{|Tf| : \|f\|_\infty \leq 1\}$.)

Sketch of proof. Prove first that T can be decomposed as $T_+ - T_-$, where $T_\pm$ are positive functionals, that is, functionals with the property that $T_+ f \geq 0$ if f is a nonnegative function.

Once this is done, the problem is reduced to showing that the Riesz representation theorem holds for positive functionals. We would like to define the measure μ by

$$\mu(A) = T\chi_A$$

where A is a Borel measure. This is of course not possible, since χ_A is not continuous, so $T\chi_A$ is not defined. We note that characteristic functions of intervals can be approximated from below by nonnegative continuous functions (make a plot!). Hence, we define a set function ρ on intervals by

$$\rho((a, b]) = \sup\{Tf : f \geq 0, f \leq \chi_{(a,b]}\}.$$

Since ρ is defined on half open intervals, it extends to a measure on $[a, b]$. The task is then to show that this measure satisfies $Tf = \int f d\mu$. □

The Riesz representation theorem for continuous functions of compact support is valid under much more general assumptions. A version is proved in Folland, Real Analysis, Theorem 7.17 and Corollary 7.18 on page 223. The original paper (written in French) by F. Riesz that deals with continuous functions on intervals, is about four pages long and completely elementary. A translation can be found at http://nonagon.org/ExLibris/rieszs-equations-integrales .