# Another Riesz Representation Theorem

In these notes we prove (one version of) a theorem known as the Riesz Representation Theorem. Some people also call it the Riesz–Markov Theorem. It expresses positive linear functionals on $C(X)$ as integrals over $X$. For simplicity, we will here only consider the case that $X$ is a compact metric space. We denote the metric $d(x,y)$. For more general versions of the theorem see

- Gerald B. Folland, *Real Analysis – Modern Techniques and Their Applications*, Wiley, theorems 7.2 and 7.17.
- Michael Reed and Barry Simon, *Functional Analysis (Methods of modern mathematical physics, volume 1)*, Academic Press, theorems IV.14 and IV.18.
- H. L. Royden, Real Analysis, Macmillan, chapter 13, sections 4 and 5.
- Walter Rudin, Real and Complex Analysis, McGraw–Hill, theorem 2.14.

The background definitions that we need are

**Definition 1** $C(X)$ is the Banach space of continuous functions on $X$ with the norm $\|\varphi\|_{C(X)} = \sup_{x\in X} |\varphi(x)|$.

**Definition 2** A map $\ell : C(X) \to \mathbb{C}$ is a *positive linear functional* if
(a) $\ell(\alpha\varphi + \beta\psi) = \alpha\,\ell(\varphi) + \beta\,\ell(\psi)$ for all $\alpha, \beta \in \mathbb{C}$ and all $\varphi, \psi \in C(X)$ and
(b) $\ell(\varphi) \ge 0$ for all $\varphi \in C(X)$ that obey $\varphi(x) \ge 0$ for all $x \in X$.

**Problem 1** Let $\ell : C(X) \to \mathbb{C}$ be a positive linear functional. Prove that

$$\big|\ell(\varphi)\big| \le \ell(1)\,\|\varphi\|_{C(X)}$$

for all $\varphi \in C(X)$. Here 1 is of course the function on $X$ that always takes the value 1.

**Definition 3**

(a) The set, $\mathcal{B}_X$, of *Borel subsets* of $X$ is the smallest $\sigma$–algebra that contains all open subsets of $X$.

(b) A *Borel measure* on $X$ is a measure $\mu : \mathcal{B}_X \to [0,\infty]$.

(c) A Borel measure $\mu$ on $X$ is said to be *regular* if, for all $A \in \mathcal{B}_X$,
  (i) $\mu(A) = \inf\big\{\ \mu(\mathcal{O})\ \big|\ A \subset \mathcal{O},\ \mathcal{O} \text{ open}\ \big\}$
  (ii) $\mu(A) = \sup\big\{\ \mu(C)\ \big|\ C \subset A,\ C \text{ compact}\ \big\}$

We shall prove

**Theorem 4 (Riesz Representation)** *Let $X$ be a compact metric space. If $\ell : C(X) \to \mathbb{C}$ is a positive linear functional on $C(X)$, then there exists a unique regular Borel measure $\mu$ on $X$ such that*

$$\ell(f) = \int f(x)\ d\mu(x)$$

*The measure $\mu$ is finite.*

**Motivation:** By way of motivation for the proof, let's guess what the measure is. To do so, we pretend that we already have a measure $\mu$ with $\ell(f) = \int f(x)\ d\mu(x)$ and derive a formula for $\mu$ in terms of $\ell$. We start off by considering any open $\mathcal{O} \subset X$. Of course $\mu(\mathcal{O}) = \int_X \chi_{\mathcal{O}}(x)\ d\mu(x)$. As the characteristic function $\chi_{\mathcal{O}}$ is not continuous, we cannot express $\mu(\mathcal{O}) = \ell(\chi_{\mathcal{O}})$. But we can express $\chi_{\mathcal{O}}$ as a limit of continuous functions. For each $n \in \mathbb{N}$, set

$$f_n(x) = \begin{cases} 0 & \text{if } x \in X \setminus \mathcal{O} \\ n\, d\big(x, X \setminus \mathcal{O}\big) & \text{if } x \in \mathcal{O} \text{ and } d(x, X \setminus \mathcal{O}) \le \frac{1}{n} \\ 1 & \text{if } x \in \mathcal{O} \text{ and } d(x, X \setminus \mathcal{O}) \ge \frac{1}{n} \end{cases}$$

This is a sequence of continuous functions on $X$ with

- $0 \le f_n(x) \le 1$ for all $n \in \mathbb{N}$ and $x \in X$ and
- $f_n \restriction X \setminus \mathcal{O} = 0$ for all $n \in \mathbb{N}$. Here "$\restriction$" means "restricted to".
- Because $\mathcal{O}$ is open, $X \setminus \mathcal{O}$ is compact, so that, for each $x \in \mathcal{O}$,
  - $d(x, X \setminus \mathcal{O}) > 0$ and
  - $f_n(x) = 1$ for all $n \ge \frac{1}{d(x, X \setminus \mathcal{O})}$.

  Consequently, $\lim_{n\to\infty} f_n(x) = \chi_{\mathcal{O}}(x)$ for all $x \in X$.

So by the dominated convergence theorem (or, if you prefer, the monotone convergence theorem)

$$\mu(\mathcal{O}) = \lim_{n\to\infty} \ell(f_n)$$

Of course, this only determines $\mu$ on open sets. But if $\mu$ is regular, it is completely determined by its values on opens sets. We are now ready to start the proof itself. See the notes "Review of Measure Theory" for a collection of measure theory definitions and theorems.

**Start of the proof of Theorem 4:** Define, for any open set $\mathcal{O} \subset X$,

$$\mu^*(\mathcal{O}) = \sup \big\{\ \ell(f)\ \big|\ f \in C(X),\ f \restriction X \setminus \mathcal{O} = 0,\ 0 \le f(x) \le 1 \text{ for all } x \in X\ \big\}$$

and, for any $A \subset X$,

$$\mu^*(A) = \inf \big\{\ \mu^*(\mathcal{O})\ \big|\ \mathcal{O} \subset X,\ \mathcal{O} \text{ open },\ A \subset \mathcal{O}\ \big\}$$

**Lemma 5**

*(a)* $\mu^*$ *is a well–defined outer measure on* $X$ *with* $\mu^*(A) \le \ell(1)$ *for all* $A \subset X$.

*(b)* $\mu^*(A) = \inf\big\{\ \mu^*(\mathcal{O})\ \big|\ \mathcal{O} \subset X,\ \mathcal{O}\ \text{open}\ ,\ A \subset \mathcal{O}\ \big\}$ *for all* $A \subset X$.

*(c) If* $\mathcal{O} \subset \mathcal{X}$ *is open, then* $\mathcal{O}$ *is measurable.*

*(d) If* $\mathcal{U} \subset \mathcal{X}$ *is Borel, then* $\mathcal{U}$ *is measurable.*

*(e) If* $A \subset \mathcal{X}$ *is measurable, then* $\mu^*(A) = \sup\big\{\ \mu^*(C)\ \big|\ C \subset A,\ C\ \text{compact}\ \big\}$.

**Proof:** (a) This follows almost directly from the definitions and the fact that, since $\ell$ is a positive linear functional, $0 \le \ell(f) \le \ell(1)$ for any function $f \in C(X)$ that obeys $0 \le f(x) \le 1$ for all $x \in X$.

(b) This is of course part of the definition of $\mu^*$.

(c) Recall that, by definition, $\mathcal{O} \subset X$ is measurable with respect to $\mu^*$ if we have $\mu^*(A) = \mu^*(A \cap \mathcal{O}) + \mu^*\big(A \cap (X \setminus \mathcal{O})\big)$ for all $A \subset X$. So let $A \subset \mathcal{X}$. That $\mu^*(A) \le \mu^*(A \cap \mathcal{O}) + \mu^*\big(A \cap (X \setminus \mathcal{O})\big)$ is part of the definition of "outer measure". So it suffices to prove that, for any $\varepsilon > 0$,

$$\mu^*(A) \ge \mu^*(A \cap \mathcal{O}) + \mu^*\big(A \cap (X \setminus \mathcal{O})\big) - \varepsilon$$

So fix any $\varepsilon > 0$. The definition of $\mu^*$ is more direct when applied to open sets than to general sets, so we start by using the following argument to replace the general set $A$ with an open set $\tilde{A}$. By the definition of $\mu^*(A)$, there is an open set $\tilde{A} \subset X$ such that $A \subset \tilde{A}$ and $\mu^*(A) \ge \mu^*(\tilde{A}) - \frac{\varepsilon}{2}$. As $A \subset \tilde{A}$, we have that $\mu^*(\tilde{A} \cap \mathcal{O}) \ge \mu^*(A \cap \mathcal{O})$ and $\mu^*\big(\tilde{A} \cap (X \setminus \mathcal{O})\big) \ge \mu^*\big(A \cap (X \setminus \mathcal{O})\big)$. So it suffices to prove that

$$\mu^*(\tilde{A}) \ge \mu^*(\tilde{A} \cap \mathcal{O}) + \mu^*\big(\tilde{A} \cap (X \setminus \mathcal{O})\big) - \tfrac{\varepsilon}{2}$$

Here is the idea of the rest of the proof. We are going to construct three continuous functions $f_1, f_2, f_3 : X \to [0, 1]$ and an open set $\widetilde{\mathcal{O}^c}$ that contains, but is only a tiny bit bigger than $X \setminus \mathcal{O}$ (remember that $\tilde{A} \cap (X \setminus \mathcal{O})$ is not open), such that

$$\begin{aligned} \mu^*(\tilde{A} \cap \mathcal{O}) &\le \ell(f_1) + \tfrac{\varepsilon}{4} \quad f_1 \text{ nonzero only on } \tilde{A} \cap \mathcal{O} \\ \mu^*\big(\tilde{A} \cap (X \setminus \mathcal{O})\big) &\le \ell(f_2) + \tfrac{\varepsilon}{4} \quad f_2 \text{ nonzero only on } \tilde{A} \cap \widetilde{\mathcal{O}^c} \\ f_3 &= f_1 + f_2 \end{aligned}$$

Once we have succeeded in doing so, we have finished, since then $f_3$ is nonzero only on $\tilde{A}$, takes values in $[0,1]$ and is continuous so that

$$\mu^*(\tilde{A}\cap\mathcal{O})+\mu^*\big(\tilde{A}\cap(X\setminus\mathcal{O})\big)\le\ell(f_1+f_2)+\tfrac{\varepsilon}{2}=\ell(f_3)+\tfrac{\varepsilon}{2}\le\mu^*(\tilde{A})+\tfrac{\varepsilon}{2}$$

So we now only have to construct $f_1$, $f_2$, $f_3$ and $\widetilde{\mathcal{O}^c}$. The principal hazard that we must avoid arises from fact that $\mathcal{O}$ and $\widetilde{\mathcal{O}^c}$ overlap a bit. So there is a danger that $f_1+f_2$ is larger than 1 somewhere on $\mathcal{O}\cap\widetilde{\mathcal{O}^c}$. Fortunately, $f_1$ is zero on $X\setminus\mathcal{O}$ and all of $\mathcal{O}\cap\widetilde{\mathcal{O}^c}$ is very close to $X\setminus\mathcal{O}$, so $f_1$ is very small on $\mathcal{O}\cap\widetilde{\mathcal{O}^c}$. Here are the details.

Since $\tilde{A}\cap\mathcal{O}$ is open, the definition of $\mu^*(\tilde{A}\cap\mathcal{O})$ implies that there is a continuous function $F_1:X\to[0,1]$ that is nonzero only on $\tilde{A}\cap\mathcal{O}$ and obeys $\mu^*(\tilde{A}\cap\mathcal{O})\le\ell(F_1)+\frac{\varepsilon}{5}$. Since $\ell(F_1)\le\ell(1)<\infty$, we can pick a $\delta>0$ such that $\frac{\delta}{1+\delta}\ell(1)\le\frac{\varepsilon}{20}$. Set $f_1=\frac{F_1}{1+\delta}$. Then

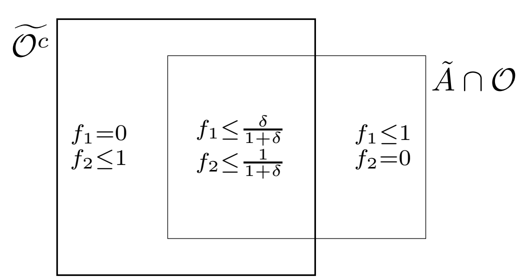
$$\mu^*(\tilde{A}\cap\mathcal{O})\le\ell(F_1)+\tfrac{\varepsilon}{5}=\tfrac{1}{1+\delta}\ell(F_1)+\tfrac{\delta}{1+\delta}\ell(F_1)+\tfrac{\varepsilon}{5}\le\ell(f_1)+\tfrac{\varepsilon}{4}$$

Since $F_1$ is continuous and vanishes on $X\setminus(\tilde{A}\cap\mathcal{O})$, there is an open neighbourhood $\widetilde{\mathcal{O}^c}$ of $X\setminus(\tilde{A}\cap\mathcal{O})\supset X\setminus\mathcal{O}$ such that $F_1(x)\le\delta$ for $x\in\widetilde{\mathcal{O}^c}$. Since $\tilde{A}\cap\widetilde{\mathcal{O}^c}$ is open, the definition of $\mu^*(\tilde{A}\cap\widetilde{\mathcal{O}^c})$ implies that there is a continuous function $F_2:X\to[0,1]$ that is nonzero only on $\tilde{A}\cap\widetilde{\mathcal{O}^c}$ and obeys

$$\mu^*\big(\tilde{A}\cap(X\setminus\mathcal{O})\big)\le\mu^*(\tilde{A}\cap\widetilde{\mathcal{O}^c})\le\ell(F_2)+\tfrac{\varepsilon}{5}=\tfrac{1}{1+\delta}\ell(F_2)+\tfrac{\delta}{1+\delta}\ell(F_2)+\tfrac{\varepsilon}{5}\le\tfrac{1}{1+\delta}\ell(F_2)+\tfrac{\varepsilon}{4}$$

Set $f_2=\frac{F_2}{1+\delta}$ and $f_3=f_1+f_2$. It remains only to verify that $0\le f_3(x)\le 1$, This follows from (see the figure below) the facts that

- $f_1$ is nonzero at most on $\tilde{A}\cap\mathcal{O}$ and $f_2$ is nonzero at most on $\tilde{A}\cap\widetilde{\mathcal{O}^c}$
- on $(\tilde{A}\cap\mathcal{O})\cap\widetilde{\mathcal{O}^c}$, we have $F_1\le\delta$ and $F_2\le 1$ so that $f_3=\frac{F_1}{1+\delta}+\frac{F_2}{1+\delta}\le\frac{\delta}{1+\delta}+\frac{1}{1+\delta}=1$.
- on $(\tilde{A}\cap\mathcal{O})\setminus\widetilde{\mathcal{O}^c}$, we have $f_1\le 1$ and $f_2=0$
- on $\widetilde{\mathcal{O}^c}\setminus(\tilde{A}\cap\mathcal{O})$, we have $f_1=0$ and $f_2\le 1$




(d) By Carathéodory's theorem, the set of all measurable sets is always a $\sigma$–algebra. In our case it contains all open sets and hence must contain all Borel sets.

(e) Let $A \subset X$ be measurable. Then

$$\begin{aligned}\mu^*(A) &= \mu^*(X) - \mu^*(X \setminus A) \\ &= \mu^*(X) - \inf\big\{\ \mu^*(\mathcal{O})\ \big|\ \mathcal{O} \subset X,\ \mathcal{O} \text{ open },\ X \setminus A \subset \mathcal{O}\ \big\} \\ &= \mu^*(X) - \inf\big\{\ \mu^*(X) - \mu^*(X \setminus \mathcal{O})\ \big|\ X \setminus \mathcal{O} \text{ compact },\ X \setminus \mathcal{O} \subset A\ \big\} \\ &= \sup\big\{\ \mu^*(C)\ \big|\ C \subset A,\ C \text{ compact}\ \big\}\end{aligned}$$

■

**Completion of the proof of Theorem 4:** Define $\mu$ to be the restriction of $\mu^*$ to the Borel sets. By Carathéodory's theorem, $\mu$ is a measure. By parts (b) and (d) of Lemma 5, it is a regular Borel measure. Since $\mu(X) = \mu^*(X) = \ell(1)$, it is a finite measure. That $\ell(f) = \int f(x)\ d\mu(x)$ is proven in Lemma 6, below.

That just leaves the uniqueness. If $\nu$ is a regular Borel measure and $\ell(f) = \int f(x)\ d\nu(x)$ for all $f \in C(X)$, then we must have

$$\nu(\mathcal{O}) = \sup\big\{\ \ell(f)\ \big|\ f \in C(X),\ f \restriction X \setminus \mathcal{O} = 0,\ 0 \le f(x) \le 1 \text{ for all } x \in X\ \big\} = \mu^*(\mathcal{O}) = \mu(\mathcal{O})$$

for all open sets $\mathcal{O}$. This was proven in the motivation leading up to the definition of $\mu^*$. The regularity of $\nu$ the forces $\nu(A) = \mu(A)$ for all Borel sets $A$. ■

**Lemma 6** *If $f \in C(X)$, then*

$$\ell(f) = \int f(x)\ d\mu(x)$$

**Proof:** We first observe that it suffices to prove that $\ell(f) \le \int f(x)\ d\mu(x)$ for all real–valued $f \in C(X)$. (Then $\ell(-f) \le \int (-f)(x)\ d\mu$ too, so that $\ell(f) = \int f(x)\ d\mu(x)$ for all real–valued $f \in C(X)$.) We then observe that, since $\mu(X) = \ell(1) < \infty$, it suffices to consider $f \ge 0$. (Otherwise replace $f$ by $f + \|f\|_\infty$.)

So fix any nonnegative $f \in C(X)$ and any $n \in \mathbb{N}$. Define, for each $m \in \mathbb{N}$,

$$\mathcal{B}_m = f^{-1}\Big(\big[\tfrac{m-1}{n}, \tfrac{m}{n}\big)\Big)$$

This $\mathcal{B}_m$ is the intersection of $f^{-1}\big(\big[\frac{m-1}{n}, \infty\big)\big)$, which is closed, and $f^{-1}\big(\big(-\infty, \frac{m}{n}\big)\big)$, which is open. So $\mathcal{B}_m$ is Borel. Since $f$ is bounded, there is an $M \in \mathbb{N}$ such that $\mathcal{B}_m = \emptyset$ for all $m > M$. For each $1 \le m \le M$, there is an open set $\mathcal{O}_m \subset X$ such that $\mathcal{B}_m \subset \mathcal{O}_m$,

$\mu(\mathcal{B}_m) \geq \mu(\mathcal{O}_m) - \frac{1}{nM}$ and $0 \leq f \restriction \mathcal{O}_m \leq \frac{m+1}{n}$, since $\mu$ is regular and $f$ is continuous. Again, for each $1 \leq m \leq M$, define

$$h_m(y) = \frac{d(y\,,\, X \setminus \mathcal{O}_m)}{\sum_{m'=1}^{M} d(y\,,\, X \setminus \mathcal{O}_{m'})}$$

and observe that

$$h_m \in C(X) \qquad 0 \leq h_m \leq 1 \qquad h_m(y) \neq 0 \iff y \in \mathcal{O}_m \qquad \sum_{m=1}^{M} h_m(y) = 1$$

In particular, the denominator $\sum_{m'=1}^{M} d(y\,,\, X \setminus \mathcal{O}_{m'})$ never vanishes because each $y \in X$ is in $\mathcal{O}_{m'}$ for some $1 \leq m' \leq M$. (So $\big\{h_m\big\}_{1\leq m\leq M}$ is a partition of unity. The only reason that it isn't subordinate to the open cover $\big\{\mathcal{O}_m\big\}_{1\leq m\leq M}$ of $X$ is that the support of $h_m$ is $\overline{\mathcal{O}_m}$.) Hence

$$\begin{aligned}
\ell(f) &= \sum_{m=1}^{M} \ell\big(h_m f\big) \leq \sum_{m=1}^{M} \tfrac{m+1}{n}\ell\big(h_m\big) \leq \sum_{m=1}^{M} \tfrac{m+1}{n}\mu(\mathcal{O}_m) \leq \sum_{m=1}^{M} \tfrac{m+1}{n}\big[\mu(\mathcal{B}_m) + \tfrac{1}{nM}\big] \\
&= \sum_{m=1}^{M} \tfrac{m-1}{n}\mu(\mathcal{B}_m) + \tfrac{2}{n}\sum_{m=1}^{M} \mu(\mathcal{B}_m) + \sum_{m=1}^{M} \tfrac{m+1}{n^2 M} \\
&\leq \int f(x)\ d\mu(x) + \tfrac{2}{n}\mu(X) + \max_{1\leq m\leq M} \tfrac{m+1}{n^2} \\
&\leq \int f(x)\ d\mu(x) + \tfrac{2}{n}\mu(X) + \tfrac{1}{n}\big(\|f\|_\infty + \tfrac{1}{n}\big)
\end{aligned} \tag{1}$$

For the first inequality we used the assumption that $\ell$ is positive. As (1) is true for all $n \in \mathbb{N}$, we have that $\ell(f) \leq \int f(x)\ \mu(x)$. ■