CHAPTER **2**

Measures

In this chapter I will introduce the notion of a measure, give a procedure for constructing one, and apply that procedure to construct Lebesgue's measure on $\mathbb{R}^N$ as well as the Bernoulli measures for coin tossing.

§ 2.1 Some Generalities

In this section I give a mathematically precise definition of what a measure is and prove a few elementary properties that follow from the definition. However, to avoid getting lost in the formalities, it will be important to keep the ultimate goal in mind, and for this reason I will begin with a brief summary of what that goal is.

§ **2.1.1. The Idea:** The essence of any theory of integration is a *divide and conquer* strategy. That is, given a space E and a family $\mathcal{B}$ of subsets $\Gamma \subseteq E$ for which one has a *reasonable notion* of *measure* assignment $\Gamma \in \mathcal{B} \longmapsto \mu(\Gamma) \in [0, \infty]$, the *integral* of a function $f : E \longrightarrow \mathbb{R}$ with respect to μ is computed by a prescription that contains the following ingredients. First, one has to choose a partition $\mathcal{P}$ of the space E into subsets $\Gamma \in \mathcal{B}$. Second, having chosen $\mathcal{P}$, one has to select for each $\Gamma \in \mathcal{P}$ a *typical* value a_Γ of f on Γ. Third, given both the partition $\mathcal{P}$ and the selection

$$\Gamma \in \mathcal{P} \longmapsto a_\Gamma \in f(\Gamma) \equiv \operatorname{Range}\big(f \restriction \Gamma\big),$$

one forms the sum

$$\sum_{\Gamma \in \mathcal{P}} a_\Gamma \, \mu(\Gamma). \tag{2.1.1}$$

Finally, using a limit procedure if necessary, one removes the ambiguity (inherent in the notion of *typical*) by choosing the partitions $\mathcal{P}$ in such a way that the restriction of f to each Γ is increasingly close to a constant.

Obviously, even if one ignores all questions of convergence, the only way in which one can make sense out of (2.1.1) is to restrict oneself to partitions $\mathcal{P}$ that are either finite or, at worst, countable. Hence, in general, the final limit procedure will be essential. Be that as it may, when E is itself countable and $\{x\} \in \mathcal{B}$ for every $x \in E$, there is an *obvious* way to avoid the limit step; namely, one chooses $\mathcal{P} = \big\{\{x\} : x \in E\big\}$ and takes

$$\sum_{x \in E} f(x)\mu(\{x\}) \tag{2.1.2}$$

D.W. Stroock, *Essentials of Integration Theory for Analysis*, Graduate Texts in Mathematics 262,
DOI 10.1007/978-1-4614-1135-2_2,

to be the *integral*. (I continue, for the present, to systematically ignore all problems arising from questions of convergence.) Clearly, this is the idea on which Riemann based his theory of integration. On the other hand, Riemann's is not the only *obvious* way to proceed, even in the case of countable spaces E. For example, again assume that E is countable, and take $\mathcal{B}$ to be the set of all subsets of E. Given $f : E \longrightarrow \mathbb{R}$, set $\Gamma(a) = \{x \in E : f(x) = a\} \in \mathcal{B}$ for every $a \in \mathbb{R}$. Then Lebesgue would say that

$$\sum_{a\in \mathrm{Range}(f)} a\,\mu\bigl(\Gamma(a)\bigr) \tag{2.1.3}$$

is an equally *obvious* candidate for the *integral* of f.

In order to reconcile these two *obvious* definitions, one has to examine the assignment $\Gamma \in \mathcal{B} \longmapsto \mu(\Gamma) \in [0,\infty]$ of *measure*. Indeed, even if E is countable and $\mathcal{B}$ contains every subset of E, (2.1.2) and (2.1.3) give the same answer only if one knows that, for any countable collection $\bigl\{\Gamma_n\bigr\} \subseteq \mathcal{B}$,

$$\mu\left(\bigcup_n \Gamma_n\right) = \sum_n \mu\bigl(\Gamma_n\bigr) \quad \text{when } \Gamma_m \cap \Gamma_n = \emptyset \text{ for } m \neq n. \tag{2.1.4}$$

The property in (2.1.4) is called **countable additivity**, and, as will become increasingly apparent, it is crucial. When E is countable, (2.1.4) is equivalent to taking

$$\mu(\Gamma) = \sum_{x\in\Gamma} \mu\bigl(\{x\}\bigr), \quad \Gamma \subseteq E.$$

However, when E is uncountable, the property in (2.1.4) becomes highly non-trivial. In fact, it is unquestionably Lebesgue's most significant achievement to have shown that there are non-trivial *assignments of measure* that enjoy this property.

Having compared Lebesgue's ideas to Riemann's in the countable setting, I close this introduction to Lebesgue's theory with a few words about the same comparison for uncountable spaces. For this purpose, suppose that $E = [0,1]$ and, without worrying about exactly which subsets of E are included in $\mathcal{B}$, assume that $\Gamma \in \mathcal{B} \longmapsto \mu(\Gamma) \in [0,1]$ is a mapping that satisfies (2.1.4).

Now let $f : [0,1] \longrightarrow \mathbb{R}$ be given. In order to integrate f, Riemann says that one should divide up $[0,1]$ into small intervals, choose a representative value of f from each interval, form the associated Riemann sum, and then take the limit as the mesh size of the division tends to 0. As we know, his procedure works beautifully as long as the function f respects the topology of the real line: that is, as long as f is sufficiently continuous. However, Riemann's procedure is doomed to failure when f does not respect the topology of $\mathbb{R}$. The problem is, of course, that Riemann's partitioning procedure is tied to the topology of the reals and is therefore too rigid to accommodate functions that pay little or no attention to that topology. To get around this problem, Lebesgue tailors his partitioning procedure to the particular function f under consideration. Thus, for a given function f, Lebesgue might consider the sequence of partitions $\mathcal{P}_n$, $n \in \mathbb{N}$, consisting of the sets

$$\Gamma_{n,k} = \left\{x \in E : f(x) \in \left[k2^{-n}, (k+1)2^{-n}\right)\right\}, \quad k \in \mathbb{Z}.$$

Obviously, all values of f restricted to any one of the $\Gamma_{n,k}$'s can differ from one another by at most $\frac{1}{2^n}$. Hence, assuming that $\Gamma_{n,k} \in \mathcal{B}$ for every $n \in \mathbb{N}$ and $k \in \mathbb{Z}$ and ignoring convergence problems,

$$\lim_{n\to\infty} \sum_{k\in\mathbb{Z}} \frac{k}{2^n}\, \mu\big(\Gamma_{n,k}\big)$$

simply must be the *integral* of f!

When one hears Lebesgue's ideas for the first time, one may well wonder what there is left to be done. On the other hand, after a little reflection, some doubts begin to emerge. For example, what is so sacrosanct about the partitioning suggested in the preceding paragraph and, for instance, why should one not have done the same thing relative to powers of 3 rather than 2? The answer is, of course, that there is nothing to recommend 2 over 3 and that it should make no difference which of them is used. Thus, one has to check that it really does not matter, and, once again, the verification entails repeated application of countable additivity. In fact, it will become increasingly evident that Lebesgue's entire program rests on countable additivity.

§ **2.1.2. Measures and Measure Spaces:** With the preceding discussion in mind, the following should seem quite natural.

Given a non-empty set E, the **power set** $\mathcal{P}(E)$ is the collection of all subsets of E, and a **σ-algebra** $\mathcal{B}$ is any subset of $\mathcal{P}(E)$ with the properties that $E \in \mathcal{B}$, $\mathcal{B}$ is **closed under countable unions** (i.e., $\{B_n : n \geq 1\} \subseteq \mathcal{B} \implies \bigcup_{n=1}^{\infty} B_n \in \mathcal{B}$), and $\mathcal{B}$ is **closed under complementation** (i.e., $B \in \mathcal{B} \implies B^{\complement} = E \setminus B \in \mathcal{B}$). Observe that if $\{B_n : n \geq 1\} \subseteq \mathcal{B}$, then

$$\bigcap_{n=1}^{\infty} B_n = \left(\bigcup_{n=1}^{\infty} B_n^{\complement}\right)^{\complement} \in \mathcal{B},$$

and so $\mathcal{B}$ is also **closed under countable intersections**. Given E and a σ-algebra $\mathcal{B}$ of its subsets, the pair $(E, \mathcal{B})$ is called a **measurable space**. Finally, if $(E, \mathcal{B})$ and $(E', \mathcal{B}')$ are measurable spaces, then a map $\Phi : E \longrightarrow E'$ is said to be **measurable** if (cf. Exercise 2.1.19 below) $\Phi^{-1}(B') \in \mathcal{B}$ for every $B' \in \mathcal{B}'$. Notice the analogy between the definitions of measurability and continuity. In particular, it is clear that if Φ is a measurable map on $(E_1, \mathcal{B}_1)$ into $(E_2, \mathcal{B}_2)$ and Ψ is a measurable map on $(E_2, \mathcal{B}_2)$ into $(E_3, \mathcal{B}_3)$, then $\Psi \circ \Phi$ is a measurable map on $(E_1, \mathcal{B}_1)$ into $(E_3, \mathcal{B}_3)$.

Obviously both $\{\emptyset, E\}$ and $\mathcal{P}(E)$ are σ-algebras over E. In fact, they are, respectively, the smallest and largest σ-algebras over E. More generally, given any[1] $\mathcal{C} \subseteq \mathcal{P}(E)$, there is a smallest σ-algebra over E, denoted by $\sigma(\mathcal{C})$ and known as the σ-algebra **generated** by $\mathcal{C}$. To construct $\sigma(\mathcal{C})$, note that there is at least one, namely $\mathcal{P}(E)$, σ-algebra containing $\mathcal{C}$, and check that the intersection of all the σ-algebras containing $\mathcal{C}$ is again a σ-algebra that contains $\mathcal{C}$. When E is a topological space, the σ-algebra generated by its open subsets is called the **Borel σ-algebra** and is denoted by $\mathcal{B}_E$.

Given a σ-algebra $\mathcal{B}$ over E, the reason why the pair $(E, \mathcal{B})$ is called a measurable space is that it is the natural structure on which measures are defined. Namely, a **measure** on $(E, \mathcal{B})$ is a map $\mu : \mathcal{B} \longrightarrow [0, \infty]$ that assigns 0 to $\emptyset$ and is countably additive in the sense that (2.1.4) holds whenever $\{\Gamma_n\}$ is a sequence of mutually disjoint elements of $\mathcal{B}$. If μ is a measure on $(E, \mathcal{B})$, then the triple $(E, \mathcal{B}, \mu)$ is called a **measure space**. A measure μ on $(E, \mathcal{B})$ is said to be **finite** if $\mu(E) < \infty$, and it is said to be a **probability measure** if $\mu(E) = 1$, in which case $(E, \mathcal{B}, \mu)$ is called a **probability space**.

Note that if $B, C \in \mathcal{B}$, then $B \cup C = B \cup \big(C \setminus (B \cap C)\big)$ and therefore

$$\mu(B) \le \mu(B) + \mu(C \setminus B) = \mu(C) \quad \text{for all } B, C \in \mathcal{B} \text{ with } B \subseteq C. \tag{2.1.5}$$

In addition, because $C = (B \cap C) \cup \big(C \setminus (B \cap C)\big)$ and $B \cup C = B \cup \big(C \setminus (B \cap C)\big)$,

$$\mu(B) + \mu(C) = \mu(B) + \mu\big(B \cap C) + \mu\big(C \setminus (B \cap C)\big) = \mu(B \cup C) + \mu(B \cap C),$$

and so

$$\begin{gathered} \mu(B \cup C) = \mu(C) + \mu(B) - \mu(B \cap C) \\ \text{for all } B, C \in \mathcal{B} \text{ with } \mu(B \cap C) < \infty \end{gathered} \tag{2.1.6}$$

and

$$\begin{gathered} \mu(C \setminus B) = \mu(C) - \mu(B) \\ \text{for all } B, C \in \mathcal{B} \text{ with } B \subseteq C \text{ and } \mu(B) < \infty. \end{gathered} \tag{2.1.7}$$

Finally, μ is **countably subadditive** in the sense that

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) \le \sum_{n=1}^{\infty} \mu(B_n) \quad \text{for any } \{B_n : n \ge 1\} \subseteq \mathcal{B}. \tag{2.1.8}$$

To check this, set $A_1 = B_1$ and $A_{n+1} = B_{n+1} \setminus \bigcup_{m=1}^{n} B_m$, note that the A_n's are mutually disjoint elements of $\mathcal{B}$ whose union is the same as that of the B_n's, and apply (2.1.4) and (2.1.5) to conclude that

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n) \le \sum_{n=1}^{\infty} \mu(B_n).$$

[1] Even if $E = \mathbb{R}^N$, the elements of $\mathcal{C}$ need not be rectangles.

As a consequence, the countable union of B_n's with μ-measure 0 again has μ-measure 0. More generally,

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) \tag{2.1.9}$$
$$\text{for any } \{B_n : n \geq 1\} \subseteq \mathcal{B} \text{ with } \mu(B_m \cap B_n) = 0 \text{ when } m \neq n.$$

To see this, take $C = \bigcup\{B_m \cap B_n : m \neq n\}$, use the preceding to see that $\mu(C) = 0$, set $B'_n = B_n \setminus C$, and apply (2.1.4) to $\{B'_n : n \geq 1\}$.

Another important property of measures is that they are continuous under non-decreasing limits. To explain this property, say that $\{B_n : n \geq 1\}$ **increases** to B and write $B_n \nearrow B$ if $B_{n+1} \supseteq B_n$ for all $n \in \mathbb{Z}^+$ and $B = \bigcup_{n=1}^{\infty} B_n$. Then

$$\{B_n : n \geq 1\} \subseteq \mathcal{B} \text{ and } B_n \nearrow B \implies \mu(B_n) \nearrow \mu(B). \tag{2.1.10}$$

To check this, set $A_1 = B_1$ and $A_{n+1} = B_{n+1} \setminus B_n$, note that the A_n's are mutually disjoint and $B_n = \bigcup_{m=1}^{n} A_m$, and conclude that

$$\mu(B_n) = \sum_{m=1}^{n} \mu(A_m) \nearrow \sum_{m=1}^{\infty} \mu(A_m) = \mu(B).$$

Next say that $\{B_n : n \geq 1\}$ **decreases** to B and write $B_n \searrow B$ if $B_{n+1} \subseteq B_n$ for each $n \in \mathbb{Z}^+$ and $B = \bigcap_{n=1}^{\infty} B_n$. Obviously, $B_n \searrow B$ if and only if $B_1 \setminus B_n \nearrow B_1 \setminus B$. Hence, by combining (2.1.10) with (2.1.7), one finds that

$$\{B_n : n \geq 1\} \subseteq \mathcal{B},\ B_n \searrow B, \text{ and } \mu(B_1) < \infty \implies \mu(B_n) \searrow \mu(B). \tag{2.1.11}$$

To see that the condition $\mu(B_1) < \infty$ cannot be dispensed with in general, define μ on $\big(\mathbb{Z}^+, \mathcal{P}(\mathbb{Z}^+)\big)$ to be the counting measure (i.e., $\mu(B) = \mathrm{card}(B)$ for $B \subseteq \mathbb{Z}^+$), and take $B_m = \{n \in \mathbb{Z}^+ : n \geq m\}$. Clearly $B_m \searrow \emptyset$, and yet $\mu(B_m) = \infty$ for all m.

Very often one encounters a situation in which two measures agree on a collection of sets and one wants to know that they agree on the σ-algebra generated by those sets. To handle such a situation, the following concepts are sometimes useful. A collection $\mathcal{C} \subseteq \mathcal{P}(E)$ is called a **Π-system** if it is closed under finite intersections. Given a Π-system $\mathcal{C}$, it is important to know what additional properties a Π-system must possess in order to be a σ-algebra. For this reason one introduces a notion that complements that of a Π-system. Namely, say that $\mathcal{H} \subseteq \mathcal{P}(E)$ is a **Λ-system over E** if

(**a**) $E \in \mathcal{H}$,
(**b**) $\Gamma, \Gamma' \in \mathcal{H}$ and $\Gamma \cap \Gamma' = \emptyset \implies \Gamma \cup \Gamma' \in \mathcal{H}$,
(**c**) $\Gamma, \Gamma' \in \mathcal{H}$ and $\Gamma \subseteq \Gamma' \implies \Gamma' \setminus \Gamma \in \mathcal{H}$,
(**d**) $\{\Gamma_n : n \geq 1\} \subseteq \mathcal{H}$ and $\Gamma_n \nearrow \Gamma \implies \Gamma \in \mathcal{H}$.

Notice that the collection of sets on which two finite measures agree satisfies (**b**), (**c**), and (**d**). Hence, if they agree on E, then they agree on a Λ-system.[2]

LEMMA 2.1.12. *The intersection of an arbitrary collection of* Π-*systems or of* Λ-*systems is again a* Π-*system or a* Λ-*system. Moreover,* $\mathcal{B} \subseteq \mathcal{P}(E)$ *is a* σ-*algebra over* E *if and only if it is both a* Π-*system and a* Λ-*system over* E. *Finally, if* $\mathcal{C} \subseteq \mathcal{P}(E)$ *is a* Π-*system, then* $\sigma(\mathcal{C})$ *is the smallest* Λ-*system over* E *containing* $\mathcal{C}$.

PROOF: The first assertion requires no comment. To prove the second one, it suffices to prove that if $\mathcal{B}$ is both a Π-system and a Λ-system over E, then it is a σ-algebra over E. To this end, first note that $A^{\complement} = E \setminus A \in \mathcal{B}$ for every $A \in \mathcal{B}$ and therefore that $\mathcal{B}$ is closed under complementation. Second, if $\Gamma_1, \Gamma_2 \in \mathcal{B}$, then $\Gamma_1 \cup \Gamma_2 = \Gamma_1 \cup (\Gamma_2 \setminus \Gamma_3)$ where $\Gamma_3 = \Gamma_1 \cap \Gamma_2$. Hence $\mathcal{B}$ is closed under finite unions. Finally, if $\{\Gamma_n : n \geq 1\} \subseteq \mathcal{B}$, set $A_n = \bigcup_1^n \Gamma_m$ for $n \geq 1$. Then $\{A_n : n \geq 1\} \subseteq \mathcal{B}$ and $A_n \nearrow \bigcup_1^\infty \Gamma_m$. Hence $\bigcup_1^\infty \Gamma_m \in \mathcal{B}$, and so $\mathcal{B}$ is a σ-algebra.

To prove the final assertion, let $\mathcal{C}$ be a Π-system and $\mathcal{H}$ the smallest Λ-system over E containing $\mathcal{C}$. Clearly $\sigma(\mathcal{C}) \supseteq \mathcal{H}$, and so all that we have to do is show that $\mathcal{H}$ is Π-system over E. To this end, first set

$$\mathcal{H}_1 = \{\Gamma \subseteq E : \Gamma \cap \Delta \in \mathcal{H} \text{ for all } \Delta \in \mathcal{C}\}.$$

It is then easy to check that $\mathcal{H}_1$ is a Λ -system over E. Moreover, since $\mathcal{C}$ is a Π-system, $\mathcal{C} \subseteq \mathcal{H}_1$, and therefore $\mathcal{H} \subseteq \mathcal{H}_1$. In other words, $\Gamma \cap \Delta \in \mathcal{H}$ for all $\Gamma \in \mathcal{H}$ and $\Delta \in \mathcal{C}$. Next set

$$\mathcal{H}_2 = \{\Gamma \subseteq E : \Gamma \cap \Delta \in \mathcal{H} \text{ for all } \Delta \in \mathcal{H}\}.$$

Again it is clear that $\mathcal{H}_2$ is a Λ-system. Also, by the preceding, $\mathcal{C} \subseteq \mathcal{H}_2$. Hence $\mathcal{H} \subseteq \mathcal{H}_2$, and so $\mathcal{H}$ is a Π-system. $\square$

As a consequence of Lemma 2.1.12 and the remark preceding it, one has the following important result.

THEOREM 2.1.13. *Let* $(E, \mathcal{B})$ *be a measurable space and* $\mathcal{C}$ *is* Π-*system that generates* $\mathcal{B}$. *If* μ *and* ν *are a pair of finite measures on* $(E, \mathcal{B})$ *and* $\mu(\Gamma) = \nu(\Gamma)$ *for all* $\Gamma \in \{E\} \cup \mathcal{C}$, *then* $\mu = \nu$.

PROOF: As was remarked above, additivity, (2.1.7), and (2.1.10) imply that $\mathcal{H} = \{\Gamma \in \mathcal{B} : \mu(\Gamma) = \nu(\Gamma)\}$ is a Λ-system. Hence, since $\mathcal{B} \supseteq \mathcal{H} \supseteq \mathcal{C}$, it follows from Lemma 2.1.12 that $\mathcal{H} = \mathcal{B}$. $\square$

A measure space $(E, \mathcal{B}, \mu)$ is said to be **complete** if $\Gamma \in \mathcal{B}$ whenever there exist C, $D \in \mathcal{B}$ such that $C \subseteq \Gamma \subseteq D$ with $\mu(D \setminus C) = 0$. The following simple lemma shows that every measure space can be "completed."

[2] I learned these ideas from E. B. Dynkin's treatise on Markov processes. In fact, the Λ- and Π-system scheme is often attributed to Dynkin, who certainly deserves the credit for its exploitation by a whole generation of probabilists. On the other hand, Richard Gill has informed me that, according to A.J. Lenstra, their origins go back to W. Sierpiński's article *Un théorème général sur les familles d'ensembles*, which appeared in *Fund. Math.* 12 (1928), pp. 206–210.

LEMMA 2.1.14. *Given a measure space $(E, \mathcal{B}, \mu)$, define $\overline{\mathcal{B}}^{\mu}$ to be the set of $\Gamma \subseteq E$ for which there exist C, $D \in \mathcal{B}$ satisfying $C \subseteq \Gamma \subseteq D$ and $\mu(D \setminus C) = 0$. Then $\overline{\mathcal{B}}^{\mu}$ is a σ-algebra over E and there is a unique extension $\bar{\mu}$ of μ to $\overline{\mathcal{B}}^{\mu}$ as a measure on $(E, \overline{\mathcal{B}}^{\mu})$. Furthermore, $(E, \overline{\mathcal{B}}^{\mu}, \bar{\mu})$ is a complete measure space, and if $(E, \mathcal{B}, \mu)$ is complete, then $\overline{\mathcal{B}}^{\mu} = \mathcal{B}$.*

PROOF: To see that $\overline{\mathcal{B}}^{\mu}$ is a σ-algebra, suppose that $\{\Gamma_n : n \geq 1\} \subseteq \overline{\mathcal{B}}^{\mu}$, and choose $\{C_n : n \geq 1\} \cup \{D_n : n \geq 1\} \subseteq \mathcal{B}$ accordingly. Then $C = \bigcup_{n=1}^{\infty} C_n$ and $D = \bigcup_{n=1}^{\infty} D_n$ are elements of $\mathcal{B}$, $C \subseteq \bigcup_{n=1}^{\infty} \Gamma_n \subseteq D$, and $\mu(D \setminus C) = 0$. Also, if $\Gamma \in \overline{\mathcal{B}}^{\mu}$ and C and D are associated elements of $\mathcal{B}$, then $D^{\complement}$, $C^{\complement} \in \mathcal{B}$, $D^{\complement} \subseteq \Gamma^{\complement} \subseteq C^{\complement}$, and $\mu(C^{\complement} \setminus D^{\complement}) = \mu(D \setminus C) = 0$.

Next, given $\Gamma \in \overline{\mathcal{B}}^{\mu}$, suppose that C, C', D, $D' \in \mathcal{B}$ satisfy $C \cup C' \subseteq \Gamma \subseteq D \cap D'$ and $\mu(D \setminus C) = 0 = \mu(D' \setminus C')$. Then $\mu(D' \setminus D) \leq \mu(D' \setminus C') = 0$ and so $\mu(D') \leq \mu(D) + \mu(D' \setminus D) = \mu(D)$. Similarly, $\mu(D) \leq \mu(D')$, which means that $\mu(D) = \mu(D')$, and, because $\mu(C) = \mu(D)$ and $\mu(C') = \mu(D')$, it follows that μ assigns the same measure to C, C', D and D'. Hence, we can unambiguously define $\bar{\mu}(\Gamma) = \mu(C) = \mu(D)$ when $\Gamma \in \overline{\mathcal{B}}^{\mu}$ and C, $D \in \mathcal{B}$ satisfy $C \subseteq \Gamma \subseteq D$ with $\mu(D \setminus C) = 0$. Furthermore, if $\{\Gamma_n : n \geq 1\}$ are mutually disjoint elements of $\overline{\mathcal{B}}^{\mu}$ and $\{C_n : n \geq 1\} \cup \{D_n : n \geq 1\} \subseteq \mathcal{B}$ are chosen accordingly, then the C_n's are mutually disjoint, and so

$$\bar{\mu}\left(\bigcup_{n=1}^{\infty} \Gamma_n\right) = \mu\left(\bigcup_{n=1}^{\infty} C_n\right) = \sum_{n=1}^{\infty} \mu(C_n) = \sum_{n=1}^{\infty} \bar{\mu}(\Gamma_n).$$

Hence, $\bar{\mu}$ is a measure on $(E, \overline{\mathcal{B}}^{\mu})$ whose restriction to $\mathcal{B}$ coincides with μ.

Finally, suppose that ν is any measure on $(E, \overline{\mathcal{B}}^{\mu})$ that extends μ. If Γ is a subset of E for which there exist Γ', $\Gamma'' \in \overline{\mathcal{B}}^{\mu}$ satisfying $\Gamma' \subseteq \Gamma \subseteq \Gamma''$ and $\nu(\Gamma'' \setminus \Gamma') = 0$, there exist C, $D \in \mathcal{B}$ satisfying $C \subseteq \Gamma'$ and $\Gamma'' \subseteq D$ such that $\nu(D \setminus \Gamma'') = 0 = \nu(\Gamma' \setminus C)$ and therefore

$$\mu(D \setminus C) = \nu(D \setminus C) = \nu(D \setminus \Gamma'') + \nu(\Gamma'' \setminus \Gamma') + \nu(\Gamma' \setminus C) = 0.$$

Hence, $\Gamma \in \overline{\mathcal{B}}^{\mu}$ and $\nu(\Gamma) = \mu(C) = \bar{\mu}(\Gamma)$. Thus, we now know that $\bar{\mu}$ is the only extension of μ as a measure on $(E, \overline{\mathcal{B}}^{\mu})$ and that $\overline{\mathcal{B}}^{\mu} = \mathcal{B}$ if $(E, \mathcal{B}, \mu)$ is complete. $\square$

The measure space $(E, \overline{\mathcal{B}}^{\mu}, \bar{\mu})$ is called the **completion** of $(E, \mathcal{B}, \mu)$, and Lemma 2.1.14 says that every measure space has a unique completion. Elements of $\overline{\mathcal{B}}^{\mu}$ are said to be **μ-measurable**.

Given a topological space E, use $\mathfrak{G}(E)$ to denote the class of all open subsets of E and $\mathfrak{G}_{\delta}(E)$ the class of subsets that can be written as the countable intersection of open subsets. Analogously, $\mathfrak{F}(E)$ and $\mathfrak{F}_{\sigma}(E)$ will denote, respectively, the class of all closed subsets of E and the class of subsets that can be written as the countable union of closed subsets. Clearly $B \in \mathfrak{G}(E) \iff B^{\complement} \in \mathfrak{F}(E)$, $B \in \mathfrak{G}_{\delta}(E) \iff B^{\complement} \in \mathfrak{F}_{\sigma}(E)$, and

$\mathfrak{G}_\delta(E) \cup \mathfrak{F}_\sigma(E) \subseteq \mathcal{B}_E$. Moreover, when the topology of E admits a metric ρ, it is easy to check that $\mathfrak{F}(E) \subseteq \mathfrak{G}_\delta(E)$ and therefore $\mathfrak{G}(E) \subseteq \mathfrak{F}_\sigma(E)$. Indeed, if $F \in \mathfrak{F}(E)$, then

$$\mathfrak{G}(E) \ni \{x : \rho(x, F) < \tfrac{1}{n}\} \searrow F \quad \text{as } n \to \infty.$$

Finally, a measure μ on $(E, \mathcal{B}_E)$ is called a **Borel measure** on E, and if μ is a Borel measure on E, a set $\Gamma \subseteq E$ is said to be **μ-regular** when, for each $\epsilon > 0$, there exist $F \in \mathfrak{F}(E)$ and $G \in \mathfrak{G}(E)$ such that $F \subseteq \Gamma \subseteq G$ and $\mu(G \setminus F) < \epsilon$. A Borel measure μ is said to be **regular** if every element of $\mathcal{B}_E$ is μ-regular.

Theorem 2.1.15. *Let E be a topological space and μ a Borel measure on E. If $\Gamma \subseteq E$ is μ-regular, then there exist $C \in \mathfrak{F}_\sigma(E)$ and $D \in \mathfrak{G}_\delta(E)$ for which $C \subseteq \Gamma \subseteq D$ and $\mu(D \setminus C) = 0$. In particular, $\Gamma \in \overline{\mathcal{B}_E}^\mu$ if Γ is μ-regular. Conversely, if μ is regular, then every element of $\overline{\mathcal{B}_E}^\mu$ is μ-regular. Moreover, if the topology on E admits a metric space and μ is a finite Borel measure on E, then μ is regular. (See Exercise 2.1.20 for a small extension.)*

Proof: To prove the first part, suppose that $\Gamma \subseteq E$ is μ-regular. Then, for each $n \geq 1$, there exist $F_n \in \mathfrak{F}(E)$ and $G_n \in \mathfrak{G}(E)$ such that $F_n \subseteq \Gamma \subseteq G_n$ and $\mu(G_n \setminus F_n) < \frac{1}{n}$. Thus, if $C = \bigcup_{n=1}^\infty F_n$ and $D = \bigcap_{n=1}^\infty G_n$, then $C \in \mathfrak{F}_\sigma(E)$, $D \in \mathfrak{G}_\delta(E)$, $C \subseteq \Gamma \subseteq D$, and, because $D \setminus C \subseteq G_n \setminus F_n$ for all n, $\mu(D \setminus C) = 0$. Obviously, this means that $\Gamma \in \overline{\mathcal{B}_E}^\mu$. Conversely, if μ is regular and $\Gamma \in \overline{\mathcal{B}_E}^\mu$, then there exist $\Gamma', \Gamma'' \in \mathcal{B}_E$ for which $\Gamma' \subseteq \Gamma \subseteq \Gamma''$, $\mu(\Gamma'' \setminus \Gamma') = 0$. By regularity, for each $\epsilon > 0$, there exist $F \in \mathfrak{F}(E)$, $G \in \mathfrak{G}(E)$ such that $F \subseteq \Gamma'$, $\Gamma'' \subseteq G$, and $\bar{\mu}(G \setminus \Gamma'') \vee \bar{\mu}(\Gamma' \setminus F) < \frac{\epsilon}{2}$. Hence, $F \subseteq \Gamma \subseteq G$ and

$$\mu(G \setminus F) = \bar{\mu}(G \setminus \Gamma'') + \mu(\Gamma'' \setminus \Gamma') + \mu(\Gamma' \setminus C) < \epsilon,$$

and so Γ is μ-regular.

Now suppose the E admits a metric and that μ is finite, and let $\mathcal{R}$ be the collection of $B \in \mathcal{B}_E$ that are μ-regular. If we show that $\mathcal{R}$ is a σ-algebra that contains $\mathfrak{G}(E)$, then we will know that $\mathcal{R} = \mathcal{B}_E$ and therefore that μ is regular. Obviously $\mathcal{R}$ is closed under complementation. Next, suppose that $\{B_n : n \geq 1\} \subseteq \mathcal{R}$, and set $B = \bigcup_{n=1}^\infty B_n$. Given $\epsilon > 0$, for each n choose $F_n \in \mathfrak{F}(E)$ and $G_n \in \mathfrak{G}(E)$ such that $F_n \subseteq B_n \subseteq G_n$ and $\mu(G_n \setminus F_n) < 2^{-n-1}\epsilon$. Then $\mathfrak{G}(E) \ni G = \bigcup_{m=1}^\infty G_m \supseteq B$, $\mathfrak{F}_\sigma(E) \ni C = \bigcup_{m=1}^\infty F_m \subseteq B$, and

$$\mu(G \setminus C) \leq \mu\left(\bigcup_{m=1}^\infty (G_m \setminus F_m)\right) \leq \sum_{m=1}^\infty \mu(G_m \setminus F_m) < \frac{\epsilon}{2}.$$

Finally, because $\mu(C) < \infty$ and $C \setminus \bigcup_{m=1}^n F_m \searrow \emptyset$, (2.1.11) allows us to choose an $n \in \mathbb{Z}^+$ for which $\mu(C \setminus F) < \frac{\epsilon}{2}$ when $F = \bigcup_{m=1}^n F_m \in \mathfrak{F}(E)$. Hence, since $\mu(G \setminus F) = \mu(G \setminus C) + \mu(C \setminus F) < \epsilon$, we know that $B \in \mathcal{R}$ and therefore that $\mathcal{R}$ is a σ-algebra.

To complete the proof, it remains to show that $\mathfrak{G}(E) \subseteq \mathcal{R}$, and clearly this comes down to showing that for each open G and $\epsilon > 0$ there is a closed

$F \subseteq G$ for which $\mu(G\backslash F) < \epsilon$. But, because E has a metric topology, we know that $\mathfrak{G}(E) \subseteq \mathfrak{F}_\sigma(E)$. Hence, if G is open, then there exists a non-decreasing sequence $\{F_n : n \geq 1\} \subseteq \mathfrak{F}(E)$ such that $F_n \nearrow G$, which, because $\mu(G) < \infty$, means that $\mu(G \setminus F_n) \searrow 0$. Thus, for any $\epsilon > 0$, there is an n for which $\mu(G \setminus F_n) < \epsilon$. $\square$

Exercises for § 2.1

EXERCISE 2.1.16. The decomposition of the properties of a σ-algebra in terms of Π-systems and Λ-systems is not the traditional one. Instead, most of the early books on measure theory used algebras instead of Π-systems as the standard source of generating sets. An **algebra** over E is a collection $\mathcal{A} \subseteq \mathcal{P}(E)$ that contains E and is closed under finite unions and complementation. If one starts with an algebra $\mathcal{A}$, then the complementary notion is that of a **monotone class**: $\mathcal{M}$ is said to be a monotone class if $\Gamma \in \mathcal{M}$ whenever there exists $\{\Gamma_n : n \geq 1\} \subseteq \mathcal{M}$ such that $\Gamma_n \nearrow \Gamma$. Show that $\mathcal{B}$ is a σ-algebra over E if and only if it is both an algebra over E and a monotone class. In addition, show that if $\mathcal{A}$ is an algebra over E, then $\sigma(\mathcal{A})$ is the smallest monotone class containing $\mathcal{A}$.

EXERCISE 2.1.17. If $f : \mathbb{R} \longrightarrow \mathbb{R}$ is either right continuous or left continuous, show that f is $\mathcal{B}_\mathbb{R}$-measurable.

EXERCISE 2.1.18. Given a measurable space $(E, \mathcal{B})$ and $\emptyset \neq E' \subseteq E$, show that $\mathcal{B}[E'] \equiv \{B \cap E' : B \in \mathcal{B}\}$ is a σ-algebra over E'. Further, show that if E is a topological space, then $\mathcal{B}_E[E'] = \mathcal{B}_{E'}$ when E' is given the topology that it inherits from E. Finally, if $\emptyset \neq E' \in \mathcal{B}$, show that $\mathcal{B}[E'] \subseteq \mathcal{B}$ and that the restriction to $\mathcal{B}[E']$ of any measure on $(E, \mathcal{B})$ is a measure on $\big(E', \mathcal{B}[E']\big)$. In particular, if E is a topological space and μ is a Borel measure on E, show that $\mu \restriction \mathcal{B}_{E'}$ is a Borel measure on E' and that it is regular if μ is regular.

EXERCISE 2.1.19. Given a map $\Phi : E \longrightarrow E'$, define $\Phi(\Gamma) = \{\Phi(x) : x \in \Gamma\}$ for $\Gamma \subseteq E$ and $\Phi^{-1}(\Gamma') = \{x \in E : \Phi(x) \in \Gamma'\}$ for $\Gamma' \subseteq E'$.

(**i**) Show that Φ and Φ^{-1} preserve unions in the sense that $\Phi\left(\bigcup_\alpha B_\alpha\right) = \bigcup_\alpha \Phi(B_\alpha)$ and $\Phi^{-1}\left(\bigcup_\alpha B'_\alpha\right) = \bigcup_\alpha \Phi^{-1}(B'_\alpha)$. In addition, show that Φ^{-1} preserves differences in the sense that $\Phi^{-1}(B' \setminus A') = \Phi^{-1}(B') \setminus \Phi^{-1}(A')$. On the other hand, show that Φ need not preserve differences, but that it will if it is one-to-one.

(**ii**) Suppose that $\mathcal{B}$ and $\mathcal{B}'$ are σ-algebras over, respectively, E and E' and that $\Phi : E \longrightarrow E'$. If $\mathcal{B}' = \sigma(\mathcal{C}')$ and $\Phi^{-1}(C') \in \mathcal{B}$ for every $C' \in \mathcal{C}'$, show that Φ is measurable. In particular, if E and E' are topological spaces and Φ is continuous, show that Φ is measurable as a map from $(E, \mathcal{B}_E)$ to $(E', \mathcal{B}_{E'})$. Similarly, if Φ is one-to-one and $\mathcal{B} = \sigma(\mathcal{C})$, show that $\Phi(B) \in \mathcal{B}'$ for all $B \in \mathcal{B}$ if $\Phi(C) \in \mathcal{B}'$ for all $C \in \mathcal{C} \cup \{E\}$.

(**iii**) Now suppose that μ is a measure on $(E, \mathcal{B})$ and that Φ is a measurable map from $(E, \mathcal{B})$ into $(E', \mathcal{B}')$. Define $\mu'(B') = \mu\big(\Phi^{-1}(B')\big)$ for $B' \in \mathcal{B}'$, and show that μ' is a measure on $(E', \mathcal{B}')$. This measure μ' is called the **pushforward** or **image** of μ under Φ and is denoted by either $\Phi_*\mu$ or $\mu \circ \Phi^{-1}$. Similarly, if $\Phi : E \longrightarrow E'$ is one-to-one and takes elements of $\mathcal{B}$ to elements of $\mathcal{B}'$ and if μ' is a measure on $(E', \mathcal{B}')$, show that $\Gamma \in \mathcal{B} \longmapsto \mu'\big(\Phi(\Gamma)\big) \in [0, \infty]$ is a measure on $(E, \mathcal{B})$. This measure is the **pullback** of μ' under Φ.

EXERCISE 2.1.20. Let E be a topological space and μ a Borel measure on E. Show that μ is regular if, for every $\Gamma \in \mathcal{B}_E$ and $\epsilon > 0$, there is an open $G \supseteq \Gamma$ for which $\mu(G \setminus \Gamma) < \epsilon$. In addition, if E is a metric space and there exists a non-decreasing sequence $\{G_n : n \geq 1\}$ of open sets such that $G_n \nearrow E$ and $\mu(G_n) < \infty$ for each $n \in \mathbb{Z}^+$, show that μ is regular.

EXERCISE 2.1.21. Let $(E, \mathcal{B}, \mu)$ be a finite measure space. Given $n \geq 2$ and $\{\Gamma_m : 1 \leq m \leq n\} \subseteq \mathcal{B}$, use (2.1.6) and induction to show that

$$\mu(\Gamma_1 \cup \cdots \cup \Gamma_n) = -\sum_F (-1)^{\mathrm{card}(F)} \mu\big(\Gamma_F\big),$$

where the summation is over non-empty subsets F of $\{1, \ldots, n\}$ and $\Gamma_F \equiv \bigcap_{i \in F} \Gamma_i$. Although this formula is seldom used except in the case $n = 2$, the following is an interesting application of the general result. Let E be the group of permutations of $\{1, \ldots, n\}$, $\mathcal{B} = \mathcal{P}(E)$, and $\mu\big(\{\pi\}\big) = \frac{1}{n!}$ for each $\pi \in E$. Denote by A the set of $\pi \in E$ such that $\pi(i) \neq i$ for any $1 \leq i \leq n$. Then one can interpret $\mu(A)$ as the probability that, when the numbers $1, \ldots, n$ are randomly ordered, none of them is placed in the correct position. On the basis of this interpretation, one might suspect that $\mu(A)$ should tend to 0 as $n \to \infty$. However, by direct computation, one can see that this is not the case. Indeed, let Γ_i be the set of $\pi \in E$ for which $\pi(i) = i$. Then $A = \big(\Gamma_1 \cup \cdots \cup \Gamma_n\big)^{\complement}$, and therefore

$$\mu(A) = 1 - \mu(\Gamma_1 \cup \cdots \cup \Gamma_n) = 1 + \sum_F (-1)^{\mathrm{card}(F)} \mu\big(\Gamma_F\big).$$

Show that $\mu(\Gamma_F) = \frac{(n-m)!}{n!}$ if $\mathrm{card}(F) = m$, and conclude from this that $\mu(A) = \sum_0^n \frac{(-1)^m}{m!} \longrightarrow \frac{1}{e}$ as $n \to \infty$.

EXERCISE 2.1.22. Given a sequence $\{B_n : n \geq 1\}$ of sets, define their **limit inferior** to be the set

$$\varliminf_{n \to \infty} B_n = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} B_n,$$

or, equivalently, the set of $x \in E$ that are in all but finitely many of the B_n's. Also, define their **limit superior** to be the set

$$\varlimsup_{n \to \infty} B_n = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} B_n,$$

or, equivalently, the set of x that are in infinitely many of the B_n's. Show that $\varliminf_{n\to\infty} B_n \subseteq \varlimsup_{n\to\infty} B_n$, and say that the sequence $\{B_n : n \geq 1\}$ has a **limit** if equality holds, in which case $\varliminf_{n\to\infty} B_n = \varlimsup_{n\to\infty} B_n$ is said to be the limit $\lim_{n\to\infty} B_n$ of $\{B_n : n \geq 1\}$. Show that if $(E, \mathcal{B}, \mu)$ is a measure space and $\{B_n : n \geq 1\} \subseteq \mathcal{B}$, both $\varliminf_{n\to\infty} B_n \in \mathcal{B}$ and $\varliminf_{n\to\infty} B_n$ are elements of $\mathcal{B}$. Further, show that

$$\mu\left(\varliminf_{n\to\infty} B_n\right) \leq \varliminf_{n\to\infty} \mu(B_n) \tag{2.1.23}$$

and that

$$\mu\left(\varlimsup_{n\to\infty} B_n\right) \geq \varlimsup_{n\to\infty} \mu(B_n) \quad \text{if } \mu\left(\bigcup_{n=1}^{\infty} B_n\right) < \infty. \tag{2.1.24}$$

Conclude that

$$\lim_{n\to\infty} B_n \text{ exists } \& \ \mu\left(\bigcup_{n=1}^{\infty} B_n\right) < \infty \implies \mu\left(\lim_{n\to\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n). \tag{2.1.25}$$

Hint: Note that $\bigcap_{n=m}^{\infty} B_n \nearrow \varliminf_{n\to\infty} B_n$ and $\bigcup_{n=m}^{\infty} B_n \searrow \varlimsup_{n\to\infty} B_n$ as $m \to \infty$.

EXERCISE 2.1.26. Let $(E, \mathcal{B}, \mu)$ be a measure space and $\{B_n : n \geq 1\} \subseteq \mathcal{B}$. Show that

$$\sum_{n=1}^{\infty} \mu(B_n) < \infty \implies \mu\left(\varlimsup_{n\to\infty} B_n\right) = 0.$$

This useful observation is usually attributed to E. Borel. More profound is the following converse statement, which is due to F. Cantelli. Assume that μ is a probability measure. Sets $\{B_n : n \geq 1\} \subseteq \mathcal{B}$ are said to be **independent** under μ or **μ-independent** if, for all $n \geq 1$ and choices of $C_m \in \{B_m, B_m^{\complement}\}$, $1 \leq m \leq n$, $\mu(C_1 \cap \cdots \cap C_n) = \mu(C_1) \cdots \mu(C_n)$. Cantelli's result says that if $\{B_n : n \geq 1\} \subseteq \mathcal{B}$ are μ-independent sets, then

$$\sum_{n=1}^{\infty} \mu(B_n) = \infty \implies \mu\left(\varlimsup_{n\to\infty} B_n\right) = 1.$$

Thus, for μ-independent sets, $\mu\left(\varlimsup_{n\to\infty} B_n\right)$ is either 0 or 1 according to whether $\sum_{n=1}^{\infty} \mu(B_n)$ is finite or infinite, a result that is referred to as the **Borel–Cantelli Lemma**. Give a proof of Cantelli's result. In doing so, the following outline might be helpful.

(**i**) Show that it suffices to prove that $\lim_{N\to\infty} \mu\left(\bigcap_{n=m}^{N} B_n^{\complement}\right) = 0$ for each $m \in \mathbb{Z}^+$.

(**ii**) Show that $1 - x \le e^{-x}$ for all $x \ge 0$, and use this to check that

$$\mu\left(\bigcap_{n=m}^{N} B_n^{\complement}\right) \le e^{-\sum_{n=m}^{N} \mu(B_n)} \quad \text{for } N \ge m.$$

EXERCISE 2.1.27. Given a pair of measures μ and ν on a measurable space $(E, \mathcal{B})$, one says that μ is **absolutely continuous** with respect to ν and writes $\mu \ll \nu$ if, for all $B \in \mathcal{B}$, $\nu(B) = 0 \implies \mu(B) = 0$. Assuming that μ is finite and that $\mu \ll \nu$, show that for each $\epsilon > 0$ there exists a $\delta > 0$ such that $\mu(B) < \epsilon$ whenever $\nu(B) < \delta$. Next, assume that E is a metric space, that both μ and ν are regular Borel measures on E, and that μ is finite. Show that $\mu \ll \nu$ if and only if for every $\epsilon > 0$ there exists a $\delta > 0$ for which $\nu(G) < \delta \implies \mu(G) < \epsilon$ whenever $G \in \mathfrak{G}(E)$.

EXERCISE 2.1.28. A pair of measures μ and ν on a measurable space $(E, \mathcal{B})$ are said to be **singular** to one another and one writes $\mu \perp \nu$ if there exists a $B \in \mathcal{B}$ such that $\mu(B) = 0 = \nu(B^{\complement})$. In words, μ and ν are singular to one another if they live on disjoint parts of E. Assuming that E is a metric space, that ν is regular Borel measure on E, and that μ is finite, show that $\mu \perp \nu$ if and only if for every $\delta > 0$ there is a $G \in \mathfrak{G}(E)$ for which $\nu(G) < \delta$ and $\mu(G^{\complement}) = 0$.

§ 2.2 A Construction of Measures

In this section I will first develop a procedure for constructing measures and will then apply that procedure to three important examples.

§ **2.2.1. A Construction Procedure:** Suppose that $\mathfrak{R}$ is a collection of compact subsets I of a metric space (E, ρ) and that V is a map from $\mathfrak{R}$ to $[0, \infty)$ that satisfy the following conditions:

(1) $\emptyset \in \mathfrak{R}$ and $I, I' \in \mathfrak{R} \implies I \cap I' \in \mathfrak{R}$.
(2) $V(\emptyset) = 0$ and $V(I) \le V(J)$ if $I, J \in \mathfrak{R}$ and $I \subseteq J$.
(3) For any $J \in \mathfrak{R}$, $n \in \mathbb{Z}^+$, and $\{I_1, \dots, I_n\} \subseteq \mathfrak{R}$, $V(J) \le \sum_{m=1}^{n} V(I_m)$ if $J \subseteq \bigcup_{m=1}^{n} I_m$ and $V(J) \ge \sum_{m=1}^{n} V(I_m)$ if the I_m's are non-overlapping (i.e., their interiors are mutually disjoint) and $J \supseteq \bigcup_{m=1}^{n} I_m$.
(4) For any $I \in \mathfrak{R}$ and $\epsilon > 0$, there exist $I', I'' \in \mathfrak{R}$ such that $I'' \subseteq \mathring{I}$, $I \subseteq \mathring{I}'$, and $V(I') \le V(I'') + \epsilon$.
(5) For any $G \in \mathfrak{G}(E)$, there is a sequence $\{I_n : n \ge 1\}$ of non-overlapping elements of $\mathfrak{R}$ such that $G = \bigcup_{n=1}^{\infty} I_n$.

An example to keep in mind is that for which $E = \mathbb{R}^N$, $\mathfrak{R}$ is the collection of all closed rectangles in $\mathbb{R}^N$ (one should consider $\emptyset$ to be a rectangle), and $V(I) = \mathrm{vol}(I)$.

The goal of this subsection is to prove that there is a unique Borel measure μ on E such that $\mu(I) = V(I)$ for all $I \in \mathfrak{R}$. Before getting started with the proof, recall the following elementary fact about double sums of non-negative numbers.

LEMMA 2.2.1. *If* $\{a_{m,n} : (m,n) \in (\mathbb{Z}^+)^2\} \subseteq [0,\infty)$, *then*

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{m,n} = \sum_{(m,n)\in(\mathbb{Z}^+)^2} a_{m,n} = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{m,n}.$$

PROOF: For each M, $N \in \mathbb{Z}^+$,

$$\left(\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{m,n}\right) \wedge \left(\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{m,n}\right) \geq \sum_{\{(m,n):\, m\leq M \;\&\; n\leq N\}} a_{m,n},$$

and therefore

$$\left(\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{m,n}\right) \wedge \left(\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{m,n}\right) \geq \sum_{(m,n)\in(\mathbb{Z}^+)^2} a_{m,n}.$$

Similarly,

$$\sum_{(m,n)\in(\mathbb{Z}^+)^2} a_{m,n} \geq \left(\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{m,n}\right) \vee \left(\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{m,n}\right),$$

and so all three must be equal. $\square$

Now define $\tilde{\mu}(\Gamma)$ for $\Gamma \subseteq E$ to be the infimum of $\sum_{m=1}^{\infty} V(I_m)$ for all choices of $\{I_m : m \geq 1\} \subseteq \mathfrak{R}$ that cover Γ (i.e., $\Gamma \subseteq \bigcup_{m=1}^{\infty} I_m$). My strategy is to find a σ-algebra $\mathcal{L} \supseteq \mathcal{B}_E$ for which the restriction μ of $\tilde{\mu}$ to $\mathcal{L}$ is a measure. Thus, the first thing that I have to check is that $\tilde{\mu}(I) = V(I)$ for all $I \in \mathfrak{R}$.

LEMMA 2.2.2. *If* $L \in \mathbb{Z}^+$ *and* $\Gamma = \bigcup_{\ell=1}^{L} J_\ell$ *where the* J_m*'s are non-overlapping elements of* $\mathfrak{R}$, *then* $\tilde{\mu}(\Gamma) = \sum_{\ell=1}^{L} V(J_\ell)$. *In particular,* $\tilde{\mu}(I) = V(I)$ *for all* $I \in \mathfrak{R}$.

PROOF: Obviously $\tilde{\mu}(\Gamma) \leq \sum_{\ell=1}^{L} V(J_\ell)$. To prove the opposite inequality, let $\{I_m : m \geq 1\}$ be a cover of Γ by elements of $\mathfrak{R}$. Given an $\epsilon > 0$, choose I'_m for each $m \in \mathbb{Z}^+$ so that $I_m \subseteq \mathring{I}'_m$ and $V(I'_m) \leq V(I_m) + 2^{-m}\epsilon$. Because Γ is compact, there exists an $n \in \mathbb{Z}^+$ such that $\{I'_1, \dots, I'_n\}$ covers Γ.

Next, set $I_{m,\ell} = I'_m \cap J_\ell$ for $1 \leq m \leq n$ and $1 \leq \ell \leq L$. Then, for each ℓ, $J_\ell = \bigcup_{m=1}^{n} I_{m,\ell}$, and, for each m, the $I_{m,\ell}$'s are non-overlapping elements of $\mathfrak{R}$ with $\bigcup_{\ell=1}^{L} I_{m,\ell} \subseteq I'_m$. Hence, by (3),

$$\sum_{m=1}^{\infty} V(I_m) + \epsilon \geq \sum_{m=1}^{n} V(I'_m) \geq \sum_{m=1}^{n}\sum_{\ell=1}^{L} V(I_{m,\ell}) \geq \sum_{\ell=1}^{L} V(J_\ell). \quad \square$$

In view of Lemma 2.2.2, I am justified in replacing $V(I)$ by $\tilde{\mu}(I)$ for $I \in \mathfrak{R}$.

The next result shows that half the equality in (2.1.4) is automatic, even before one restricts to Γ's from $\mathcal{L}$.

Lemma 2.2.3. *If $\Gamma \subseteq \Gamma'$, then $\tilde{\mu}(\Gamma) \le \tilde{\mu}(\Gamma')$. In fact, if $\Gamma \subseteq \bigcup_{n=1}^{\infty} \Gamma_n$, then $\tilde{\mu}(\Gamma) \le \sum_{n=1}^{\infty} \tilde{\mu}(\Gamma_n)$. In particular, if $\Gamma \subseteq \bigcup_{n=1}^{\infty} \Gamma_n$ and $\tilde{\mu}(\Gamma_n) = 0$ for each $n \ge 1$, then $\tilde{\mu}(\Gamma) = 0$.*

Proof: The first assertion follows immediately from the fact that every cover of Γ' is also a cover of Γ.

In order to prove the second assertion, let $\epsilon > 0$ be given, and choose for each $n \ge 1$ a cover $\{I_{m,n} : m \ge 1\} \subseteq \mathfrak{R}$ of Γ_n satisfying $\sum_{m=1}^{\infty} V(I_{m,n}) \le \tilde{\mu}(\Gamma_n) + 2^{-n}\epsilon$. It is obvious that $\{I_{m,n} : (m,n) \in (\mathbb{Z}^+)^2\}$ is a countable cover of Γ. Hence, by Lemma 2.2.1,

$$\tilde{\mu}(\Gamma) \le \sum_{(m,n)\in(\mathbb{Z}^+)^2} V(I_{m,n}) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} V(I_{m,n}) \le \sum_{n=1}^{\infty} \tilde{\mu}(\Gamma_n) + \epsilon. \quad \square$$

As a consequence of Lemma 2.2.3, one has that

$$\tilde{\mu}(\Gamma) = \inf\{\tilde{\mu}(G) : \Gamma \subseteq G \in \mathfrak{G}(E)\}. \tag{2.2.4}$$

To see this, note that the left-hand side is certainly dominated by the right. Thus, it suffices to show that if $\{I_m : m \ge 1\}$ is a cover of Γ by elements of $\mathfrak{R}$ and $\epsilon > 0$, then there is a $\mathfrak{G}(E) \ni G \supseteq \Gamma$ such that $\tilde{\mu}(G) \le \sum_{m=1}^{\infty} V(I_m) + \epsilon$. To this end, for each m choose $I'_m \in \mathfrak{R}$ such that $I_m \subseteq \mathring{I}'_m$ and $V(I'_m) \le V(I_m) + 2^{-m}\epsilon$, and take $G = \bigcup_{m=1}^{\infty} \mathring{I}'_m$. Clearly $\Gamma \subseteq G \in \mathfrak{G}(E)$ and

$$\tilde{\mu}(G) \le \sum_{m=1}^{\infty} V(I'_m) \le \sum_{m=1}^{\infty} V(I_m) + \epsilon.$$

One important consequence of (2.2.4) is that it shows that for any $\Gamma \subseteq E$ there is a $\mathfrak{G}_\delta(E) \ni D \supseteq \Gamma$ for which $\tilde{\mu}(D) = \tilde{\mu}(\Gamma)$. Indeed, simply choose $\Gamma \subseteq G_n \in \mathfrak{G}(E)$ for which $\tilde{\mu}(G_n) \le \tilde{\mu}(\Gamma) + \frac{1}{n}$, and take $D = \bigcap_{n=1}^{\infty} G_n$.

Another virtue of (2.2.4) is that it facilitates the proof of the second part of the following preliminary additivity result about $\tilde{\mu}$.

Lemma 2.2.5. *If G and G' are disjoint open subsets of E, then $\tilde{\mu}(G \cup G') = \tilde{\mu}(G) + \tilde{\mu}(G')$. Also, if K and K' are disjoint compact subsets of E, then $\tilde{\mu}(K \cup K') = \tilde{\mu}(K) + \tilde{\mu}(K')$.*

Proof: We begin by showing that if $\{I_m : m \ge 1\}$ is a sequence of non-overlapping elements of $\mathfrak{R}$, then

$$\tilde{\mu}\left(\bigcup_{m=1}^{\infty} I_m\right) = \sum_{m=1}^{\infty} V(I_m). \tag{2.2.6}$$

Because the left-hand side is dominated by the right, it suffices to show that the right-hand side is dominated by the left. However, by Lemmas 2.2.3 and 2.2.2, for each $n \in \mathbb{Z}^+$,

$$\tilde{\mu}\left(\bigcup_{m=1}^{\infty} I_m\right) \ge \tilde{\mu}\left(\bigcup_{m=1}^{n} I_m\right) = \sum_{m=1}^{n} V(I_m),$$

which completes the proof of (2.2.6).

Next, suppose that G and G' are disjoint open sets. By (5), there exist non-overlapping sequences $\{I_m : m \geq 1\}$ and $\{I'_m : m \geq 1\}$ of elements of $\mathfrak{R}$ such that $G = \bigcup_{m=1}^{\infty} I_m$ and $G' = \bigcup_{m=1}^{\infty} I'_{m'}$. Thus, if $I''_{2m-1} = I_m$ and $I''_{2m} = I'_m$ for $m \geq 1$, the I''_m's are non-overlapping elements of $\mathfrak{R}$ whose union is $G \cup G'$. Hence, by (2.2.6),

$$\tilde{\mu}(G \cup G') = \sum_{m=1}^{\infty} V(I''_m) = \sum_{m=1}^{\infty} V(I_m) + \sum_{m=1}^{\infty} V(I'_m) = \tilde{\mu}(G) + \tilde{\mu}(G').$$

To complete the proof, let K and K' be given. Because they are disjoint and compact, there exist disjoint open sets G and G' such that $K \subseteq G$ and $K' \subseteq G'$. Thus, for any open $H \supseteq K \cup K'$,

$$\tilde{\mu}(H) \geq \tilde{\mu}\big((H \cap G) \cup (H \cap G')\big) = \tilde{\mu}(H \cap G) + \tilde{\mu}(H \cap G') \geq \tilde{\mu}(K) + \tilde{\mu}(K'),$$

and therefore, by (2.2.4), $\tilde{\mu}(K \cup K') \geq \tilde{\mu}(K) + \tilde{\mu}(K')$. Because the opposite inequality always holds, there is nothing more to do. $\square$

I am at last ready to describe the σ-algebra $\mathcal{L}$, although it will not be immediately obvious that it is a σ-algebra or that $\tilde{\mu}$ is countably additive on it. Be that as it may, take $\mathcal{L}$ to be the collection of $\Gamma \subseteq E$ with the property that, for each $\epsilon > 0$, there is an open $G \supseteq \Gamma$ for which $\tilde{\mu}(G \setminus \Gamma) < \epsilon$.

At first, one might be tempted to say that, in view of (2.2.4), every subset Γ is an element of $\mathcal{L}$. This is because one is inclined to think that $\tilde{\mu}(G) = \tilde{\mu}(G \setminus \Gamma) + \tilde{\mu}(\Gamma)$ when, in fact, $\tilde{\mu}(G) \leq \tilde{\mu}(G \setminus \Gamma) + \tilde{\mu}(\Gamma)$ is all that we know in general. Therein lies the subtlety of the definition! Nonetheless, it is clear that $\mathfrak{G}(E) \subseteq \mathcal{L}$. Furthermore, if $\tilde{\mu}(\Gamma) = 0$, then $\Gamma \in \mathcal{L}$, since, by (2.2.4), one can choose, for any $\epsilon > 0$, an open $G \supseteq \Gamma$ such that $\tilde{\mu}(G \setminus \Gamma) \leq \tilde{\mu}(G) < \epsilon$. Finally, if $\Gamma \in \mathcal{L}$, then there is a $D \in \mathfrak{G}_\delta(E)$ for which $\Gamma \subseteq D$ and $\tilde{\mu}(D \setminus \Gamma) = 0$. Indeed, simply choose $\{G_n : n \geq 1\} \subseteq \mathfrak{G}(\mathbb{R}^N)$ for which $\Gamma \subseteq G_n$ and $\tilde{\mu}(G_n \setminus \Gamma) < \frac{1}{n}$, and take $D = \bigcap_1^{\infty} G_n$.

The next result shows that $\mathcal{L}$ is closed under countable unions.

LEMMA 2.2.7. *If $\{\Gamma_n : n \geq 1\} \subseteq \mathcal{L}$, then $\Gamma = \bigcup_1^{\infty} \Gamma_n \in \mathcal{L}$, and, of course (cf. Lemma 2.2.3), $\tilde{\mu}(\Gamma) \leq \sum_1^{\infty} \tilde{\mu}(\Gamma_n)$.*

PROOF: For each $n \geq 1$, choose $\mathfrak{G}(E) \ni G_n \supseteq \Gamma_n$ so that $\tilde{\mu}(G_n \setminus \Gamma_n) < 2^{-n}\epsilon$. Then $G \equiv \bigcup_1^{\infty} G_n$ is open, contains Γ, and (by Lemma 2.2.3) satisfies

$$\tilde{\mu}(G \setminus \Gamma) \leq \tilde{\mu}\left(\bigcup_{n=1}^{\infty} (G_n \setminus \Gamma_n)\right) \leq \sum_1^{\infty} \tilde{\mu}(G_n \setminus \Gamma_n) < \epsilon. \quad \square$$

Knowing that $\mathcal{L}$ is closed under countable unions and that it contains $\mathfrak{G}(E)$, we will know that it is a σ-algebra and that $\mathcal{B}_E \subseteq \mathcal{L}$ as soon I show that $\mathcal{L}$ is closed under complementation.

LEMMA 2.2.8. *If $K \subset\subset E$,*[1] *then $K \in \mathcal{L}$ and $\tilde{\mu}(K) < \infty$.*

PROOF: The first step is to show that $\tilde{\mu}(K) < \infty$. For this purpose, choose a non-overlapping cover $\{I_m : m \geq 1\} \subseteq \mathfrak{R}$ of K, and then use (4) to choose $\{I'_m : m \geq 1\} \subseteq \mathfrak{R}$ so that $I_m \subseteq \mathring{I}'_m$ and $V(I'_m) \leq V(I_m) + 1$ for each m. Now apply the Heine–Borel property to find an $n \in \mathbb{Z}^+$ such that $K \subseteq \bigcup_{m=1}^n \mathring{I}'_m$. Then $\tilde{\mu}(K) \leq n + \sum_{m=1}^n V(I_m) < \infty$.

We will now show that $K \in \mathcal{L}$. To this end, let $\epsilon > 0$ be given, and choose an open set $G \supseteq K$ so that $\tilde{\mu}(G) \leq \tilde{\mu}(K) + \epsilon$. Set $H = G \setminus K \in \mathfrak{G}(E)$, and choose a non-overlapping sequence $\{I_n : n \geq 1\} \subseteq \mathfrak{R}$ so that $H = \bigcup_1^\infty I_n$. Then, by (2.2.6), $\tilde{\mu}(H) = \sum_{m=1}^\infty V(I_m)$. In addition, for each $n \in \mathbb{Z}^+$, $K_n \equiv \bigcup_{m=1}^n I_m$ is compact and disjoint from K. Hence, by Lemmas 2.2.7 and 2.2.3,

$$\tilde{\mu}(K_n) + \tilde{\mu}(K) = \tilde{\mu}(K_n \cup K) \leq \tilde{\mu}(G),$$

and so, because $\tilde{\mu}(K) < \infty$, $\sum_{m=1}^n V(I_m) = \tilde{\mu}(K_n) \leq \epsilon$ for all n, and therefore $\tilde{\mu}(G \setminus K) = \tilde{\mu}(H) \leq \sum_{m=1}^\infty \tilde{\mu}(I_m) \leq \epsilon$. $\square$

LEMMA 2.2.9. *$\mathcal{L}$ is a σ-algebra over E that contains $\mathcal{B}_E$. Moreover, if $\Gamma \subseteq E$, then $\Gamma \in \mathcal{L}$ if and only if for every $\epsilon > 0$ there exist $F \in \mathfrak{F}(E)$ and $G \in \mathfrak{G}(E)$ such that $F \subseteq \Gamma \subseteq G$ and $\tilde{\mu}(G \setminus F) < \epsilon$. Alternatively, $\Gamma \in \mathcal{L}$ if there exist $C, D \in \mathcal{L}$ such that $C \subseteq \Gamma \subseteq D$ and $\tilde{\mu}(D \setminus C) = 0$, and if $\Gamma \in \mathcal{L}$, then there exist $C \in \mathfrak{F}_\sigma(E)$ and $D \in \mathfrak{G}_\delta(E)$ such that $C \subseteq \Gamma \subseteq D$ and $\tilde{\mu}(D \setminus C) = 0$.*

PROOF: Because of Lemma 2.2.7, proving that $\mathcal{L}$ is a σ-algebra comes down to showing that it is closed under complementation. For this purpose, begin by observing that $\mathfrak{F}_\sigma(E) \subseteq \mathcal{L}$. To check this, choose $\{I_m : m \geq 1\} \subseteq \mathfrak{R}$ such that $E = \bigcup_{m=1}^\infty I_m$, and set $K_n = \bigcup_{m=1}^n I_m$. Then K_n is compact for each n. Given $F \in \mathfrak{F}(E)$ and $n \in \mathbb{Z}^+$, set $F_n = F \cap K_n$. Clearly F_n is compact and is therefore an element of $\mathcal{L}$. Hence, since $F = \bigcup_{n=1}^\infty F_n$ and $\mathcal{L}$ is closed under countable unions, we see first that $\mathfrak{F}(E) \subseteq \mathcal{L}$ and then that $\mathfrak{F}_\sigma(E) \subseteq \mathcal{L}$.

Next, suppose that $\Gamma \in \mathcal{L}$, and choose $D \in \mathfrak{G}_\delta(E)$ for which $\Gamma \subseteq D$ and $\tilde{\mu}(D \setminus \Gamma) = 0$. Then $C \equiv D^\complement \in \mathfrak{F}_\sigma(E)$, $C \subseteq \Gamma^\complement$, and $\tilde{\mu}(\Gamma^\complement \setminus C) = \tilde{\mu}(D \setminus \Gamma) = 0$. Hence, $\Gamma^\complement \setminus C \in \mathcal{L}$, and therefore so is $\Gamma^\complement = C \cup (\Gamma^\complement \setminus C)$, which means that $\mathcal{L}$ is closed under complementation and is therefore a σ-algebra over E.

Knowing that $\mathcal{L}$ contains $\mathfrak{G}(E)$ and is a σ-algebra, we know that $\mathcal{B}_E \subseteq \mathcal{L}$. In addition, if $\Gamma \in \mathcal{L}$, then for each $\epsilon > 0$ we can find an open $G \supseteq \Gamma$ and a closed F with $F^\complement \supseteq \Gamma^\complement$ for which $\tilde{\mu}(G \setminus \Gamma) \vee \tilde{\mu}(F^\complement \setminus \Gamma^\complement) < \frac{\epsilon}{2}$, which means that $F \subseteq \Gamma \subseteq G$ and $\tilde{\mu}(G \setminus F) < \epsilon$.

Finally, given the preceding, it is clear that if $\Gamma \in \mathcal{L}$ then there exist $C \in \mathfrak{F}_\sigma(E)$ and $D \in \mathfrak{G}_\delta(E)$ such that $C \subseteq \Gamma \subseteq D$ and $\tilde{\mu}(D \setminus C) = 0$. Conversely, if there exist such $C, D \in \mathcal{L}$, $C \subseteq \Gamma \subseteq D$, and $\tilde{\mu}(D \setminus C) = 0$, then $\tilde{\mu}(\Gamma \setminus C) \leq \tilde{\mu}(D \setminus C) = 0$, and so $\Gamma = C \cup (\Gamma \setminus C) \in \mathcal{L}$. $\square$

---

[1] I will often use the notation $K \subset\subset E$ to mean that K is a compact subset of E. When E is discrete, the notation means that K is a finite subset of E.

Theorem 2.2.10. *Refer to the preceding. Then there is a unique Borel measure μ on E for which $\mu(I) = V(I)$ for all $I \in \mathfrak{R}$. Moreover, μ is regular, $\mathcal{L} = \overline{\mathcal{B}_E}^{\mu}$, and (cf. Lemma 2.1.14) $\bar{\mu}$ is the restriction of $\tilde{\mu}$ to $\overline{\mathcal{B}_E}^{\mu}$.*

Proof: We will first show that $\tilde{\mu}$ is countably additive on $\mathcal{L}$. To this end, let $\{\Gamma_n : n \geq 1\} \subseteq \mathcal{L}$ be a sequence of mutually disjoint, relatively compact (i.e., their closures are compact) sets. By Lemma 2.2.9, for each $\epsilon > 0$ we can find a sequence $\{K_n : n \geq 1\}$ of compact sets such that $K_n \subseteq \Gamma_n$ and $\tilde{\mu}(\Gamma_n) \leq \tilde{\mu}(K_n) + 2^{-n}\epsilon$ for each n. Hence, by Lemma 2.2.7, for each $n \in \mathbb{Z}^+$,

$$\tilde{\mu}\left(\bigcup_{m=1}^{\infty} \Gamma_m\right) \geq \tilde{\mu}\left(\bigcup_{m=1}^{n} K_m\right) = \sum_{m=1}^{n} \tilde{\mu}(K_m),$$

and therefore

$$\tilde{\mu}\left(\bigcup_{m=1}^{\infty} \Gamma_m\right) \geq \sum_{m=1}^{\infty} \tilde{\mu}(K_m) \geq \sum_{m=1}^{\infty} \tilde{\mu}(\Gamma_m) - \epsilon,$$

which proves that

$$\tilde{\mu}\left(\bigcup_{m=1}^{\infty} \Gamma_m\right) \geq \sum_{m=1}^{\infty} \tilde{\mu}(\Gamma_m).$$

Since the opposite inequality is trivial, this proves the countable additivity of $\tilde{\mu}$ for relatively compact elements of $\mathcal{L}$.

To treat the general case, choose $\{I_m : m \geq 1\} \subseteq \mathfrak{R}$ for which that $E = \bigcup_{m=1}^{\infty} I_m$, and set $A_1 = I_1$ and $A_{n+1} = I_{n+1} \setminus \bigcup_{m=1}^{n} I_m$. Then the A_n's are mutually disjoint, relatively compact elements of $\mathcal{L}$. Hence, if $\Gamma_{m,n} = A_m \cap \Gamma_n$, then the $\Gamma_{m,n}$'s are also mutually disjoint and relatively compact. Furthermore, $\Gamma_n = \bigcup_{m=1}^{\infty} \Gamma_{m,n}$ for each n, and so, by the preceding and Lemma 2.2.1,

$$\tilde{\mu}\left(\bigcup_{n=1}^{\infty} \Gamma_n\right) = \sum_{(m,n)\in(\mathbb{Z}^+)^2} \tilde{\mu}(\Gamma_{m,n}) = \sum_{n=1}^{\infty} \tilde{\mu}(\Gamma_n).$$

Knowing that $\tilde{\mu}$ is countably additive on $\mathcal{L}$ and that $\mathcal{L} \supseteq \mathcal{B}_E$, we can take μ to be the restriction of $\tilde{\mu}$ to $\mathcal{B}_E$. Furthermore, the results in Lemma 2.2.9 show that this μ is regular, $\mathcal{L} = \overline{\mathcal{B}_E}^{\mu}$, and that $\tilde{\mu} \restriction \mathcal{L}$ equals $\bar{\mu}$.

To complete the proof, suppose that ν is a second Borel measure on E for which $\nu(I) = V(I)$ whenever $I \in \mathfrak{R}$. Given an open G, choose a non-overlapping $\{I_m : m \geq 1\} \subseteq \mathfrak{R}$ whose union is G, and apply (2.1.8) and (2.2.6) to conclude that

$$\nu(G) \leq \sum_{m=1}^{\infty} \nu(I_m) = \sum_{m=1}^{\infty} V(I_m) = \mu(G).$$

Next, given $\epsilon > 0$, choose for $m \in \mathbb{Z}^+$ an $I''_m \in \mathfrak{R}$ so that $I''_m \subseteq \mathring{I}_m$ and $V(I_m) \le V(I''_m) + 2^{-m}\epsilon$. Then, because the I''_m's are disjoint,

$$\nu(G) \ge \nu\left(\bigcup_{m=1}^{\infty} I''_m\right) = \sum_{m=1}^{\infty} V(I''_m) \ge \sum_{m=1}^{\infty} V(I_m) - \epsilon \ge \mu(G) - \epsilon.$$

Hence, ν and μ are equal on $\mathfrak{G}(E)$. Finally, note that, by combining (4) and (5), we can produce a non-decreasing sequence of open sets $G_n \nearrow E$ such that $\mu(G_n) < \infty$. Hence, by Theorem 2.1.13, ν equals μ on $\mathcal{B}_{G_n}$ for each n, from which it follows easily that ν equals μ on $\mathcal{B}_E$. $\square$

Corollary 2.2.11. *Suppose that $T : E \longrightarrow E$ is a transformation with the property that $T^{-1}(I) \in \mathfrak{R}$ and $V\big(T^{-1}(I)\big) = V(I)$ for all $I \in \mathfrak{R}$. Then $T^{-1}(\Gamma) \in \mathcal{B}_E$ and $\mu\big(T^{-1}(\Gamma)\big) = \mu(\Gamma)$ for all $\Gamma \in \mathcal{B}_E$. Equivalently, $T_*\mu = \mu$.*

Proof: By part (**i**) of Exercise 2.1.19, T^{-1} of a union of sets is the union of T^{-1} of each set over which the union is taken, and T^{-1} of a difference of sets is the difference of T^{-1} of each set. Next, let $\mathcal{B}$ be the set of $\Gamma \in \mathcal{B}_E$ with the property that $T^{-1}(\Gamma) \in \mathcal{B}_E$. By the preceding observation, $\mathcal{B}$ is a σ-algebra over E. In addition, $\mathfrak{R} \subseteq \mathcal{B}$. Thus, because for any open G there is a sequence $\{I_m : m \ge 1\} \subseteq \mathfrak{R}$ whose union is G, $\mathfrak{G}(E) \subseteq \mathcal{B}$. Since this means that $\mathcal{B}$ is a σ-algebra contained in $\mathcal{B}_E$ and containing $\mathfrak{G}(E)$, it follows that $\mathcal{B} = \mathcal{B}_E$.

Next, set $\nu(\Gamma) = \mu\big(T^{-1}(\Gamma)\big)$ for $\Gamma \in \mathcal{B}_E$. By part (**iii**) of Exercise 2.1.19, ν is a Borel measure on E. Moreover, by assumption, $\nu(I) = V(I)$ for $I \in \mathfrak{R}$. Hence, by the uniqueness statement in Theorem 2.2.10, $\nu = \mu$. $\square$

§ **2.2.2. Lebesgue Measure on $\mathbb{R}^N$:** My first application of Theorem 2.2.10 will be to the construction of the father of all measures, Lebesgue measure on $\mathbb{R}^N$.

Endow $\mathbb{R}^N$ with the standard Euclidean metric, the one given by the Euclidean distance between points. Next, take $\mathfrak{R}$ to be the set of all rectangles I in $\mathbb{R}^N$ relative to a fixed orthonormal coordinate system, include the empty rectangle in $\mathfrak{R}$, and define $V(I) = \mathrm{vol}(I)$ if $I \ne \emptyset$ and $V(\emptyset) = 0$. In order to apply the results in § 2.2.1, I have to show that this choice of $\mathfrak{R}$ and V satisfies the hypotheses (1)–(5) listed at the beginning of that subsection. It is clear that they satisfy (1), (2), and (4). In addition, (3) follows from Lemma 1.1.1. To check (5), I will use the following lemma. In its statement and elsewhere, a **square** will be a (multi-dimensional) rectangle all of whose edges have the same length. That is, a non-empty square is a set Q of the form $x + [a, b]^N$ for some $x \in \mathbb{R}^N$ and $a \le b$.

Lemma 2.2.12. *If G is an open set in $\mathbb{R}$, then G is the union of a countable number of mutually disjoint open intervals. More generally, if G is an open set in $\mathbb{R}^N$, then, for each $\delta > 0$, G admits a countable, non-overlapping, exact cover $\mathcal{C}$ by closed squares Q with* $\mathrm{diam}\,(Q) < \delta$.

Proof: If $G \subseteq \mathbb{R}$ is open and $x \in G$, let $\mathring{I}_x$ be the open connected component of G containing x. Then $\mathring{I}_x$ is an open interval and, for any $x, y \in G$, either

$\mathring{I}_x \cap \mathring{I}_y = \emptyset$ or $\mathring{I}_x = \mathring{I}_y$. Hence, $\mathcal{C} \equiv \{\mathring{I}_x : x \in G \cap \mathbb{Q}\}$ ($\mathbb{Q}$ here denotes the set of rational numbers) is the required cover.

To handle the second assertion, set $Q_n = \left[0, 2^{-n}\right]^N$ and $\mathcal{K}_n = \{\frac{\mathbf{k}}{2^n} + Q_n : \mathbf{k} \in \mathbb{Z}^N\}$. Note that if $m \le n$, $Q \in \mathcal{K}_m$, and $Q' \in \mathcal{K}_n$, then either $Q' \subseteq Q$ or $\mathring{Q} \cap \mathring{Q}' = \emptyset$. Now let $G \in \mathfrak{G}(\mathbb{R}^N)$ and $\delta > 0$ be given, take n_0 to be the smallest $n \in \mathbb{Z}$ for which $2^{-n}\sqrt{N} < \delta$, and set $\mathcal{C}_{n_0} = \{Q \in \mathcal{K}_{n_0} : Q \subseteq G\}$. Next, define $\mathcal{C}_n$ inductively for $n > n_0$ so that

$$\mathcal{C}_{n+1} = \left\{Q' \in \mathcal{K}_{n+1} : Q' \subseteq G \text{ and } \mathring{Q}' \cap \mathring{Q} = \emptyset \text{ for any } Q \in \bigcup_{m=n_0}^{n} \mathcal{C}_m\right\}.$$

Since if $m \le n$, $Q \in \mathcal{C}_m$, and $Q' \in \mathcal{C}_n$, either $Q = Q'$ or $\mathring{Q} \cap \mathring{Q}' = \emptyset$, the squares in $\mathcal{C} \equiv \bigcup_{n=n_0}^{\infty} \mathcal{C}_n$ are non-overlapping, and certainly $\bigcup \mathcal{C} \subseteq G$. Finally, if $x \in G$, choose $n \ge n_0$ and $Q' \in \mathcal{K}_n$ such that $x \in Q' \subseteq G$. If $Q' \notin \mathcal{C}_n$, then there exist an $n_0 \le m < n$ and a $Q \in \mathcal{C}_m$ for which $\mathring{Q} \cap \mathring{Q}' \neq \emptyset$. But this means that $Q' \subseteq Q$ and therefore that $x \in Q \subseteq \bigcup \mathcal{C}$. Thus $\mathcal{C}$ covers G. $\Box$

Knowing that $\mathfrak{R}$ and V satisfy hypotheses (1)–(5), we can apply Theorem 2.2.10 and thereby derive the following fundamental result.

THEOREM 2.2.13. *There is one and only one Borel measure $\lambda_{\mathbb{R}^N}$ on $\mathbb{R}^N$ with the property that $\lambda_{\mathbb{R}^N}(Q) = \mathrm{vol}(Q)$ for all squares Q in $\mathbb{R}^N$. Moreover, $\lambda_{\mathbb{R}^N}$ is regular.*

PROOF: The existence of $\lambda_{\mathbb{R}^N}$ as well as its regularity are immediate consequences of Theorem 2.2.10. Furthermore, that theorem says that $\lambda_{\mathbb{R}^N}$ is the only Borel measure ν with the property that $\nu(I) = \mathrm{vol}(I)$ for all $I \in \mathfrak{R}$. Thus, to prove the uniqueness statement here, it suffices to check that $\nu(I) = \mathrm{vol}(I)$ for all rectangles I if it does for squares. To this end, first note that if I is a rectangle, then there exists a sequence $\{I_n : n \ge 1\}$ of rectangles such that $I_n \nearrow \mathring{I}$ and $\mathrm{vol}(I_n) \nearrow \mathrm{vol}(I)$. Hence, by (2.1.10), $\nu(I) = \nu(\mathring{I})$. In particular, by (2.1.7), this means that[2] $\nu(\partial I) = \nu(I) - \nu(\mathring{I}) = 0$ for all rectangles I. Now apply Lemma 2.2.12 to write $\mathring{I} = \bigcup_{n=1}^{\infty} Q_n$, where the Q_n's are non-overlapping squares, use the preceding to check that $\nu(Q_m \cap Q_n) = 0$ for $m \neq n$, and apply (2.1.9) to see that $\nu(I) = \nu(\mathring{I}) = \sum_{n=1}^{\infty} \mathrm{vol}(Q_n)$. Since the same reasoning applies to $\lambda_{\mathbb{R}^N}$ and $\lambda_{\mathbb{R}^N}(Q_n) = \mathrm{vol}(Q_n)$, we have now shown that $\nu(I) = \mathrm{vol}(I)$ as well. $\Box$

The Borel measure $\lambda_{\mathbb{R}^N}$ described in Theorem 2.2.13 is called **Lebesgue measure** on $\mathbb{R}^N$. In addition, elements of $\overline{\mathcal{B}_{\mathbb{R}^N}}^{\lambda_{\mathbb{R}^N}}$ are said to be **Lebesgue measurable** sets.

An important property of Lebesgue measure is that it is translation invariant. To be precise, for each $x \in \mathbb{R}^N$, define the **translation map**

[2] I use $\partial\Gamma$ to denote the boundary $\bar{\Gamma} \setminus \mathring{\Gamma}$ of a set Γ.

$T_x : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ by $T_x(y) = x + y$. Obviously, T_x is one-to-one and onto. In fact, $T_x^{-1} = T_{-x}$. In addition, because $T_x = T_{-x}^{-1}$ and T_{-x} is continuous, T_x takes $\mathcal{B}_{\mathbb{R}^N}$ into itself. Finally, say that a Borel measure μ on $\mathbb{R}^N$ is **translation invariant** if $\mu\big(T_x(\Gamma)\big) = \nu(\Gamma)$ for all $x \in \mathbb{R}^N$ and $\Gamma \in \mathcal{B}_{\mathbb{R}^N}$. The following corollary provides an important characterization of Lebesgue measure in terms of translation invariance.

Corollary 2.2.14. *Lebesgue measure is the one and only translation invariant Borel measure on $\mathbb{R}^N$ that assigns the unit square $[0,1]^N$ measure 1. Thus, if ν is a translation invariant Borel measure on $\mathbb{R}^N$ and $\alpha = \nu\big([0,1]^N\big) < \infty$, then $\nu = \alpha\lambda_{\mathbb{R}^N}$.*

Proof: That $\lambda_{\mathbb{R}^N}$ is translation invariant follows immediately from Corollary 2.2.11 and the fact that, for any rectangle I and $x \in \mathbb{R}^N$, $\mathrm{vol}\big(T_x(I)\big) = \mathrm{vol}(I)$.

To prove the uniqueness assertion, suppose that μ is a translation invariant Borel measure that gives measure 1 to $[0,1]^N$. We first show that $\mu(H) = 0$ for every H of the form $\{x \in \mathbb{R}^N : x_i = c\}$ for some $1 \le i \le N$ and $c \in \mathbb{R}$. Indeed, by translation invariance, it suffices to treat the case $c = 0$. In addition, by countable subadditivity and translation invariance, it is sufficient to show that $\mu(R) = 0$ when $R = \{x : x_i = 0 \text{ and } x_j \in [0,1] \text{ for } j \ne i\}$. But if $\mathbf{e}_i$ is the unit vector whose ith coordinate is 1 and whose other coordinates are 0, then, for any $n \ge 1$, the sets $\big\{\frac{m}{n}\mathbf{e}_i + R : 0 \le m \le n\big\}$ are mutually disjoint, have the same μ-measure as R, and are contained in $[0,1]^N$. Hence, $n\mu(R) \le 1$ for all $n \ge 1$, and so $\mu(R) = 0$.

Given the preceding, we know that $\mu(\partial I) = 0$ for all rectangles $I \subseteq \mathbb{R}^N$. Hence, if $(n_1, \dots, n_N) \in (\mathbb{Z}^+)^N$, then

$$\begin{aligned}\mu\big([0,1]^N\big) &= \mu\left(\bigcup\left\{\prod_{i=1}^N \left[\tfrac{k_i-1}{n_i}, \tfrac{k_i}{n_i}\right] : 1 \le k_i \le n_i \text{ for } 1 \le i \le N\right\}\right)\\ &= \left(\prod_{i=1}^N n_i\right)\mu\left(\prod_{i=1}^N \left[0, \tfrac{1}{n_i}\right]\right),\end{aligned}$$

and so $\mu\left(\prod_{i=1}^N \left[0, \frac{1}{n_i}\right]\right) = \prod_{i=1}^N \frac{1}{n_i}$. Starting from this, the same line of reasoning can be used to show that $\mu\left(\prod_{i=1}^N \left[0, \frac{m_i}{n_i}\right]\right) = \prod_{i=1}^N \frac{m_i}{n_i}$ for any pair $(m_1, \dots, m_N)$, $(n_1, \dots, n_N) \in (\mathbb{Z}^+)^N$. Hence, by translation invariance, for any rectangle whose sides have rational lengths, $\mu(I) = \mathrm{vol}(I)$. Finally, for any rectangle I, we can choose a non-increasing sequence $\{I_n : n \ge 1\}$ of rectangles with rational side lengths such that $I_n \searrow I$, and so $\mu(I) = \lim_{n\to\infty} \mathrm{vol}(I_n) = \mathrm{vol}(I)$. Now apply Corollary 2.2.11.

To prove the concluding assertion, first suppose that $\alpha = 0$. Then, because $\mathbb{R}^N$ can be covered by a countable number of translates of $[0,1]^N$, it follows that $\nu(\mathbb{R}^N) = 0$ and therefore that $\nu = \alpha\lambda_{\mathbb{R}^N}$. Next suppose that $\alpha > 0$. Then $\alpha^{-1}\nu$ is a translation invariant Borel measure on $\mathbb{R}^N$ and ν assigns the unit square measure 1. Hence, by the earlier part, $\alpha^{-1}\nu = \lambda_{\mathbb{R}^N}$. $\square$

Although the property of translation invariance was built into the construction of Lebesgue measure, it is not immediately obvious how Lebesgue measure responds to rotations of $\mathbb{R}^N$. One suspects that, as *the natural* measure on $\mathbb{R}^N$, Lebesgue measure should be invariant under the full group of Euclidean transformations (i.e., rotations as well as translations). However, because my description of Lebesgue's measure was based on rectangles and the rectangles were inextricably tied to a fixed set of coordinate axes, rotation invariance is not as obvious as translation invariance was.

The following corollary shows how Lebesgue measure transforms under an arbitrary linear transformation of $\mathbb{R}^N$, and rotation invariance will follow as an immediate corollary.

Given an $N \times N$, real matrix $A = \big((a_{ij})\big)_{1 \le i,j \le N}$, define $T_A : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ to be the action of A on x. That is, $(T_A x)_i = \sum_{j=1}^N a_{ij} x_j$ for $1 \le i \le N$. We can now prove the following.

THEOREM 2.2.15. *For any $N \times N$, real matrix A and $\Gamma \in \overline{\mathcal{B}_{\mathbb{R}^N}}^{\lambda_{\mathbb{R}^N}}$, $T_A(\Gamma) \in \overline{\mathcal{B}_{\mathbb{R}^N}}^{\lambda_{\mathbb{R}^N}}$ and $\overline{\lambda_{\mathbb{R}^N}}(T_A \Gamma) = |\det(A)| \overline{\lambda_{\mathbb{R}^N}}(\Gamma)$. Moreover, if A is non-singular, then T_A takes $\mathcal{B}_{\mathbb{R}^N}$ into itself.*

PROOF: We begin with the case in which A is non-singular. Then $T_{A^{-1}}$ is a continuous, one-to-one map from $\mathbb{R}^N$ onto itself, and $T_A = (T_{A^{-1}})^{-1}$. Hence, by (**iii**) of Exercise 2.1.19, T_A takes $\mathcal{B}_{\mathbb{R}^N}$ into itself. Next, define ν_A on $\mathcal{B}_{\mathbb{R}^N}$ by $\nu_A(\Gamma) = \lambda_{\mathbb{R}^N}\big(T_A(\Gamma)\big)$. Then, again, by part (**iii**) of Exercise 2.1.19, ν_A is a Borel measure on $\mathbb{R}^N$. Now set $\alpha(A) = \nu_A\big([0,1]^N\big)$. Because $T_A([0,1]^N)$ is compact, $\alpha(A) < \infty$. In addition, because

$$\nu_A\big(T_x(\Gamma)\big) = \lambda_{\mathbb{R}^N}\big(T_{Ax} + T_A(\Gamma)\big) = \lambda_{\mathbb{R}^N}\big(T_A(\Gamma)\big) = \nu_A(\Gamma),$$

ν_A is translation invariant. Thus Corollary 2.2.14 says that $\nu_A = \alpha(A)\lambda_{\mathbb{R}^N}$, and so all that we have to do is show that $\alpha(A) = |\det(A)|$. To this end, observe that there are cases in which $\alpha(A)$ can be computed by hand. The first of these is when A is diagonal with positive diagonal elements, in which case $T_A\big([0,1]\big) = \prod_{j=1}^N [0, a_{jj}]$ and therefore $\alpha(A) = \prod_{j=1}^N a_{jj} = \det(A)$. The second case is the one in which A is an orthogonal matrix. Then $T_A\big(\overline{B(0,1)}\big) = \overline{B(0,1)}$ and therefore, since $\lambda_{\mathbb{R}^N}\big(\overline{B(0,1)}\big) \in (0,\infty)$, $\alpha(A) = 1$. To go further, notice that, since $T_{AA'} = T_A \circ T_{A'}$, $\alpha(AA') = \alpha(A)\alpha(A')$. Hence, if A is symmetric and positive definite (i.e., all its eigenvalues are positive) and $\mathcal{O}$ is an orthogonal matrix for which[3] $D = \mathcal{O}^\top A \mathcal{O}$ is diagonal, then the diagonal entries of D are positive, $\det(A) = \det(D) = \alpha(D)$, and therefore $\alpha(A) = \alpha(\mathcal{O}^\top)\alpha(D)\alpha(\mathcal{O}) = \alpha(D) = \det(A)$. Finally, for any non-singular A, $A^\top A$ is a symmetric, positive definite matrix. Moreover, if $\mathcal{O} = A^{-1}(AA^\top)^{\frac{1}{2}}$, where $(AA^\top)^{\frac{1}{2}}$ denotes the symmetric square root of $AA^\top$, then $\mathcal{O}$ satisfies $\mathcal{O}\mathcal{O}^\top = A^{-1}AA^\top(A^\top)^{-1} = I$ and is therefore orthogonal. Hence, since $A = (AA^\top)^{\frac{1}{2}}\mathcal{O}^\top$, we find that $\alpha(A) = \det\big((AA^\top)^{\frac{1}{2}}\big) =$

[3] Given a matrix A, I use $A^\top$ to denote the transpose matrix.

$|\det(\mathrm{A})|$. Finally, to show that T_A takes a $\Gamma \in \overline{\mathcal{B}_{\mathbb{R}^N}}^{\lambda_{\mathbb{R}^N}}$ into $\overline{\mathcal{B}_{\mathbb{R}^N}}^{\lambda_{\mathbb{R}^N}}$ and that $\overline{\lambda_{\mathbb{R}^N}}\big(T_A(\Gamma)\big) = |\det(A)|\overline{\lambda_{\mathbb{R}^N}}(\Gamma)$, choose $C, D \in \mathcal{B}_{\mathbb{R}^N}$ so that $C \subseteq \Gamma \subseteq D$ and $\lambda_{\mathbb{R}^N}(D \setminus C) = 0$. Then $T_A(C), T_A(D) \in \mathcal{B}_{\mathbb{R}^N}$, $T_A(C) \subseteq T_A(\Gamma) \subseteq T_A(D)$, $\lambda_{\mathbb{R}^N}\big(T_A(D) \setminus T_A(C)\big) = \lambda_{\mathbb{R}^N}\big(T_A(D \setminus C)\big) = 0$, and therefore $T_A(\Gamma) \in \overline{\mathcal{B}_{\mathbb{R}^N}}^{\lambda_{\mathbb{R}^N}}$ and $\overline{\lambda_{\mathbb{R}^N}}\big(T_A(\Gamma)\big) = \lambda_{\mathbb{R}^N}\big(T_A(C)\big) = |\det(A)|\lambda_{\mathbb{R}^N}(C) = |\det(A)|\overline{\lambda_{\mathbb{R}^N}}(\Gamma)$.

To treat the singular case, first observe that there is nothing to do when $N = 1$, since the singularity of A means that $T_A(\mathbb{R}) = \{0\}$ and $\lambda_{\mathbb{R}}(\{0\}) = 0$. Thus, assume that $N \geq 2$. Then, if A is singular, $T_A(\mathbb{R}^N)$ is contained in an $(N-1)$-dimensional subspace of $\mathbb{R}^N$. Therefore, what remains is to show that $\lambda_{\mathbb{R}^N}$ assigns measure 0 to an $(N-1)$-dimensional subspace H. This is clear if $H = \mathbb{R}^{N-1} \times \{0\}$, since in that case one can obviously cover H with a countable number of rectangles each of which has volume 0. To handle general H's, choose an orthogonal matrix $\mathcal{O}$ so that $H = T_{\mathcal{O}}\big(\mathbb{R}^{N-1} \times \{0\}\big)$, and use the preceding to conclude that $\lambda_{\mathbb{R}^N}(H) = \lambda_{\mathbb{R}^N}\big(\mathbb{R}^{N-1} \times \{0\}\big) = 0$. $\square$

Before concluding this preliminary discussion of Lebesgue measure, it may be appropriate to examine whether there are any sets that are not Lebesgue measurable. It turns out that the existence of such sets brings up some extremely delicate issues about the foundations of mathematics. Indeed, if one is willing to abandon the full axiom of choice, then R. Solovay has shown that there is a model of mathematics in which *every* subset of $\mathbb{R}^N$ is Lebesgue measurable. However, if one accepts the full axiom of choice, then the following argument, due to Vitali, shows that there are sets that are not Lebesgue measurable. The use of the axiom of choice comes in Lemma 2.2.17 below. It is not used in the proof of the next lemma, a result that is interesting in its own right. See Exercise 2.2.35 for a somewhat surprising application and Exercise 6.3.17 for another derivation of it.

Lemma 2.2.16. *If $\Gamma \in \overline{\mathcal{B}_{\mathbb{R}}}^{\lambda_{\mathbb{R}}}$ has positive Lebesgue measure, then the set $\Gamma - \Gamma \equiv \{y - x : x, y \in \Gamma\}$ contains an open interval $(-\delta, \delta)$ for some $\delta > 0$.*

Proof: Without loss of generality, we will assume that $\Gamma \in \mathcal{B}_{\mathbb{R}}$ and that $\lambda_{\mathbb{R}}(\Gamma) \in (0, \infty)$.

Choose an open set $G \supseteq \Gamma$ for which $\lambda_{\mathbb{R}}(G \setminus \Gamma) < \frac{1}{3}\lambda_{\mathbb{R}}(\Gamma)$, and let (cf. the first part of Lemma 2.2.12) $\mathcal{C}$ be a countable collection of mutually disjoint, non-empty, open intervals $\mathring{I}$ whose union is G. Then

$$\sum_{\mathring{I} \in \mathcal{C}} \lambda_{\mathbb{R}}(\mathring{I} \cap \Gamma) = \lambda_{\mathbb{R}}(\Gamma) \geq \tfrac{3}{4}\lambda_{\mathbb{R}}(G) = \tfrac{3}{4}\sum_{\mathring{I} \in \mathcal{C}} \lambda_{\mathbb{R}}(\mathring{I}).$$

Hence, there must be an $\mathring{I} \in \mathcal{C}$ for which $\lambda_{\mathbb{R}}(\mathring{I} \cap \Gamma) \geq \frac{3}{4}\lambda_{\mathbb{R}}(\mathring{I})$. Set $A = \mathring{I} \cap \Gamma$. If $d \in \mathbb{R}$ and $(d + A) \cap A = \emptyset$, then

$$2\lambda_{\mathbb{R}}(A) = \lambda_{\mathbb{R}}(d + A) + \lambda_{\mathbb{R}}(A) = \lambda_{\mathbb{R}}\big((d + A) \cup A\big) \leq \lambda_{\mathbb{R}}\big((d + \mathring{I}) \cup \mathring{I}\big).$$

At the same time, $(d+\mathring{I}) \cup \mathring{I} \subseteq (I^-, d+I^+)$ if $d \geq 0$ and $(d+\mathring{I}) \cup \mathring{I} \subseteq (d+I^-, I^+)$ if $d < 0$, where I^- and I^+ denote the left and right endpoints of $\mathring{I}$. Thus, in either case, $\lambda_{\mathbb{R}}\big((d+\mathring{I})\cup\mathring{I}\big)\le |d|+\lambda_{\mathbb{R}}(\mathring{I})$. Hence, if $(d+A)\cap A=\emptyset$, then $\frac{3}{2}\lambda_{\mathbb{R}}(\mathring{I})\le 2\lambda_{\mathbb{R}}(A)\le |d|+\lambda_{\mathbb{R}}(\mathring{I})$, from which one sees that $|d|\ge \frac{1}{2}\lambda_{\mathbb{R}}(\mathring{I})$. In other words, if $|d|<\frac{1}{2}\lambda_{\mathbb{R}}(\mathring{I})$, then $(d+A)\cap A\ne\emptyset$. But this means that for every $d\in\big(-\frac{1}{2}\lambda_{\mathbb{R}}(\mathring{I}),\frac{1}{2}\lambda_{\mathbb{R}}(\mathring{I})\big)$ there exist $x,y\in A\subseteq\Gamma$ for which $d=y-x$. $\square$

LEMMA 2.2.17. *Let $\mathbb{Q}$ denote the set of rational real numbers. Assuming the axiom of choice, there is a subset A of $\mathbb{R}$ such that $(A-A)\cap\mathbb{Q}=\{0\}$ and yet $\mathbb{R}=\bigcup_{q\in\mathbb{Q}}(q+A)$.*

PROOF: Write $x\sim y$ if $y-x\in\mathbb{Q}$. Then "$\sim$" is an equivalence relation on $\mathbb{R}$, and, for each $x\in\mathbb{R}$, the equivalence class $[x]^\sim$ of x is $x+\mathbb{Q}$. Now, using the axiom of choice, choose A to be a set that contains precisely one element from each of the equivalence classes $[x]^\sim$, $x\in\mathbb{R}$. It is then clear that A has the required properties. $\square$

THEOREM 2.2.18. *Assuming the axiom of choice, every $\Gamma\in\overline{\mathcal{B}_{\mathbb{R}}}^{\lambda_{\mathbb{R}}}$ with positive Lebesgue measure contains a subset that is not Lebesgue measurable. (See part* (**iii**) *of Exercise 2.2.36 for another construction of non-measurable quantities.)*

PROOF: Let A be the set constructed in Lemma 2.2.17, and suppose that $\Gamma\cap T_q(A)$ were Lebesgue measurable for each $q\in\mathbb{Q}$. Then we would have that $0<\overline{\lambda_{\mathbb{R}}}(\Gamma)\le\sum_{q\in\mathbb{Q}}\overline{\lambda_{\mathbb{R}}}\big(\Gamma\cap T_q(A)\big)$, and so there would exist a $q\in\mathbb{Q}$ such that $\overline{\lambda_{\mathbb{R}}}\big(\Gamma\cap T_q(A)\big)>0$. But, by Lemma 2.2.16, we would then have that $(-\delta,\delta)\subseteq\{y-x:x,y\in T_q(A)\}\subseteq\{0\}\cup\mathbb{Q}^{\complement}$ for some $\delta>0$, which cannot be. $\square$

§ **2.2.3. Distribution Functions and Measures:** Given a finite Borel measure on $\mathbb{R}$, set $F_\mu(x)=\mu\big((-\infty,x]\big)$. Clearly F_μ is a non-negative, bounded, right-continuous, non-decreasing function that tends to 0 as $x\to-\infty$. The function F_μ is called the **distribution function** for μ. In this subsection I will show that every bounded, right-continuous, non-decreasing function F that tends to 0 at $-\infty$ is the distribution of a unique finite Borel measure on $\mathbb{R}$.

Let F be given. By Exercise 1.2.23, $F=F_{\mathrm{c}}+F_{\mathrm{d}}$, where F_{c} and F_{d} are bounded and non-decreasing, F_{c} is continuous, and F_{d} is a pure jump function. Further, it is easy to check that both F_{c} and F_{d} can be taken so that they tend to 0 at $-\infty$. Hence, to prove the existence of a μ for which $F=F_\mu$, it suffices to do so when F is either a continuous or a pure jump function and then take the sum of the measures corresponding to F_{c} and F_{d}.

When F is a pure jump function, there is nearly nothing to do. Simply define μ by

$$\mu_F(\Gamma)=\sum_{\{x\in\Gamma\cap D\}}\big(F(x)-F(x-)\big)\quad\text{for }\Gamma\in\mathcal{B}_{\mathbb{R}},$$

where D is the countable set consisting of the discontinuities of F. Without difficulty, one can check that μ_F is a finite Borel measure on $\mathbb{R}$ for which F is the distribution function.

Now assume that F is continuous, and take $\mathfrak{R}$ to be the set of all (including the empty interval) closed intervals I in $\mathbb{R}$, and define $V\big([a,b]\big) = F(b) - F(a)$ for $a \leq b$. Checking that this choice of $\mathfrak{R}$ and V satisfies that hypotheses at the start of § 2.2.1 is easy. The same argument as we used to prove Lemma 1.1.1 when $N = 1$ shows that (3) holds, (4) follows from the continuity of F, and Lemma 2.2.12 proves (5). Thus, by Theorem 2.2.10, there is a Borel measure μ_F on $\mathbb{R}$ for which $\mu_F\big([a,b]\big) = F(b) - F(a)$ for all $a < b$. In particular,

$$F(x) = F(x) - \lim_{y \searrow -\infty} F(y) = \lim_{y \to -\infty} \mu_F\big([y,x]\big) = \mu_F\big((-\infty,x]\big),$$

and so F is the distribution function for μ.

Theorem 2.2.19. *Let F be a bounded, right-continuous, non-decreasing function on $\mathbb{R}$ that tends to 0 at $-\infty$. Then there is a unique Borel measure μ_F on $\mathbb{R}$ for which F is the distribution function. In particular, μ_F is finite and regular. (See Exercises 2.2.37 and 8.2.18 for other approaches.)*

Proof: The only assertions that have not been covered already are those of uniqueness and finiteness. However, the finiteness follows from $\mu_F(\mathbb{R}) = \lim_{x \nearrow \infty} F(x) < \infty$. To prove the uniqueness, suppose that ν is a second Borel measure on $\mathbb{R}$ satisfying $\nu\big((-\infty,x]\big) = F(x)$ for all $x \in \mathbb{R}$. Then, by the argument just given, ν is finite. In addition, for $a < b$,

$$\nu\big((a,b)\big) = \lim_{x \nearrow b}\big(F(x) - F(a)\big) = F(b-) - F(a) = \mu_F\big((a,b)\big).$$

Hence, $\nu(\mathring{I}) = \mu_F(\mathring{I})$ for all open intervals $\mathring{I}$, and so, by the first part of Lemma 2.2.12, $\nu(G) = \mu_F(G)$ for all $G \in \mathfrak{G}(\mathbb{R})$. By Theorem 2.1.13, this means that $\nu = \mu_F$ on $\mathcal{B}_{\mathbb{R}}$. $\square$

§ **2.2.4. Bernoulli Measure:** Here is an application of the material in § 2.2.1 to a probabilistic model of coin tossing.

Set $\Omega = \{0,1\}^{\mathbb{Z}^+}$, the space of maps $\omega : \mathbb{Z}^+ \longrightarrow \{0,1\}$. In the model, Ω is thought of as the set of all possible outcomes of a countably infinite number of coin tosses: $\omega(i) = 1$ if the ith toss came up heads and $\omega(i) = 0$ if it came up tails. Similarly, given $\emptyset \neq S \subseteq \mathbb{Z}^+$, take $\Omega(S) = \{0,1\}^S$, think of $\Omega(S)$ as the outcomes of those tosses that occurred during S, and define $\Pi_S\Omega \longrightarrow \Omega(S)$ to be the projection map given by $\Pi_S\omega = \omega \restriction S$. Then, for each S, $\mathcal{A}(S) \equiv \{\Pi_S^{-1}\Gamma : \Gamma \subseteq \Omega(S)\}$ is a σ-algebra over Ω, and, in the model, elements of $\mathcal{A}(S)$ are events (the probabilistic term for subsets) that depend only on the outcome of tosses corresponding to the i's in S.

Now suppose that, on each toss, the coin comes up heads with probability $p \in (0,1)$ and tails with probability $q = 1 - p$. Further, assume that the outcomes of distinct tosses are independent of one another. That is, if $\eta \in \Omega(S)$, then the probability of the event $\Pi_S^{-1}(\{\eta\}) = \{\omega \in \Omega : \omega(i) = \eta(i) \text{ for } i \in S\}$ is

$$p^{\sum_{i\in S}\eta(i)}q^{\sum_{i\in S}(1-\eta(i))}. \tag{2.2.20}$$

Obviously, when S is infinite, this quantity is 0. On the other hand, if $\emptyset \neq F \subset\subset \mathbb{Z}^+$ (i.e., F is a non-empty, finite subset of $\mathbb{Z}^+$), and $\beta_p^F : \mathcal{A}(F) \longrightarrow [0,1]$ is defined by the prescription

$$\beta_p^F\left(\Pi_F^{-1}\Gamma\right) = \sum_{\eta\in\Gamma} p^{\sum_{i\in F}\eta(i)}q^{\sum_{i\in F}(1-\eta(i))}, \tag{2.2.21}$$

where the sum over the empty set is taken to be 0, then β_p^F is a measure on $\big(\Omega, \mathcal{A}(F)\big)$ that measures the probability of events that depend only on outcomes during F.

Next set $\mathcal{A} = \bigcup\{\mathcal{A}(F) : \emptyset \neq F \subset\subset \mathbb{Z}^+\}$. Then $\mathcal{A}$ is an (cf. Exercise 2.1.16) algebra over Ω. However, $\mathcal{A}$ is *not* a σ-algebra. Nonetheless, we can define $\beta_p : \mathcal{A} \longrightarrow [0,1]$ so that $\beta_p(A) = \beta_p^F(A)$ if $A \in \mathcal{A}(F)$. To know that this definition is justified, we must make sure that if $A \in \mathcal{A}(F)\cap\mathcal{A}(F')$, where $F \neq F'$, then $\beta_p^F(A) = \beta_p^{F'}(A)$. To this end, note that $A \in \mathcal{A}(F\cap F')$, and therefore that it suffices to handle the case in which $F \subset F'$. Further, this case reduces to the one in which $F' = F\cup\{j\}$, where $j \notin F$. But if $\Gamma \subseteq \Omega(F)$, then $\Pi_F^{-1}\Gamma = \Pi_{F'}^{-1}\Gamma_0 \cup \Pi_{F'}^{-1}\Gamma_1$, where

$$\Gamma_k = \big\{\eta \in \Omega(F') : \eta \restriction F \in \Gamma \text{ and } \eta(j) = k\big\} \quad \text{for } k \in \{0,1\},$$

and so

$$\begin{aligned}\beta_p^{F'}\left(\Pi_{F'}^{-1}\Gamma\right) &= \beta_p^{F'}\left(\Pi_{F'}^{-1}\Gamma_0\right) + \beta_p^{F'}\left(\Pi_{F'}^{-1}\Gamma_1\right)\\ &= q\beta_p^F\left(\Pi_F^{-1}\Gamma\right) + p\beta_p^F\left(\Pi_F^{-1}\Gamma\right) = \beta_p^F\left(\Pi_F^{-1}\Gamma\right).\end{aligned}$$

Thus, we now know that β_p is well-defined and $\beta_p(\Omega) = 1$. In addition, β_p is finitely additive in the sense that, for any $n \in \mathbb{Z}^+$,

$$\beta_p\left(\bigcup_{m=1}^n A_m\right) = \sum_{m=1}^n \beta_p(A_m)$$

if the A_m's are mutually disjoint elements of $\mathcal{A}$. Indeed, by choosing $F \subset\subset \mathbb{Z}^+$ so that $\{A_m : 1 \leq m \leq n\} \subseteq \mathcal{A}(F)$, one can do the computation with β_p^F instead of β_p.

The preceding paragraphs summarize the presentation of coin tossing given in an elementary probability theory course. What is not usually covered in such a course is the extension of β_p to events like

$$A = \left\{\omega \in \Omega : \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n \omega(k) \text{ exists}\right\}$$

that depend on an infinite number of tosses. My aim here is to show that such an extension exists and that it can be constructed using the results in § 2.2.1.

The first step is to introduce a topology on Ω, and the one that I will choose is the one corresponding to pointwise convergence. That is, I want the sequence $\{\omega_m : m \geq 1\}$ to converge to ω if and only if for each $i \in \mathbb{Z}^+$ there is an m_i such that $\omega_m(i) = \omega(i)$ for all $m \geq m_i$. One way to describe this topology is to define

$$\rho(\omega, \omega') = \sum_{i=1}^{\infty} 2^{-i}|\omega(i) - \omega'(i)|,$$

and check that ρ is a metric on Ω for which convergence is the same as pointwise convergence. Further, as a topological space with metric ρ, Ω is compact. To see this, let $\{\omega_m : m \geq 1\} \subseteq \Omega$ be given. Then there exists a strictly increasing sequence $\{m_{1,\ell} : \ell \geq 1\} \subseteq \mathbb{Z}^+$ such that $\omega_{m_{1,\ell}}(1) = \omega_{m_{1,1}}(1)$ for all $\ell \in \mathbb{Z}^+$. Knowing $\{m_{1,\ell} : \ell \geq 1\}$, choose a strictly increasing subsequence $\{m_{2,\ell} : \ell \geq 1\}$ of $\{m_{1,\ell} : \ell \geq 1\}$ for which $\omega_{m_{2,\ell}}(2) = \omega_{m_{2,1}}(2)$ for all $\ell \in \mathbb{Z}^+$. Proceeding by induction on $k \in \mathbb{Z}^+$, produce $\{m_{k,\ell} : (k, \ell) \in (\mathbb{Z}^+)^2\}$ such that $\{m_{k+1,\ell} : \ell \geq 1\}$ is a strictly increasing subsequence of $\{m_{k,\ell} : \ell \geq 1\}$ and $\omega_{m_{k,\ell}}(k) = \omega_{m_{k,1}}(k)$ for all $\ell \geq 1$. If $m_k = m_{k,k}$ and $\omega(i) = \omega_{m_i}(i)$, then $\{\omega_{m_i} : i \geq 1\}$ is a subsequence of $\{\omega_m : m \geq 1\}$ and $\omega_{m_i} \longrightarrow \omega$ as $i \to \infty$.

It is clear that every $A \in \mathcal{A}$ is closed. In addition, if $\emptyset \neq F \subset\subset \mathbb{Z}^+$, $\omega \in A \in \mathcal{A}(F)$, and $i_F \equiv \max\{i : i \in F\}$, then $\rho(\omega', \omega) < 2^{-i_F} \implies \omega' \restriction F = \omega \restriction F$ and therefore $\omega' \in A$. Hence, every element of $\mathcal{A}$ is both open and closed. Moreover, for each $\omega \in \Omega$, $\left\{\Pi_F^{-1}(\{\omega \restriction F\}) : \emptyset \neq F \subset\subset \mathbb{Z}^+\right\}$ forms a countable neighborhood basis at ω. Indeed, given $\omega \in \Omega$ and $r > 0$, choose $n \geq 1$ such that $2^{-n} < r$, and observe that $\{\omega' : \omega'(i) = \omega(i) \text{ for } 1 \leq i \leq n\}$ is contained in the ρ-ball of radius r centered at ω.

Having made these preparations, we can turn to the construction. Take $\mathfrak{R} = \mathcal{A}$, and define $V(A) = \beta_p(A)$ for $A \in \mathcal{A}$. Then $\mathfrak{R}$ and V satisfy the hypotheses (1)–(5) at the beginning of § 2.2.1. Indeed, (1), (2), and (3) are obvious from the facts that $\mathcal{A}$ is an algebra and that β_p is finitely additive on $\mathcal{A}$. As for (4), the fact that each $A \in \mathcal{A}$ is both open and closed means that there is nothing to check. Finally, the following lemma shows that (5) holds.

Lemma 2.2.22. *If $\emptyset \neq S \subseteq \mathbb{Z}^+$ and $G \in \mathcal{A}(S)$ is open, then there is a sequence $\{A_m : m \geq 1\}$ of mutually disjoint elements of $\mathcal{A}(S) \cap \mathcal{A}$ for which $G = \bigcup_{m=1}^{\infty} A_m$.*

Proof: To produce a sequence $\{A_m : m \geq 1\} \subseteq \mathcal{A}$ of mutually disjoint elements of $\mathcal{A}(S)$ such that $G = \bigcup_{m=1}^{\infty} A_m$, one can proceed as follows. If S is finite, then one can take $A_1 = A$ and $A_n = \emptyset$ for $n \geq 2$. If S is infinite, let $\{i_n : n \geq 1\}$ be the strictly increasing enumeration of S, and set $F_n = \{i_1, \ldots, i_n\}$. Choose A_1 to be the largest element $A \in \mathcal{A}(F_1)$ with the property that $A \subseteq G$. Equivalently, A_1 is the union of all the $A \in \mathcal{A}(F_1)$ contained in G. Next, given $A_m \in \mathcal{A}(F_m)$ for $1 \le m \le n$, choose A_{n+1} to the largest $A \in \mathcal{A}(F_{n+1})$ contained in $G \setminus \bigcup_{m=1}^{n} A_m$. Obviously, these A_m's are mutually disjoint elements of $\mathcal{A} \cap \mathcal{A}(S)$, all of which are subsets of G. To see that they cover G, suppose that $\omega \in G$, and choose $n \ge 2$ for which $\omega' \in G$ whenever $\rho(\omega', \omega) < 2^{-i_{n-1}}$. Then $A \equiv \{\omega' : \omega'(i_m) = \omega(i_m) \text{ for } 1 \le i \le n\} \in \mathcal{A}(F_n)$ is a subset of G, and so either $\omega \in \bigcup_1^{n-1} A_m$ or $A \cap \bigcup_1^{n-1} A_m = \emptyset$, in which case $\omega \in A \subseteq A_n$. $\square$

Theorem 2.2.23. *Referring to the preceding, there exists a unique extension of β_p as a Borel probability measure on Ω, and this extension is regular. Finally, β_p is the unique Borel measure ν on Ω with the property that, for each $n \in \mathbb{Z}^+$ and $\eta \in \{0,1\}^n$,*

$$\nu\big(\{\omega \in \Omega : \omega(i) = \eta(i) \text{ for } 1 \le i \le n\}\big) = p^{\sum_{i=1}^{n} \eta(i)} q^{n - \sum_{i=1}^{n} \eta(i)}.$$

Proof: The existence of the extension as well as its regularity are guaranteed by Theorem 2.2.10. Furthermore, that theorem says that there is only one extension. Finally, suppose that ν is as in the last part of the statement. Because every non-empty element of $\mathcal{A}$ is the finite union of mutually disjoint sets of the form $\{\omega : \omega(m) = \eta(m) \text{ for } 1 \le m \le n\}$, where $\eta \in \{0,1\}^n$ for some $n \in \mathbb{Z}^+$, any ν that extends $\beta_p \restriction \mathcal{A}$ is therefore equal to β_p. $\square$

Because Bernoulli (again Jacob) made seminal contributions to the study of coin tossing, the Borel probability measure β_p in Theorem 2.2.23 is called the **Bernoulli measure** with parameter p. Before closing this discussion of coin tossing, it should be pointed out that the independence on which (2.2.20) was based extends to β_p as a Borel measure. To verify this, we will need the following lemma.

Lemma 2.2.24. *Suppose that $\emptyset \ne S \subset \mathbb{Z}^+$. If $B \in \mathcal{A}(S)$ and $B \subseteq H \in \mathfrak{G}(\Omega)$, then there is a $G \in \mathfrak{G}(\Omega) \cap \mathcal{A}(S)$ such that $B \subseteq G \subseteq H$. Hence, if $B \in \mathcal{A}(S) \cap \overline{\mathcal{B}}^{\beta_p}$, then $\overline{\beta_p}(B) = \inf\{\beta_p(G) : B \subseteq G \in \mathfrak{G}(\Omega) \cap \mathcal{A}(S)\}$.*

Proof: Given $\omega \in B$, note that $\Pi_S^{-1}(\{\omega \restriction S\}) \subset\subset H$. In particular, there exists an $n(\omega) \in \mathbb{Z}^+$ for which $\rho\big(\Pi_S^{-1}(\{\omega \restriction S\}), H^{\complement}\big) > 2^{-n(\omega)}$. Thus, if $F(\omega) = \{i \in S : i \le n(\omega)\}$, then $A(\omega) \equiv \Pi_F^{-1}\big(\{\omega \restriction F(\omega)\}\big) \subseteq H$. Indeed, if $\omega' \in A(\omega)$, determine $\omega'' \in \Omega$ by taking $\omega'' \restriction S = \omega \restriction S$ and $\omega'' \restriction S^{\complement} = \omega' \restriction S^{\complement}$. Then $\omega'' \in \Pi_S^{-1}(\{\omega \restriction S\})$ and $\rho(\omega', \omega'') \le 2^{-n(\omega)}$, which means that $\omega' \in H$. Now take $G = \bigcup\big\{A(\omega) : \omega \in B\big\}$, and observe that G is an open element of $\mathcal{A}(S)$ that contains B and is contained in H.

Given the preceding, the final assertion is an easy application of the fact, coming from Theorems 2.2.23 and 2.1.15, that

$$\overline{\beta_p}(B) = \inf\{\beta_p(G) : B \subseteq G \in \mathfrak{G}(\Omega)\}. \quad \square$$

Theorem 2.2.25. *Let $S \subseteq \mathbb{Z}^+$, and suppose that $B \in \overline{\mathcal{B}_\Omega}^{\beta_p} \cap \mathcal{A}(S)$ and $B' \in \overline{\mathcal{B}_\Omega}^{\beta_p} \cap \mathcal{A}(S^\complement)$. Then $\overline{\beta_p}(B \cap B') = \overline{\beta_p}(B)\overline{\beta_p}(B')$.*

Proof: Obviously, there is nothing to do when either B or B' is either empty or the whole of Ω. Thus, we will assume that $\emptyset \neq S \subset \mathbb{Z}^+$. Next suppose that $F \subset\subset S$ and $F' \subset\subset S^\complement$. Then, for any $A \in \mathcal{A}(F)$ and $A' \in \mathcal{A}(F')$, it follows easily from (2.2.21) that $\beta_p(A \cap A') = \beta_p(A)\beta_p(A')$. Next suppose that $G \in \mathcal{A}(S)$ and $G' \in \mathcal{A}(S^\complement)$ are open. By Lemma 2.2.22, $G = \bigcup_{m=1}^\infty A_m$, where $\{A_m : m \geq 1\}$ are mutually disjoint elements of $\mathcal{A} \cap \mathcal{A}(S)$, and $G' = \bigcup_{m'=1}^\infty A'_{m'}$, where $\{A'_m : m \geq 1\}$ are mutually disjoint elements of $\mathcal{A} \cap \mathcal{A}(S^\complement)$. Thus $\{A_m \cap A_{m'} : (m, m') \in (\mathbb{Z}^+)^2\}$ is a cover of $G \cap G'$ by mutually disjoint elements of $\mathcal{A}$, and $\beta_p(A_m \cap A_{m'}) = \beta_p(A_m)\beta_p(A_{m'})$ for all (m, m'). Hence, by Lemma 2.2.1,

$$\beta_p(G \cap G') = \sum_{(m,m') \in (\mathbb{Z}^+)^2} \beta_p(A_m)\beta_p(A'_{m'}) = \beta_p(G)\beta_p(G').$$

Finally, let B and B' be as in the statement. Then $\overline{\beta_p}(B \cap B') \leq \beta_p(G \cap G') = \beta_p(G)\beta_p(G')$ for any open $G \in \mathcal{A}(S)$ containing B and open $G' \in \mathcal{A}(S^\complement)$ containing B'. Hence, by Lemma 2.2.24, $\overline{\beta_p}(B \cap B') \leq \overline{\beta_p}(B)\overline{\beta_p}(B')$.

To prove the opposite inequality, let $\epsilon > 0$ be given, and choose open sets G and G' for which $B \subseteq G$, $B' \subseteq G'$, and $\overline{\beta_p}(G \setminus B) + \overline{\beta_p}(G' \setminus B') < \epsilon$. By Lemma 2.2.24, we may and will assume that $G \in \mathcal{A}(S)$ and $G' \in \mathcal{A}(S^\complement)$. But then

$$\begin{aligned}\overline{\beta_p}(B)\overline{\beta_p}(B') &\leq \beta_p(G)\beta_p(G') = \beta_p(G \cap G') \\ &= \overline{\beta_p}(B \cap B') + \overline{\beta_p}\big((G \cap G') \setminus (B \cap B')\big) \leq \overline{\beta_p}(B \cap B') + \epsilon,\end{aligned}$$

since $(G \cap G') \setminus (B \cap B') \subseteq (G \setminus B) \cup (G' \setminus B')$. □

In the jargon of probability theory, Theorem 2.2.25 is saying that the σ-algebra $\mathcal{A}(S) \cap \overline{\mathcal{B}_\Omega}^{\beta_p}$ is *independent* under $\overline{\beta_p}$ of the σ-algebra $\mathcal{A}(S^\complement) \cap \overline{\mathcal{B}_\Omega}^{\beta_p}$.

§ **2.2.5. Bernoulli and Lebesgue Measures:** For obvious reasons, the case $p = \frac{1}{2}$ is thought of as the mathematical model of a coin tossing game in which the coin is fair (i.e., unbiased). Thus, one should hope that $\mu_{\frac{1}{2}}$ has special properties, and the purpose of this subsection is to prove one such property.

There is a natural continuous map Φ taking Ω onto $[0, 1]$, the one given by

$$\Phi(\omega) = \sum_{n=1}^\infty 2^{-n}\omega(n). \tag{2.2.26}$$

By part (**ii**) of Exercise 2.1.19, Φ is measurable as a mapping from $(\Omega, \mathcal{B}_\Omega)$ to $\big([0, 1], \mathcal{B}_{[0,1]}\big)$. What I am going to show is that (cf. part (**iii**) of Exercise

2.1.19) $\Phi_*\beta_{\frac{1}{2}} = \lambda_{[0,1]}$, where (cf. Exercise 2.1.18) $\lambda_{[0,1]}$ is the Borel measure on $[0,1]$ obtained by restricting $\lambda_{\mathbb{R}}$ to $\mathcal{B}_{[0,1]}$. In fact, I am going to prove more. Namely, I am going to show that, in spite to the fact that Φ is not one-to-one,

$$\Phi(B) \in \mathcal{B}_{[0,1]} \text{ and } \lambda_{[0,1]}\big(\Phi(B)\big) = \beta_{\frac{1}{2}}(B) \quad \text{for all } B \in \mathcal{B}_\Omega. \tag{2.2.27}$$

Once we know this, we will have

$$\Phi_*\beta_{\frac{1}{2}}(\Gamma) = \beta_{\frac{1}{2}}\big(\Phi^{-1}(\Gamma)\big) = \lambda_{[0,1]}(\Gamma) \quad \text{for } \Gamma \in \mathcal{B}_{[0,1]} \tag{2.2.28}$$

as a trivial consequence.

The interest in (2.2.27) stems from the following considerations. Let ω_∞ be the element of Ω that is equal to 1 at all $n \in \mathbb{Z}^+$. Then $\Phi(\omega_\infty) = 1$. Next, let $\hat{\Omega}$ be the subset of Ω consisting of ω_∞ and any $\omega \in \Omega$ with the property that $\omega(n) = 0$ for infinitely many $n \in \mathbb{Z}^+$. Because $\Omega \setminus \hat{\Omega}$ is equal to the set of $\omega \in \Omega \setminus \{\omega_\infty\}$ for which there is an $m \in \mathbb{Z}^+$ such that $\omega(i) = 1$ for all $i \geq m$, it is clear that $\Omega \setminus \hat{\Omega}$ is countable and therefore an element of $\mathcal{B}_\Omega$ to which $\beta_{\frac{1}{2}}$ assigns measure 0. Hence, $\hat{\Omega} \in \mathcal{B}_\Omega$ and $\beta_{\frac{1}{2}}(\hat{\Omega}) = 1$. The advantage that $\hat{\Omega}$ has over Ω is that $\hat{\Phi} \equiv \Phi \restriction \hat{\Omega}$ is one-to-one and onto $[0,1]$. To see this, first note that ω_∞ is the one and only element of $\hat{\Omega}$ that $\hat{\Phi}$ takes to 1. Next, given $x \in [0,1)$, determine ω by

$$\omega(1) = \begin{cases} 0 & \text{if } x \in \big[0, \frac{1}{2}\big) \\ 1 & \text{if } x \in \big[\frac{1}{2}, 1\big) \end{cases}$$

and, for $j \geq 2$,

$$\omega(j) = \begin{cases} 0 & \text{if } x - \sum_{i=1}^{j-1} 2^{-i}\omega(i) < 2^{-j} \\ 1 & \text{if } x - \sum_{i=1}^{j-1} 2^{-i}\omega(i) \geq 2^{-j}. \end{cases}$$

One can use induction to check that $0 \leq x - \sum_{i=1}^{j} 2^{-i}\omega(i) < 2^{-j}$ for each $j \geq 1$. In particular, $\omega \in \hat{\Omega}$, since otherwise $m = \max\{i : \omega(i) = 0\}$ would be finite and $x - \sum_{i=1}^{m} 2^{-i}\omega(i) = 2^{-m}$, which would be a contradiction. Hence, $\hat{\Phi}$ is onto. To see that it is one-to-one, suppose that $\omega, \omega' \in \hat{\Omega}$ and that $\Phi(\omega) = \Phi(\omega') \in [0,1)$. Then neither ω nor ω' is ω_∞ and so each has infinitely many i's at which it vanishes. Now suppose that $\omega \neq \omega'$ and therefore that $m = \min\{i \in \mathbb{Z}^+ : \omega(i) \neq \omega'(i)\} < \infty$. Without loss in generality, we can assume that $\omega'(m) = 1$ and $\omega(m) = 0$. But then we would have

$$2^{-m} \leq 2^{-m} + \sum_{i=m+1}^{\infty} 2^{-i}\omega'(i) = \sum_{i=m+1}^{\infty} 2^{-i}\omega(i) < 2^{-m},$$

which is impossible.

From an arithmetic perspective, when $x \in [0,1)$, $\hat{\Phi}^{-1}(x)$ gives the coefficients of the dyadic expansion of x: the representation of x as $\sum_{i=1}^{\infty} 2^{-i}\omega(i)$ for which $\{i \in \mathbb{Z}^+ : \omega(i) = 1\}$ is minimal. Thus, since

$$\lambda_{[0,1]}\big(\{x \in [0,1] : \hat{\Phi}^{-1}(x) \in B\}\big) = \lambda_{[0,1]}\big(\hat{\Phi}(B)\big) = \lambda_{[0,1]}\big(\Phi(B)\big),$$

(2.2.27) says that *the statistics under* $\lambda_{[0,1]}$ *of the dyadic coefficients of a number in* $[0,1)$ *are given by* $\beta_{\frac{1}{2}}$, an observation that E. Borel seems to have been the first to make. See part (**iii**) of Exercise 3.1.15 for an application of this observation.

With the preceding as motivation, I turn now to the proof of (2.2.27). First note that, because $\Phi(B) \setminus \hat{\Phi}(B \cap \hat{\Omega})$ is countable for any $B \subseteq \Omega$, we need show only that

$$(*) \qquad \hat{\Phi}(B \cap \hat{\Omega}) \in \mathcal{B}_{[0,1]} \text{ and } \lambda_{\mathbb{R}^N}\big(\hat{\Phi}(B \cap \hat{\Omega})\big) = \beta_{\frac{1}{2}}(B) \quad \text{for all } B \in \mathcal{B}_{\Omega}.$$

Second, observe that we only need to check $(*)$ when $B = A_n(\eta) \equiv \{\omega : \omega(i) = \eta(i),\ 1 \le i \le n\}$ for some $n \in \mathbb{Z}^+$ and $\eta \in \{0,1\}^n$. To understand why this is enough, remember that $\hat{\Phi}$ is one-to-one and therefore, by part (**i**) of Exercise 2.1.19, preserves differences as well as unions. Hence, the set $\mathcal{F}$ of $B \in \mathcal{B}_{\Omega}$ for which $(*)$ holds is closed under differences as well as countable unions. In addition, because $\hat{\Phi}$ is onto, $\Omega \in \mathcal{F}$, and therefore $\mathcal{F}$ is a σ-algebra over Ω. Thus, $(*)$ will be proved once I show it holds for B's coming from a Π-system that generate $\mathcal{B}_{\Omega}$. But, every non-empty $A \in \mathcal{A}$ is the finite union of sets of the form $A_n(\eta)$ and (cf. Lemma 2.2.24) every open set in Ω is the countable union of elements of $\mathcal{A}$. Hence $\{\emptyset\} \cup \big\{A_n(\eta) : \eta \in \{0,1\}^n \ \&\ n \in \mathbb{Z}^+\big\}$ is a Π-system that generates $\mathcal{B}_{\Omega}$. Finally, given $n \in \mathbb{Z}^+$ and $\eta \in \{0,1\}^n$, it is an easy matter to check that

$$(2.2.29) \qquad [0,1) \cap \hat{\Phi}\big(A_n(\eta) \cap \hat{\Omega}\big) = \left[\sum_{i=1}^{n} 2^{-i}\eta(i), 2^{-n} + \sum_{i=1}^{n} 2^{-i}\eta(i)\right).$$

Thus, not only is $\hat{\Phi}\big(A_n(\eta) \cap \hat{\Omega}\big) \in \mathcal{B}_{[0,1]}$ but also $\lambda_{[0,1]}$ assigns it measure 2^{-n}, which is the same as the measure $\beta_{\frac{1}{2}}$ assigns to $A_n(\eta)$.

Exercises for § 2.2

EXERCISE 2.2.30. Suppose that G is an open subset of $\mathbb{R}^N$ and that $\Phi : G \longrightarrow \mathbb{R}^{N'}$ is **uniformly Lipschitz continuous** in the sense that there is an $L < \infty$ such that $|\Phi(y) - \Phi(x)| \le L|y - x|$ for all $x, y \in G$. Because Φ is continuous, it takes compact subsets of G to compact sets, and from this conclude that Φ takes elements of $\mathfrak{F}_\sigma(G)$ to elements of $\mathcal{B}_{\mathbb{R}^{N'}}$. Next, show that if $\Gamma \in \overline{\mathcal{B}_G}^{\lambda_{\mathbb{R}^N}}$ has Lebesgue measure 0, then $\Phi(\Gamma)$ is an element of $\overline{\mathcal{B}_{\mathbb{R}^N}}^{\lambda_{\mathbb{R}^{N'}}}$ that has Lebesgue measure 0. Finally, combine these to show that $\Phi(\Gamma) \in \overline{\mathcal{B}_{\mathbb{R}^N}}^{\lambda_{\mathbb{R}^{N'}}}$ for every $\Gamma \in \overline{\mathcal{B}_G}^{\lambda_{\mathbb{R}^N}}$.

EXERCISE 2.2.31. Let $\mathcal{B}$ be a σ-algebra over E with the property that $\{x\} \in \mathcal{B}$ for all $x \in E$. A measure μ on $(E, \mathcal{B})$ is said to be **non-atomic** if $\mu(\{x\}) = 0$ for all $x \in E$. Show that if there is a non-trivial (i.e., not identically 0), non-atomic measure on $(E, \mathcal{B})$, then E must be uncountable. Next, apply this to show that the existence of $\lambda_{\mathbb{R}^N}$ and β_p implies that $\mathbb{R}^N$, Ω, and $\hat{\Omega}$ must all be uncountable.

EXERCISE 2.2.32. It is clear that any countable subset of $\mathbb{R}$ has Lebesgue measure zero. However, it is not so immediately clear that there are uncountable subsets of $\mathbb{R}$ whose Lebesgue measure is zero. The goal of this exercise is to show how to construct such a set. For this purpose, start with the set $C_0 = [0,1]$, and let C_1 be the set obtained by removing the open middle third of C_0 (i.e., $C_1 = C_0 \setminus \left(\frac{1}{3}, \frac{2}{3}\right) = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$). Next, let C_2 be the set obtained from C_1 after removing the open middle third of each of the (two) intervals of which C_1 is the disjoint union. More generally, given C_n (which is the union of 2^n disjoint, closed intervals), let C_{n+1} be the set that one gets by by removing from C_n the open middle third of each of the intervals of which C_n is the disjoint union. Finally, set $C = \bigcap_{k=0}^{\infty} C_n$. The set C is called the **Cantor set**, and it turns out to be an extremely useful source of examples. In particular, show that C is an uncountable, closed subset of $[0,1]$ that has Lebesgue measure 0. See Exercise 8.3.22 for further information.

Here are some steps that you might want to follow.

(**i**) Since each C_n is closed, C is also. Next, show that $\lambda_{\mathbb{R}}(C_n) = \left(\frac{2}{3}\right)^n$ and therefore that $\lambda_{\mathbb{R}}(C) = 0$.

(**ii**) To prove that C is uncountable, refer to the notation in § 2.2.4, and define $\Psi : \Omega \longrightarrow [0,1]$ by

$$\Psi(\omega) = \sum_{i=1}^{\infty} \frac{2\omega(i)}{3^i}.$$

Show that Ψ is one-to-one.

(**iii**) In view of Theorem 2.2.23 and Exercise 2.2.31, one will know that C must be uncountable if $\Psi(\hat{\Omega}) \subseteq C$. To this end, first show that $[0,1] \setminus C$ can be covered by open intervals of the form $\left((2k-1)3^{-n}, 2k3^{-n}\right)$, where $n \in \mathbb{Z}^+$ and $1 \le k \le \frac{3^n - 1}{2}$. Next, show that $\Psi(\omega) > (2k-1)3^{-n} \implies \Psi(\omega) \ge 2k3^{-n}$ and therefore that $\Psi(\hat{\Omega}) \subseteq C$.

EXERCISES 2.2.33. Here is a rather easy application of Theorem 2.2.15. If $B_{\mathbb{R}^N}(c,r)$ is the open ball in $\mathbb{R}^N$ of radius r and center c, show that

$$\lambda_{\mathbb{R}^N}\left(B_{\mathbb{R}^N}(c,r)\right) = \lambda_{\mathbb{R}^N}\left(\overline{B_{\mathbb{R}^N}(c,r)}\right) = \Omega_N r^N, \quad \text{where } \Omega_N \equiv \lambda_{\mathbb{R}^N}\left(B_{\mathbb{R}^N}(0,1)\right)$$

is the (cf. (**iii**) in Exercise 5.1.13) **volume of the unit ball** in $\mathbb{R}^N$.

EXERCISE 2.2.34. If $\mathbf{v}_1, \dots,$ and $\mathbf{v}_N$ are vectors in $\mathbb{R}^N$, the **parallelepiped spanned by** $\{\mathbf{v}_1, \dots, \mathbf{v}_N\}$ is the set

$$P(\mathbf{v}_1, \dots, \mathbf{v}_N) \equiv \left\{ \sum_1^N x_i \mathbf{v}_i : x \in [0,1]^N \right\}.$$

When $N \geq 2$, the classical prescription for computing the *volume* of a parallelepiped is to take the product of *the area of any one side* times the length of the corresponding *altitude*. In analytic terms, this means that the volume is 0 if the vectors $\mathbf{v}_1, \dots, \mathbf{v}_N$ are linearly dependent and that otherwise the volume of $P(\mathbf{v}_1, \dots, \mathbf{v}_N)$ can be computed by taking the product of the volume of $P\big(\mathbf{v}_1, \dots, \mathbf{v}_{N-1}\big)$, thought of as a subset of the hyperplane $H\big(\mathbf{v}_1, \dots, \mathbf{v}_{N-1}\big)$ spanned by $\mathbf{v}_1, \dots, \mathbf{v}_{N-1}$, times the distance between the vector $\mathbf{v}_N$ and the hyperplane $H\big(\mathbf{v}_1, \dots, \mathbf{v}_{N-1}\big)$. Using Theorem 2.2.15, show that this prescription is correct when the *volume* of a set is interpreted as the Lebesgue measure of that set.

Hint: Take A to be the $N \times N$ matrix whose ith column is $\mathbf{v}_i$, and use Cramer's rule to compute $\det(A)$.

Exercise 2.2.35. Cauchy posed the problem of determining which functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ are **additive** in the sense that $f(x + y) = f(x) + f(y)$ for all $x, y \in \mathbb{R}$. The goal of this exercise is to show that an additive function that is $\overline{\mathcal{B}_{\mathbb{R}}}^{\lambda_{\mathbb{R}}}$-measurable must be linear. That is, $f(x) = f(1)x$ for all $x \in \mathbb{R}$.

(**i**) Show that, for each $x \in \mathbb{R}$ and rational number q, $f(qx) = qf(x)$. In particular, conclude that any continuous, additive function is linear.

(**ii**) Show that if f is bounded on some non-empty open set, then f is linear.

(**iii**) Assume that f is a $\overline{\mathcal{B}_{\mathbb{R}}}^{\lambda_{\mathbb{R}}}$-measurable, additive function. Choose an $R > 0$ for which $\Gamma = \{x \in \mathbb{R} : |f(x)| \leq R\}$ has strictly positive $\overline{\lambda_{\mathbb{R}}}$-measure, and use Lemma 2.2.16 and additivity to conclude that there is a $\delta > 0$ for which $|f(x)| \leq 2R$ on $(-\delta, \delta)$. After combining this with (**ii**), conclude that *every* $\overline{\mathcal{B}_{\mathbb{R}}}$*-measurable, additive function is linear.*

Exercise 2.2.36. In connection with Exercise 2.2.35, one should ask whether there are solutions to Cauchy's functional equation that are not linear. Because any such solution cannot be Lebesgue measurable, one should expect that its construction must require the axiom of choice. What follows is an outline of a construction.

(**i**) Let $\mathcal{A}$ denote the set of all subsets $A \subseteq \mathbb{R}$ that are linearly independent over the rational numbers $\mathbb{Q}$ in the sense that, for every finite subset $F \subseteq A$ and every choice of $\{\alpha_x : x \in F\} \subseteq \mathbb{Q}$, $\sum_{x \in F} \alpha_x x = 0 \implies \alpha_x = 0$ for all $x \in F$. Partially order $\mathcal{A}$ by inclusion, and show that every totally ordered subset of $\mathcal{A}$ admits an upper bound. That is, if $\mathcal{T} \subseteq \mathcal{A}$ and, for all $A, B \in \mathcal{T}$, either $A \subseteq B$ or $B \subseteq A$, then there exists an $M \in \mathcal{A}$ such that $A \subseteq M$ for all $A \in \mathcal{T}$. Now apply the Zorn's Lemma, which is one of the equivalent forms of the axiom of choice, to show that there exists an $M \in \mathcal{A}$ that is maximal in the sense that, for all $A \in \mathcal{A}$, $M \subseteq A \implies M = A$.

(**ii**) Referring to (**i**), show that M is a Hamel basis for $\mathbb{R}$ over the rationals. That is, for all $y \in \mathbb{R} \setminus \{0\}$ there exist a unique finite $F(y) \subseteq M$ and a unique choice of $\{q_x(y) : x \in F(y)\} \subset \mathbb{Q} \setminus \{0\}$ for which $y = \sum_{x \in F(y)} q_x(y)x$.

(**iii**) Continuing (**i**) and (**ii**), extend the definition of $q_x(y)$ so that $q_x(y) = 0$ if either $y = 0$ or $x \notin F(y)$. Then, for all $y \in \mathbb{R}$, $y = \sum_{x\in M} q_x(y)x$, where, for each y, all but a finite number of summands are 0. Next, let ψ be any $\mathbb{R}$-valued function on M, define $f : \mathbb{R} \longrightarrow \mathbb{R}$ so that $f(y) = \sum_{x\in M} \psi(x)q_x(y)x$, and show that f is always additive but that it is linear if and only if ψ is constant. In particular, for each $x \in M$, q_x is an additive, non-linear function. Conclude from this that, for each $x \in M$, $y \rightsquigarrow q_x(y)$ cannot be Lebesgue measurable and must be unbounded on each non-empty open interval.

EXERCISE 2.2.37. Here is another construction of the measures μ_F in Theorem 2.2.19. Set $F(\infty) = \lim_{x\to\infty} F(x)$, and define $F^{-1} : \big[0, F(\infty)\big) \longrightarrow \mathbb{R}$ so that

$$F^{-1}(x) = \inf\{y \in \mathbb{R} : F(y) \ge F(x)\}.$$

Check that F^{-1} is $\mathcal{B}_{[0,F(\infty))}$-measurable, and set

$$\mu(\Gamma) = (F^{-1})_*\lambda_{\mathbb{R}}(\Gamma) = \lambda_{\mathbb{R}}\big(\{x \in [0, F(\infty)) : F^{-1}(x) \in \Gamma\}\big) \quad \text{for } \Gamma \in \mathcal{B}_{\mathbb{R}}.$$

Show that μ is a finite Borel measure on $\mathbb{R}$ whose distribution function is F. Hence, $\mu = \mu_F$.

EXERCISE 2.2.38. A right-continuous, non-decreasing function $F : \mathbb{R} \longrightarrow \mathbb{R}$ is said to be **absolutely continuous** if for every $\epsilon > 0$ there exists a $\delta > 0$ such that $\sum_{n=1}^{\infty}\big(F(b_n)-F(a_n)\big) < \epsilon$ whenever $\{(a_n, b_n) : n \ge 1\}$ is a sequence of mutually disjoint open intervals satisfying $\sum_{n=1}^{\infty}(b_n - a_n) < \delta$. Show that an absolutely continuous F is uniformly continuous. Next, assume that F is bounded and tends to 0 at $-\infty$, and let (cf. Theorem 2.2.19) μ_F be the Borel measure on $\mathbb{R}$ for which F is the distribution function. Show that F is absolutely continuous as a function if and only if μ_F is (cf. Exercise 2.1.27) absolutely continuous with respect to $\lambda_{\mathbb{R}}$.

EXERCISE 2.2.39. Given a bounded, right-continuous, non-decreasing function F on $\mathbb{R}$, say that F is **singular** if for each $\delta > 0$ there exists a sequence $\{(a_n, b_n) : n \ge 1\}$ of mutually disjoint open intervals such that $\sum_{n=1}^{\infty}(b_n - a_n) < \delta$ and $F(\infty) - F(-\infty) = \sum_{n=1}^{\infty}\big(F(b_n) - F(a_n)\big)$. Assuming that F tends to 0 at $-\infty$, show that F is singular if and only if the measure μ_F for which it is the distribution function is (cf. Exercise 2.1.28) singular to $\lambda_{\mathbb{R}}$.

EXERCISE 2.2.40. As we saw in Exercise 2.2.37, all finite Borel measures on $\mathbb{R}$ can be written as an image of $\lambda_{\mathbb{R}}$. This fact is a particular example of the much more general fact that, under mild technical conditions, nearly every measure can be written as the image of $\lambda_{\mathbb{R}}$. The purpose of this exercise is to construct a measurable $f : [0, 1] \longrightarrow [0, 1]^2$ such that $\lambda_{[0,1]^2} = f_*\lambda_{[0,1]}$, where $\lambda_{[0,1]^2}$ is the restriction of $\lambda_{\mathbb{R}^2}$ to $\mathcal{B}_{[0,1]^2}$. To construct f, first define π_0 and π_1 on Ω into itself so that $[\pi_0(\omega)](i) = \omega(2i)$ and $[\pi_1(\omega)](i) = \omega(2i - 1)$ for $i \ge 1$. Next, define $f : [0, 1] \longrightarrow [0, 1]^2$ by

$$f = \big(\Phi \circ \pi_0 \circ \hat{\Phi}^{-1}, \Phi \circ \pi_1 \circ \hat{\Phi}^{-1}\big).$$

Show that f is a measurable map from $([0,1];\mathcal{B}_{[0,1]})$ onto $([0,1]^2,\mathcal{B}_{[0,1]^2})$ and that $\lambda_{[0,1]^2} = f_*\lambda_{[0,1]}$.

EXERCISE 2.2.41. If $\{S_1,\dots,S_n\}$ are mutually disjoint subsets of $\mathbb{Z}^+$ for some $n \geq 2$, show that

$$\overline{\beta_p}\big(A_1 \cap \cdots \cap A_n\big) = \overline{\beta_p}(A_1)\cdots\overline{\beta_p}(A_n)$$

for every choice of $\{A_1,\dots,A_n\} \subseteq \overline{\mathcal{B}_\Omega}^{\beta_p}$ with $A_m \in \mathcal{A}(S_m)$ for $1 \leq m \leq n$.

Springer

http://www.springer.com/978-1-4614-1134-5

Essentials of Integration Theory for Analysis
Stroock, D.W.
2011, XII, 244 p., Hardcover
ISBN: 978-1-4614-1134-5