Measure Theory

1 Measurable Spaces

A **measurable space** is a set S, together with a nonempty collection, $\mathcal{S}$, of subsets of S, satisfying the following two conditions:

1. For any A, B in the collection $\mathcal{S}$, the set[1] $A - B$ is also in $\mathcal{S}$.
2. For any $A_1, A_2, \dots \in \mathcal{S}$, $\cup A_i \in \mathcal{S}$.

The elements of $\mathcal{S}$ are called **measurable** sets. These two conditions are summarized by saying that the measurable sets are closed under taking finite differences and countable unions.

Think of S as the arena in which all the action (integrals, etc) will take place; and of the measurable sets are those that are "candidates for having a size". In some examples, *all* the measurable sets will be assigned a "size"; in others, only the smaller measurable sets will be (with the remaining measurable sets having, effectively "infinite size").

Several properties of measurable sets are immediate from the definition.

1. The empty set, $\emptyset$, is measurable. [Since $\mathcal{S}$ is nonempty, there exists some measurable set A. So, $A - A = \emptyset$ is measurable, by condition 1 above.]

2. For A and B any two measurable sets, $A \cap B$, $A \cup B$, and $A - B$ are all measurable. [The third is just condition 1 above. For the second, apply condition 2 to the sequence $A, B, \emptyset, \emptyset, \cdots$. For the first, note that $A \cap B = A - (A - B)$: Use condition 1 twice.] It follows immediately, by repeated application of these facts, that the measurable sets are closed under taking any finite numbers of intersections, unions, and differences.

3. For $A_1, A_2, \cdots$ measurable, their intersection, $\cap A_i$, is also measurable. [First note that we have the following set-theoretic identity: $A_1 \cap A_2 \cap A_3 \cap \cdots = A_1 - \{(A_1 - A_2) \cup (A_1 - A_3) \cup (A_1 - A_4) \cup \cdots\}$. Now, on the right, apply condition 1 above to the set-differences, and condition 2 to the union.] Thus, measurable sets are closed under taking countable intersections and unions.

Here are some examples of measurable spaces.

1. Let S be any set, and let $\mathcal{S}$ consist only of the empty set $\emptyset$. This is a (rather boring) measurable space.

[1] By $A - B$, we mean $A \cap B^c$, i.e., the set of all points of A that are not in B.

2. Let S be any set, and let $\mathcal{S}$ consist of all subsets of S. This is a measurable space.

3. Let S be any set, and let $\mathcal{S}$ consist of all subsets of S that are countable (or finite). This is a measurable space.

4. Let S be any set, and fix any nonempty collection $\mathcal{P}$ of subsets of S. Let $\mathcal{S}$ be the collection of subsets of S that result from the following construction. First set $\mathcal{S} = \mathcal{P}$. Now expand $\mathcal{S}$ to include all sets that result by taking differences and countable unions of sets in $\mathcal{S}$. Next, again expand $\mathcal{S}$ to include all sets that result by taking differences and countable unions of sets in (the already expanded) $\mathcal{S}$. Continue in this way[2], and denote by $\mathcal{S}$ the collection that results. Then $(S, \mathcal{S})$ is a measurable space. Thus, you can generate measurable spaces by starting with any set S, and *any* collection $\mathcal{P}$ of subsets of S (i.e., those that you really want to turn out, in the end, to be measurable). By expanding that original collection $\mathcal{P}$, as described above, you can indeed achieve a measurable space in which the chosen sets are indeed measurable.

5. Let $(S, \mathcal{S})$ be any measurable space, and let $K \subset S$ (not necessarily measurable). Let $\mathcal{K}$ denote the collection of all subsets of K that are $\mathcal{S}$-measurable. Then $(K, \mathcal{K})$ is a measurable space. [The two properties for $(K, \mathcal{K})$ follow immediately from the corresponding properties of $(S, \mathcal{S})$.] Thus, each subset of a measurable space gives rise to a new measurable space (called a **subspace** of the original measurable space).

6. Let $(S', \mathcal{S}')$ and $(S'', \mathcal{S}'')$ be measurable spaces, based on disjoint underlying sets. Set $S = S' \cup S''$, and let $\mathcal{S}$ consist of all sets $A \subset S$ such that $A \cap S' \in \mathcal{S}'$ and $A \cap S'' \in \mathcal{S}''$. Then $(S, \mathcal{S})$ is a measurable space. [The two properties for $(S, \mathcal{S})$ follow immediately from the corresponding properties of $(S', \mathcal{S}')$ and $(S'', \mathcal{S}'')$. For instance, the first property follows from: $(A - B) \cap S' = (A \cap S') - (B \cap S')$ and $(A - B) \cap S'' = (A \cap S'') - (B \cap S'')$.]

2 Measures

Let $(S, \mathcal{S})$ be a measurable space. A **measure** on $(S, \mathcal{S})$ consists of a nonempty subset, $\mathcal{M}$, of $\mathcal{S}$, together with a mapping $\mathcal{M} \xrightarrow{\mu} R^+$ (where R^+

[2]In more detail: Set $\mathcal{S}_1 = \mathcal{P}$. Let $\mathcal{S}_2$ the set that results from applying the above process to $\mathcal{S}_1$; then by $\mathcal{S}_3$ the set that results from applying the process to $\mathcal{S}_2$, etc. Then set $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \cdots$.

denotes the set of nonnegative reals), satisfying the following two conditions:

1. For any $A \in \mathcal{M}$ and any $B \subset A$, with $B \in \mathcal{S}$, we have $B \in \mathcal{M}$.

2. Let $A_1, A_2, \cdots \in \mathcal{M}$ be disjoint, and set $A = A_1 \cup A_2 \cup \cdots$. Then: This union A is in $\mathcal{M}$ if and only if the sum $\mu(A_1) + \mu(A_2) + \cdots$ converges; and when these hold that sum is precisely $\mu(A)$.

A set $A \in \mathcal{M}$ is said to **have measure**; and $\mu(A)$ is called the **measure** of A. Think of the collection $\mathcal{M}$ as consisting of those measurable sets that actually are assigned a "size" (i.e., of those size-candidates (in $\mathcal{S}$) that were successful); and of $\mu(A)$ as that size. Then the first condition above says that all sufficiently small measurable sets are indeed assigned size. The second condition says that the only excuse a measurable set A has for not being assigned a size is that "there is already too much measure inside A", i.e., that A effectively has "infinite measure". The last part of condition 2 says that measure is additive under taking unions of disjoint sets (something we would have wanted and expected to be true).

Several properties of measures are immediate from the definition.

1. The empty set $\emptyset$ is in $\mathcal{M}$, and $\mu(\emptyset) = 0$. [There exists some set $A \in \mathcal{M}$. Set $B = \emptyset$ and apply condition 1, to conclude $\emptyset \in \mathcal{M}$. Now apply condition 2 to the sequence $\emptyset, \emptyset, \cdots$ (having union $A = \emptyset$). Since $A \in \mathcal{M}$, we have $\mu(\emptyset) + \mu(\emptyset) + \cdots = \mu(\emptyset)$, which implies $\mu(\emptyset) = 0$.]

2. For any $A, B \in \mathcal{M}$, $A \cap B$, $A \cup B$, and $A - B$ are all in $\mathcal{M}$. Furthermore, if A and B are disjoint, then $\mu(A \cup B) = \mu(A) + \mu(B)$. [The first and third follow immediately from condition 1, since $A \cap B$ and $A - B$ are both subsets of A. For the second, apply condition 2 to the sequence $A - B, B, \emptyset, \emptyset, \cdots$ of disjoint sets, with union $A \cup B$. Additivity of the measures also follows from this, since when A and B are disjoint, $A - B = A$.]

3. For any $A, B \in \mathcal{M}$, with $B \subset A$, then $\mu(B) \leq \mu(A)$. [We have, by the previous item, $\mu(A) = \mu(B) + \mu(A - B)$.] Thus, "the bigger the set, the larger its measure".

4. For any $A_1, A_2, \cdots \in \mathcal{M}$, $\cap A_i \in \mathcal{M}$. [This is immediate from condition 1 above, since $\cap A_i \in \mathcal{S}$ and $\cap A_i \subset A_1 \in \mathcal{M}$.]

Thus, the sets that have measure (i.e., those that are in $\mathcal{M}$) are closed under finite differences, intersections and unions; as well as under countable intersections. What about countable unions? Let $A_1, A_2, \cdots$ be a sequence of sets in $\mathcal{M}$, *not necessarily disjoint*. First note that $\cup A_i = A$ can always be written as a union of a collection of *disjoint* sets in $\mathcal{M}$, namely of $A_1, A_2 - A_1, A_3 - A_2 - A_1, \cdots$. If the sum of the measures of the sets in this last list converges, then, by condition 2 above, we are guaranteed that $A \in \mathcal{M}$. And if the sum doesn't converge, then we are guaranteed that A is not in $\mathcal{M}$. Note, incidentally, that convergence of this sum is guaranteed by convergence of the sum $\mu(A_1) + \mu(A_2) + \mu(A_3) + \cdots$ (but, without disjointness, this last sum may exceed $\mu(A)$). In short, the sets that have measure are not in general closed under countable unions, but failure occurs only because of excessive measure.

Here are some examples of measures.

1. Let S be any set, let $\mathcal{S}$, the collection of measurable sets, be all subsets of S, let $\mathcal{M} = \mathcal{S}$, and, for $A \in \mathcal{M}$, let $\mu(A) = 0$. This is a (boring) measure.

2. Let S be any set, $\mathcal{S}$ all countable (or finite) subsets of S, $\mathcal{M}$ the collection of all finite subsets of S, and, for $A \in \mathcal{M}$, let $\mu(A)$ be the number of elements in the set A. This is is called **counting measure** on S. Note that the set S itself could be uncountable.

3. Let S be any set and $\mathcal{S}$ the collection of all subsets of S. Fix a nonnegative function $S \xrightarrow{f} R^+$ on S. Now let $\mathcal{M}$ consist of all sets $A \in \mathcal{S}$ such that $\Sigma_A\, f$ converges. Thus, $\mathcal{M}$ includes all the finite subsets of S; and possibly some countably infinite subsets (provided there isn't too much f on the subset); and possibly even some uncountable infinite subsets (provided f vanishes a lot on the subset). For $A \in \mathcal{M}$, set $\mu(A) = \Sigma_A f$. This is a measure. For $f = 1$, it reduces to counting measure.

4. Let $(S, \mathcal{S}, \mathcal{M}, \mu)$ be any measurable space/measure. Fix any $K \in S$ (not necessarily in $\mathcal{S}$). Denote by $\mathcal{K}$ the collection of all sets in $\mathcal{S}$ that are subsets of K; and by $\mathcal{M}_K$ the collection of all sets in $\mathcal{M}$ that are subsets of K. For $A \in \mathcal{M}_K$, set $\mu_K(A) = \mu(A)$. Then $(K, \mathcal{K}, \mathcal{M}_K, \mu_K)$ is again a measurable space/measure. [This is an easy check, using for each property, the corresponding property of $(S, \mathcal{S}, \mathcal{M}, \mu)$.] Thus, any subset of the underlying set S of a space with measure gives rise to another space with measure. This is called, of course, a **measure subspace**.

5. Let $(S', \mathcal{S}', \mathcal{M}', \mu')$ and $(S'', \mathcal{S}'', \mathcal{M}'', \mu'')$ be measurable spaces/measures, with S' and S'' disjoint. Set $S = S' \cup S''$; let $\mathcal{S}$ consist of $A \subset S$ such that $A \cap S' \in \mathcal{S}'$ and $A \cap S'' \in \mathcal{S}''$. Let $\mathcal{S}$ (resp, $\mathcal{M}$) consist of $A \subset S$ such that $A \cap S' \in \mathcal{S}'$ and $A \cap S'' \in \mathcal{S}'$ (resp, $\in \mathcal{M}'$ and $\in \mathcal{M}''$). Finally, for $A \in \mathcal{M}$, set $\mu(A) = \mu'(A \cap S') + \mu''(A \cap S'')$. This is a measurable space/measure. Thus, we may take the "disjoint union" of two measurable spaces/measures.

6. Let $(S, \mathcal{S})$ be a measurable space, and let $(\mathcal{M}, \mu)$ and $(\mathcal{M}, \mu')$ be two

measures on this space. [Note that they have the same $\mathcal{M}$.] Define $\mathcal{M} \overset{\mu+\mu'}{\to} R^+$ by: $(\mu + \mu')(A) = \mu(A) + \mu'(A)$. This is a measure, too. And, similarly, for any number $a > 0$, the mapping $\mathcal{M} \overset{a\mu}{\to} R^+$ with action $(a\mu)(A) = a\mu(A)$ is a measure. Thus, we can add measures, and multiply them by positive constants. [Why did we impose "$a > 0$"?]

We now obtain two results to the effect that "if a sequence of sets approaches (in a suitable sense) another set, then their measures approach the measure of that other set". In short, the measure of a set is "a continuous function of the set".

Theorem. Fix a measure space $(S, \mathcal{S}, \mathcal{M}, \mu)$, let $A_1 \subset A_2 \subset \cdots$ with $A_i \in \mathcal{M}$; and set $A = \cup A_i$. Then: $A \in \mathcal{M}$ if and only if the sequence $\mu(A_i)$ of numbers converges (as $i \to \infty$); and when these hold that limit is precisely $\mu(A)$.
Proof. Since the A_i are nested, we have the following set-theoretic identities:

$$A = A_1 \cup (A_2 - A_1) \cup (A_3 - A_2) \cup \cdots, \tag{1}$$

$$A_i = A_1 \cup (A_2 - A_1) \cup (A_3 - A_2) \cup \cdots \cup (A_i - A_{i-1}). \tag{2}$$

Note that the sets in the unions on the right are disjoint, and in $\mathcal{M}$. Since the union on the right of Eqn. (2) is finite, we have

$$\mu(A_i) = \mu(A_1) + \mu(A_2 - A_1) + \mu(A_3 - A_2) + \cdots + \mu(A_i - A_{i-1}). \tag{3}$$

Hence: The $\mu(A_i)$ converge if and only if the sum $\mu(A_1) + \mu(A_2 - A_1) + \mu(A_3 - A_2) + \cdots$ converges {by Eqn. (3)}; which in turn holds if and only if $A \in \mathcal{M}$ {by Eqn. (1) and the definition of a measure}; and that when these hold $\mu(A) = \lim \mu(A_i)$ {by Eqns. (1) and (3) and the definition of a measure}.

Theorem. Fix a measure space $(S, \mathcal{S}, \mathcal{M}, \mu)$, let $A_1 \supset A_2 \supset \cdots$, with $A_i \in \mathcal{M}$; and set $A = \cap A_i$. Then $A \in \mathcal{M}$, and $\lim_{i\to\infty} \mu(A_i) = \mu(A)$.
Proof. The proof is similar to that above (but easier), using the fact that $A_1 = A \cup (A_1 - A_2) \cup (A_2 - A_3) \cup \cdots$, where the sets on the right are disjoint, and in $\mathcal{M}$.

As a final result on measure spaces, we show that, under certain circumstances, a $(\mathcal{M}, \mu)$ that is "not quite a measure" can be made into one by including within $\mathcal{M}$ certain additional sets.

Let $(S, \mathcal{S})$ be a measurable space. Let $\mathcal{M}$ be a nonempty subset of $\mathcal{S}$, and let μ be a mapping, $\mathcal{M} \xrightarrow{\mu} R^+$. Let us suppose that this $(\mathcal{M}, \mu)$ satisfies the following two conditions:

1*. For any $A \in \mathcal{M}$ and any $B \subset A$, with $B \in \mathcal{S}$, we have $B \in \mathcal{M}$.

2*. Let $A_1, A_2, \cdots \in \mathcal{M}$ be disjoint, and set $A = A_1 \cup A_2 \cup \cdots$, their union. Then, provided $A \in \mathcal{M}$, the sum $\mu(A_1) + \mu(A_2) + \cdots$ converges, to $\mu(A)$.

Thus, this $(\mathcal{M}, \mu)$ is practically a measure on $(S, \mathcal{S})$. Condition 1* above is identical to condition 1 for a measure; and condition 2* is only somewhat weaker than condition 2 for a measure. All that has been left out, in condition 2*, is that portion of condition 2 that states: Whenever $\Sigma\mu(A_i)$ converges, then $A \in \mathcal{M}$. That is, this $(\mathcal{M}, \mu)$ is very nearly a measure, lacking only the requirement that disjoint unions of elements of $\mathcal{M}$, if not too obese measure-wise, are themselves in $\mathcal{M}$.

The present result is that, under the circumstances of the paragraph above, we can recover from that $(\mathcal{M}, \mu)$ a measure. The idea is to enlarge the original $\mathcal{M}$ to include the missing sets. Denote by $\hat{\mathcal{M}}$ the collection of all subsets of S that are of the form $\cup A_i$, where $A_1, A_2, \cdots$ is a sequence of disjoint sets in $\mathcal{M}$ for which $\Sigma\mu(A_i)$ converges; and let $\hat{\mu}(A) = \Sigma\mu(A_i)$. Note that every set A in $\mathcal{M}$ is automatically in $\hat{\mathcal{M}}$; with $\hat{\mu}(A) = \mu(A)$. This $\hat{\mathcal{M}}$, then, is just $\mathcal{M}$, augmented by certain unions of $\mathcal{M}$-sets. The present theorem is: This $(\hat{\mathcal{M}}, \hat{\mu})$ is a measure.

The first step of the proof is to show that the function $\hat{\mu}$ is well-defined. To this end, let $A = A_1 \cup A_2 \cup \cdots$ be in $\hat{\mathcal{M}}$ via condition ii) above. Let $B_1, B_2, \cdots$ be a second disjoint collection of elements of $\mathcal{M}$, with the same union: $\cup B_j = A$. We must show that $\Sigma\mu(B_j) = \Sigma\mu(A_i)$, i.e., that $\hat{\mu}(A)$, defined via the B_j, is the same as $\hat{\mu}(A)$ defined via the A_i. To see this, set, for $i, j = 1, 2, \cdots$, $C_{ij} = A_i \cap B_j$. Then the C_{ij} are disjoint and in $\mathcal{M}$, and their union is precisely A. But by condition 2* we have $\Sigma_i\mu(C_{ij}) = \mu(B_j)$ and $\Sigma_j\mu(C_{ij}) = \mu(A_i)$. That $\Sigma\mu(A_i) = \Sigma\mu(B_j)$ follows.

To complete the proof, we must show that $(\hat{\mathcal{M}}, \hat{\mu})$ satisfies conditions 1 and 2 for a measure. For condition 1: Let $A \in \hat{\mathcal{M}}$: We have $A = \cup A_i$, where the A_i are disjoint and are in $\mathcal{M}$, and are such that $\Sigma\mu(A_i)$ converges. Let $B \subset A$, with $B \in \mathcal{S}$. We must show that $B \in \hat{\mathcal{M}}$. But this follows, since $B = \cup(B \cap A_i)$, where the $B \cap A_i$ are disjoint, are in $\mathcal{M}$, and are such that $\Sigma\mu(B \cap A_i)$ converges. We leave condition 2 as an (easy) exercise.

Here is an example of an application of this result. Let $S = Z^+$, the set of positive integers, let $\mathcal{S}$ consist of all subsets of S, let $\mathcal{M}$ consist of all finite subsets of S, and, for $A \in \mathcal{M}$, let $\mu(A) = \Sigma_{n \in A}(1/2^n)$, where the sum on the right is finite. This $(\mathcal{M}, \mu)$ satisfies conditions 1* and 2* above. But it is not a measure, for it does not satisfy condition 2 for a measure space. In this case, the $\hat{\mathcal{M}}$ constructed above consists of *all* subsets of S, and, for $A \in \hat{\mathcal{M}}$, $\hat{\mu}(A) = \Sigma_{n \in A}(1/2^n)$, where now the sum on the right is over the (possibly infinite) set A. The measure space $(\mathcal{M}, \mu)$ here constructed will be recognized as a special case of Example 3 above.

Finally, we remark that, when the original $(\mathcal{M}, \mu)$ of the previous page happens to be a measure (i.e., happens to satisfy, not only condition 2*, but also condition 2), then $\hat{\mathcal{M}} = \mathcal{M}$, and $\hat{\mu} = \mu$.

3 Lebesque Measure

We now turn to what is certainly the most important example of a measure space: Lebesque measure. Let $S = R$, the set of reals. [The case $S = R^n$ is virtually identical, line-for-line, to this case; but $S = R$ makes writing easier.]

Set $I = (a, b)$, an open interval in R. The idea is that we want this interval to be measurable, with measure its length: $\mu(I) = b - a$. Let's try to turn this idea into a measure space. By condition 2 for a measure space, our collection $\mathcal{M}$ will have to include also sets of the form $K = I_1 \cup I_2 \cup \cdots$, a union of disjoint intervals, with measure $\mu(K) = \mu(I_1) + \mu(I_2) + \cdots$ — provided the sum on the right converges. And furthermore, by condition 1 for a measure space, $\mathcal{M}$ will also have to include differences of intervals, i.e., the half-closed intervals $[a, b)$ and $(a, b]$, with measures again $b - a$. So, we expand our original $\mathcal{M}$ to include these new sets. Next, let us return, with this new, expanded $\mathcal{M}$, to condition 2. By this condition, $\mathcal{M}$ must include also countable unions of the half-closed intervals. Returning to condition 1, we find that our $\mathcal{M}$ must include differences of these unions. Continue in this way, at each stage expanding the then-current $\mathcal{M}$ by including the new sets demanded by conditions 2 and 1. Does this process terminate? That is, do we, eventually, reach a point at which applying conditions 2 and 1 to the then-current $\mathcal{M}$ does not result in any further expansion of $\mathcal{M}$? If this did occur, then we would be done. Presumably, we would at that point be able to write down some general form for a set in this final $\mathcal{M}$, as well as a general formula for its measure. We would thus have our measure space. But, unfortunately, it turns out that this process does not terminate: Each passage through condition 2 and condition 1 requires that additional, new sets be included in $\mathcal{M}$. In short, this is not a very good way to construct our measure space. So, let's try a new strategy.

Fix any set $X \subset R$. Let $I_1, I_2, \cdots$ be any countable collection of open intervals that covers X [i.e., that are such that $X \subset \cup I_i$. Note that we do *not* require that the I_i be disjoint.] There always exists at least one such collection, e.g., $(-1,1), (-2,2), \cdots$. Now set $m = \Sigma\mu(I_i)$, the sum of the lengths of the I_i. This m is either a nonnegative number or "∞" (in case the sum fails to converge). We define the **outer measure** of X, written $\mu^*(X)$ to be the greatest lower bound of these m's, taken over all countable collections of open intervals that cover X; so $\mu^*(X)$ is either a nonnegative number, or "∞" (in case X is covered by no countable collection of intervals the sum of whose lengths converges).

The outer measure of X reflects "how much open-interval is required to cover X", i.e., is a rough measure of the "size" of X. For example, for X already an interval, $X = (a, b)$, we have $\mu^*(X) = (b - a)$, its length (an assertion that seems rather obvious, but is in fact a bit tricky to prove). As a second example, let X be the set of rational numbers. Order the rationals in any way, e.g., $3/5,\ -398/57,\ 3, \cdots$. Now fix any $\epsilon > 0$. Let I_1 be the interval of length ϵ centered on the first rational $(3/5)$; I_2 the interval of length $\epsilon/2$ centered on the second rational $(-398/57)$; and so on. Then these I_i cover X; and $\mu(I_1) + \mu(I_2) + \cdots = \epsilon + \epsilon/2 + \cdots = 2\epsilon$. But $\epsilon > 0$ is arbitarary: Thus, there exists a covering of X (the rationals) by open intervals the sum of whose lengths is as close to zero as we wish. We conclude: $\mu^*(X) = 0$. The same holds for any countable (or finite) subset of the reals.

The outer measure has the sort of behavior we might expect of a measure. For example: For $X \subset Y \subset R$, then $\mu^*(X) \leq \mu^*(Y)$ (which follows from the fact that any covering of Y is already a covering of X). For $X, Y \subset R$, $\mu^*(X \cup Y) \leq \mu^*(X) + \mu^*(Y)$ (which follows from the fact that the intervals in a covering of X taken together with the intervals in a covering of Y yields a collection of intervals that covers $X \cup Y$). Thus, it is tempting to try to construct our measure space using outer measure: Let $\mathcal{M}$ consist of all subsets X of $S = R$ with finite outer measure, and set $\mu(X) = \mu^*(X)$. But, unfortunately, this does not work, as the following example illustrates.

For a and b and two numbers in the interval $[0,1)$, write $a \sim b$ provided $a - b$ is a rational number. This is an equivalence relation. Consider its equivalence classes. One, for example, is the set of all numbers in $[0,1)$ of the form $\sqrt{2} + r$ with r rational; another, the set of all numbers in $[0,1)$ of the form $\pi + r$ with r rational; etc. Let $X \subset [0,1)$ be a set consisting of precisely one element from each equivalence class. So, for example, X contains exactly one number of the form $\sqrt{2}+r$ with r rational; exactly one of the form $\pi + r$ with r rational, etc. Next set, for r any rational number in $[0,1)$, $X_r = \{x \in [0,1) | (x-r) \in X \text{ or } (x - r + 1) \in X\}$. Thus, X_r is simply the set X, translated by amount r along the real line; and with that portion that is thereby moved outside $[0,1)$ reattached at the back. We now claim that these X_r (as r ranges through the rationals in $[0,1)$) have the following three properties:

1. The X_r are disjoint. {Let $x \in X_r$ and $x \in X_{r'}$. Then $(x - r)$ and $(x - r')$ (possibly after adding one in each case) is in X. But X contains exactly one element of the equivalence class "x plus a rational". Therefore, $r = r'$.}

2. The union of the X_r is all of $[0,1)$. {Let $x \in [0,1)$. Then X contains exactly one element of the equivalence class "x plus a rational", i.e., $x+r \in X$ for some rational r. That is, x is in X_r (or X_{r+1}, in case $r < 0$).}

3. For each r, $\mu^*(X_r) = \mu^*(X)$. {This follows, since outer measure is invariant under the translation and reattachment by which X_r is constructed from X.}

Now suppose, for contradiction, that we had a measure space based on outer measure. By the first two properties above, we would have $\Sigma\mu^*(X_r) = \mu^*([0,1)) = 1$, where the sum on the left is over all rationals $r \in [0,1)$. By the third property, all the $\mu^*(X_r)$ in this sum are the same number. But there exists no number with the property that, when it is summed with itself a countable number of times, the result is one. From this contradiction, we conclude: The choice for $\mathcal{M}$ of the sets of finite outer measure, and for μ the outer measure, does not result in a measure space[3].

[3]Note that the example above used (for the construction of X) the axiom of choice (one of the axioms of set theory). I am told that the rest of the axioms of set theory (i.e., with the axiom of choice omitted) are consistent with the statement "this $(\mathcal{M}, \mu)$ is a measure".

Thus, the outer measure is somewhat flawed as a representative of the "size" of a set, in the following sense. Certain sets (such as the X above) are, roughly speaking, so frothy that they cannot be covered efficiently by open intervals, and for these the outer measure is "too large".

This observation is the key to finding our measure space. For X and Y any two subsets of $S = R$, set $d(X,Y) = \mu^*(X - Y) + \mu^*(Y - X)$, so $d(X,Y)$ is a nonnegative number (or possibly "∞"). Think of $d(X,Y)$ as reflecting "the extent to which X and Y differ as sets", i.e., as an effective "distance" between the sets X and Y. This interpretation is supported by the following properties:

1. We have $d(X,Y) = 0$ whenever $X = Y$. [But note, that the converse fails, e.g., with Y consisting of X together with any one number not in X.]

2. For any subsets X, Y, Z of R, we have $d(X,Z) \leq d(X,Y) + d(Y,Z)$. {This follows from the facts that $X - Z \subset (X - Y) \cup (Y - Z)$ and $Z - X \subset (Z - Y) \cup (Y - X)$.} That is, $d(\ ,\ )$ satisfies the triangle inequality.

3. For any subsets X, X', Y, Y' of R, $d(X \cup Y, X' \cup Y') \leq d(X, X') + d(Y, Y')$, and similarly with "$\cup$" replaced by "$\cap$" or "$-$". {This follows from the fact that the set-difference of $X \cup Y$ and $X' \cup Y'$ is a subset of $(X - X') \cup (Y - Y')$; and similarly for "$\cap$" and "$-$".} That is, "nearby sets have nearby unions, intersections, and differences", i.e., the set operations are "continuous" as measured by $d(\ ,\ )$.

4. For any subsets X, Y of R, $|\mu * (X) - \mu * (Y)| \leq d(X,Y)$. {This follows from $X \cup (Y - X) = Y$ and $Y \cup (X - Y) = X$.} That is, outer measure is a $d(\ ,\ )$-continuous function of the set.

As we have remarked, the outer measure is sometimes "too large", and this fact renders it unsuitable as a measure. But the outer measure *is* suitable for generating an effective distance, $d(\ ,\ )$, between sets, for in this role its propensity to be "too-large" becomes merely an excess of caution.

We now turn to the key definition. Denote by $\mathcal{M}$ the collection of all subsets A of $S = R$ with the following property: Given any $\epsilon > 0$, there exists a $K \subset R$, where K is a finite union of open intervals, such that $d(A, K) \leq \epsilon$. And, for $A \in \mathcal{M}$, set $\mu(A) = \mu^*(A)$. In other words, the elements of $\mathcal{M}$ are the sets that can be "approximated" (as measured by $d(\ ,\ )$) by finite unions of open intervals. And, similarly, $\mu(A)$ is approximated by the sum of the lengths of the intervals in K (as follows from the fact that $d(A, K) \leq \epsilon$ implies $|\mu^*(A) - \mu^*(K)| \leq \epsilon$). It follows, in particular, that $\mu(A)$ is not "∞".

Here are two examples. Let A be the set of rational numbers. Then

$A \in \mathcal{M}$, with $\mu(A) = 0$. Indeed, given $\epsilon > 0$, choose $K = \emptyset$. Then $d(A, K) = \mu^*(A - \emptyset) + \mu^*(\emptyset - A) = \mu^*(A) - \mu^*(\emptyset) = 0 \leq \epsilon$. More generally, any set $A \subset R$ with $\mu^*(A) = 0$ is in $\mathcal{M}$. Next, let $A = (a, b)$, an open interval. Then $A \in \mathcal{M}$, with $\mu(A) = (b - a)$. Indeed, given $\epsilon > 0$, choose $K = A$, whence $d(A, K) = 0$.

We next show that this $(\mathcal{M}, \mu)$ has the properties necessary for a measure space.

1. Let $A, B \in \mathcal{M}$. Then $A - B \in \mathcal{M}$. Proof: Fix $\epsilon > 0$. Let K_A (resp., K_B) be a finite collection of open intervals with $d(A, K_A) \leq \epsilon$ (resp, $d(B, K_B) \leq \epsilon$). Let K be $K_A - K_B$, with any endpoints (arising from the set-difference) removed. Then $d(A - B, K) \leq d(A, K_A) + d(B, K_B) \leq 2\epsilon$.

2. Let $A_1, A_2, \cdots \in \mathcal{M}$ be disjoint, and let $\mu(A_1) + \mu(A_2) + \cdots$ converge. Then $A = \cup A_i \in \mathcal{M}$, and $\mu(A) = \Sigma\mu(A_i)$. Proof: Fix $\epsilon > 0$. Choose integer n such that $\mu(A_{n+1}) + \mu(A_{n+2}) + \cdots \leq \epsilon$. Next, choose K_1, a finite union of intervals, such that $d(A_1, K_1) \leq \epsilon$; K_2, such that $d(A_2, K_2) \leq \epsilon/2$, etc., up to K_n. Set $L = A_1 \cup A_2 \cup \cdots \cup A_n$ and $K = K_1 \cup K_2 \cup \cdots \cup K_n$, also a finite union of intervals. Then we have $d(A, K) \leq d(A, L) + d(L, K) \leq \epsilon + 2\epsilon$. This proves $A \in \mathcal{M}$. That $\mu(A) = \Sigma\mu(A_i)$ follows, first choosing the $K_1, \cdots, K_n$ disjoint so $\mu^*(K) = \Sigma_1^n \mu^*(K_i)$, and then using $|\mu^*(A) - \mu^*(L)| \leq \epsilon$, $|\mu^*(L) - \mu^*(K)| \leq 2\epsilon$, and $|\mu^*(K_i) - \mu^*(A_i)| \leq \epsilon/2^{i-1}$.

Now let $S = R$, let $\mathcal{S}$ be all subsets of R obtained by taking countable unions of elements of this $\mathcal{M}$. Then $(S, \mathcal{S}, \mathcal{M}, \mu)$ is a measure space, called **Lebesque measure**. The elements of $\mathcal{S}$ are called **Lebesque measurable** sets, and, for $A \in \mathcal{M}$, the number $\mu(A)$ is called the **Lebesque measure** of A.

In the land of measure spaces, the more sets that are measurable the better. Do there exists measures that are better, in this sense, than Lebesque measure? That is, does there exist a measure $(\hat{\mathcal{M}}, \hat{\mu})$ on R that is an extension of Lebesque measure, in the sense that $\hat{\mathcal{M}}$ is a proper superset of $\mathcal{M}$, and $\hat{\mu}$ agrees with μ on $\mathcal{M}$? It turns out that there does. Let X denote any non-measurable set of finite outer measure, e.g., the set constructed on page 9. Let $\hat{\mathcal{S}}$ consist of all subsets of R of the form $(A \cap X) \cup (B - X)$, where A and B are measurable. Thus, for example, choosing $A = B$ we conclude that $\hat{\mathcal{S}} \supset \mathcal{S}$; and, choosing $A \supset X$ and $B = \emptyset$, we conclude that $X \in \hat{\mathcal{S}}$. This collection is closed under differences and countable unions (as follows immediately from the fact that $\mathcal{S}$ is). Let $\hat{\mathcal{M}} \subset \mathcal{S}$ consist of those sets of this form with B having finite measure; and, for any such set, set

$\hat{\mu}((A \cap X) \cup (B - X)) = \mu^*(A \cap X) + \mu(B) - \mu^*(B \cap X)$. Thus, for example, $X \in \hat{\mathcal{M}}$, with $\hat{\mu}(X) = \mu^*(X)$; and, for $A \in \mathcal{M}$, $\hat{\mu}(A) = \mu(A)$. One checks that this $(\hat{\mathcal{M}}, \hat{\mu})$ is indeed a measure space, and that it is indeed an extension of Lebesque measure. Since $X \in \hat{\mathcal{M}}$ but $X \notin \mathcal{M}$, this is a proper extension.

4 Integrals

Let $(S, \mathcal{S})$ be a measurable space (not necessarily Lebesque). A real-valued function, $S \xrightarrow{f} R$, is said to be **measurable** provided: For O any open set in the reals, $f^{-1}[O]$ is measurable. This definition is hauntingly similar to that of continuity, the essential difference being that, since the domain in this case is a measurable space instead of a topological space, we require measurability, instead of open-ness, of $f^{-1}[O]$. As an example of this notion, we note that every continuous function, $R \xrightarrow{f} R$ is Lebesque-measurable (for, for f containuous, each $f^{-1}[O] \subset R$ is open, and hence Lebesque-measurable). As a second example, let, in a general measurable space, $A_1, \cdots, A_s$ be a finite number of disjoint measurable sets, whose union is S. Fix numbers $a_1, \cdots, a_s$. Let f be the function such that $f(x) = a_i$ whenever $x \in A_i$, for $i = 1, \cdots, s$. Thus, this function is constant on each of the elements of a finite, measurable partition of S. This function f is measurable. A measurable function f with finite range (i.e., a function of the form above) is called a **step function**.

Here are two elementary properties of measurable functions. For the first, let $S \xrightarrow{f} R$ be measurable. It then follows, using directly the properties of $\mathcal{S}$, that inverse images, under f, of differences of open sets (in R); of countable unions of such differences; of differences of such countable unions; etc., are all measurable. For the second property, let $S \xrightarrow{f} R$ and $S \xrightarrow{g} R$ be measurable. Then $S \xrightarrow{f+g} R$ is also measurable. To see this, let $S \xrightarrow{w} R^2$ be the map with $w(x) = (f(x), g(x))$. Then for each open rectangle, $(a, b) \times (c, d)$, in R^2, $w^{-1}[(a, b) \times (c, d)] = f^{-1}[(a, b)] \cap g^{-1}[(c, d)]$ is measurable, whence w^{-1} of any countable union of such rectangles is measurable, whence w^{-1} of any open subset of R^2 is measurable. Now let $O \subset R$ be open. Then $U = \{(\alpha, \beta) | \alpha + \beta \in O\}$ is open in R^2, whence $(f + g)^{-1}[O] = w^{-1}[U]$ is measurable. More generally, any continuous function applied to two (or, indeed, to any finite number) of measurable functions is measurable. We note also that any continuous function applied to two (or any finite number) of step functions is a step function.

We are now ready to do integrals. Let $(S, \mathcal{S}, \mathcal{M}, \mu)$ be a measure space. Let us, for the present discussion, restrict ourselves to the following case: **The set S itself has finite measure.** We shall relax this assumption (which is made solely to avoid, at this point of the discussion, possibly divergent integrals) shortly.

Let f be any step function (with values $a_1, \cdots, a_s$ on measurable (and, by the assumption above, finitely measurable) sets $A_1, \cdots, A_s$). We define the **integral** of this function f (over S, with respect to measure μ) by $\int_S f\, d\mu = a_1\mu(A_1) + \cdots + a_s\mu(A_s)$. We note that this is the right sort of expression to be the integral: It multiplies the value of f by the size of the region on which f takes that value, and sums over the (finite number of) different values that a step function can assume. As an example, let our measure space be Lebesque measure on $(0, 1)$, and let f be the (step) function with value 1 on the rationals (in $(0, 1)$), 0 on the irrationals. The integral of this function is zero (since the Lebesque measure of this set of rationals is zero).

This integral has the properties we would expect of an integral. For f any step function, $\int_S (af)\, d\mu = a \int_S f\, d\mu$ $\{$for $\sum (aa_i)\, \mu(A_i) = a\ \sum (a_i)\, \mu(A_i)\}$. Furthermore, for f and g any two step functions we have $\int_S (f + g)\, d\mu = \int_S f\, d\mu + \int_S g\, d\mu$. $\{$To see this, let f have values $a_1, \cdots, a_s$ on $A_1, \cdots A_s$; and g values $b_1, \cdots, b_t$ on $B_1, \cdots, B_t$. Then $(f + g)$ has values $a_i + b_j$ on $A_i \cap B_j$, for $i = 1, \cdots, s$ and $j = 1, \cdots, t$. So, $\int_S (f + g)\, d\mu = \sum_{i,j} (a_i + b_j)\, \mu(A_i \cap B_j) = \sum_{i,j}\, a_i\, \mu(A_i \cap B_j) + \sum_{i,j}\, b_j\, \mu(A_i \cap B_j)$. Now carry out the j-sum in the first term on the right (to obtain $\int_S f\, d\mu$), and the i-sum in the second (to obtain $\int_s g\, d\mu$).$\}$ In short, the integral is linear in the step-function integrated. A further, key, property of this integral is that "small (step) functions have small integral". We claim: Let step function f satisfy $|f| \leq \epsilon$. Then $|\int_S f\, d\mu| \leq \epsilon\mu(S)$. This is immediate, from $|\sum a_i\, \mu(A_i)| \leq$ (max $|a_i|$) $\sum\ \mu(A_i)$.

We now expand the applicability of our integral from step functions to all bounded, measurable functions on our finite measure space. This expansion is based on the following key fact.

Theorem. Let $S \xrightarrow{f} R$ be a bounded measurable function on $(S, \mathcal{S})$. Then for every $\epsilon > 0$ there exists a step function g on $(S, \mathcal{S})$ such that $|f - g| \leq \epsilon$.

Proof: Let $I_1, \cdots, I_s$ be a finite collection of disjoint half-open intervals whose union covers the range of f. For each i, set $A_i = f^{-1}[I_i] \in \mathcal{S}$; and choose $a_i \in I_i$. Then the function g on S with value a_i in A_i is a step function, and satisfies $|f - g| \leq \epsilon$.

Now fix any bounded measurable function f on our finite measure space. Fix a sequence of ϵ's approaching zero, and, for each, choose a step function g_ϵ with $|f - g_\epsilon| \leq \epsilon$. Then (since $|g_\epsilon - g_{\epsilon'}| \leq (\epsilon + \epsilon')$) we have $|\int_S g_\epsilon \, d\mu - \int_S g_{\epsilon'} \, d\mu| \leq (\epsilon + \epsilon') \, \mu(S)$, whence the $\int_S g_\epsilon \, d\mu$ form a Cauchy sequence. By the **integral** of f (over S, with respect to measure μ), we mean the number to which this sequence converges: $\int_S f \, d\mu = \lim_{\epsilon \to 0} \int_S g_\epsilon \, d\mu$ (noting that this number is independent of the particular sequence of g_ϵ chosen).

This integral (of a bounded measurable function on a finite measure space) inherits, immediately, the corresponding properties of the integral of a step function: i) $\int_S (af) \, d\mu = a \int_S f d\mu$; ii) $\int_S (f + g) \, d\mu = \int_S f \, d\mu + \int_S g \, d\mu$; and iii) $|\int_S f \, d\mu| \leq \text{lub} \, |f| \, \mu(S)$.

Let $(S, \mathcal{S}, \mathcal{M}, \mu)$ be a finite measure space, and let $A \subset S$ be measurable. Then, as we have seen earlier, we recover, in the obvious way, a measure space based on A: The measurable sets in this space are the measurable subsets of A; and the measure of such a set is simply its μ-measure. Next, let f be any measurable function on S. Then the restriction of f to A is a measurable function on this measure (sub-)space. Furthermore, if the original function f on S was bounded, then its restriction to A is also bounded. Under these circumstances (f a bounded measurable function on a finite measure space and $A \subset S$ measurable) we define the **integral** of f over A, written $\int_A f \, d\mu$, to be the integral over the measure (sub-)space A of the function f-restricted-to-A.

A key property of this integral is that it gives rise to a countable additive set function, in the following sense. Let $A_1, A_2, \cdots$ be disjoint measurable sets, with $\cup A_i = S$. Then, we claim, $\sum (\int_{A_i} f \, d\mu) = \int_S f \, d\mu$. It suffices to prove this claim for the case in which f a step function (for replacing, in the formula above, a general bounded, measurable function f by a step function g with $|f - g| \leq \epsilon$ changes neither side of that equation by more than $\epsilon \, \mu(S)$). So, let f be the step function taking values $b_1, b_2, \cdots, b_s$ on disjoint measurable sets $B_1, B_2, \cdots, B_s$. Consider the expression $\sum b_j \, \mu(A_i \cap B_j)$, where the sum is over $i = 1, 2, \cdots$ and $j = 1, 2, \cdots, s$. Fixing j and carrying out the sum over i, we obtain $b_j \, \mu(B_j)$ (by the additivity property of μ); while fixing i and carrying out the sum over j, we obtain $\int_{A_i} f\, d\mu$ (by definition of the integral of a step function). Now apply $\sum_j$ to the first (yielding $\int_S f\, d\mu$); and apply $\sum_i$ to the second (yielding $\sum(\int_{A_i} f\, d\mu)$). The result follows.

To summarize, we first defined the integral of a step function over a finite measure space (by the obvious formula); and then defined the integral of an arbitrary bounded measurable function over that space (by approximating that function by step functions). We now wish to relax the conditions that the measure space be finite; and that the function be bounded. In doing so, we shall encounter a new phenomenon: Our integrals may in some cases (i.e., for some functions and some regions of integration) fail to converge.

Let $(S, \mathcal{S}, \mathcal{M}, \mu)$ be an *any* measure space; and let f be a measurable function thereon. We assume for the present that this f is nonnegative (an assumption we shall relax in a moment). Next, denote by $\mathcal{M}'$ the collection of all measurable $A \subset S$ having finite measure, and on which f is bounded. [Note that there are many such sets: Take, e.g., $f^{-1}[(a, b)]$, and, if this set fails to have finite measure, take any finite-measure subset of it.] For any such $A \in \mathcal{M}'$, set $\mu'(A) = \int_A f\, d\mu$, noting that the integral on the right makes sense. Now, this $(\mathcal{M}', \mu')$ that we have just constructed will not in general be a measure space (as defined on page 3): While it always satisfies condition 1 of that definition, it may fail to satisfy condition 2. However, this $(\mathcal{M}', \mu')$ *does* satisfy the conditions 1* and 2* listed on page 6. [Condition 1* is immediate, while condition 2* is what we just showed two paragraphs ago.] As we showed on page 6, we may, under these conditions, expand $(\mathcal{M}', \mu')$ to a measure space. That construction, in more detail, is the following. Consider any sequence, $A_1, A_2, \cdots$ of disjoint sets in $\mathcal{M}'$ (so each A_i has finite measure, and on each f is bounded), such that $\sum \mu'(A_i)$ $(= \sum(\int_{A_i} f\, d\mu))$ converges. Let $\hat{\mathcal{M}}'$ consist of all subsets of S given as the union of such A_i; and, for $A = \cup A_i \in \hat{\mathcal{M}}'$, set $\hat{\mu}' = \sum \mu'(A_i)$. Then, as we showed earlier, this $(\hat{\mathcal{M}}', \hat{\mu}')$ is a measure space. We say that the measurable function $f \geq 0$ is **integrable** over $A \subset S$ provided $A \in \hat{\mathcal{M}}'$; and, for such a set A, we write $\int_A f\, d\mu = \mu'(A)$. Thus, a general nonnegative measurable function on a general measure space is integrable over measurable $A \subset S$ provided that A can be written as a countable union of disjoint sets, each of finite measure and on which f is bounded, such that the sum of the integrals of f over those sets converges; and in this case the integral is given by that sum.

It is easy, finally, to relax the condition $f \geq 0$ (which was made solely in order to invoke the earlier result on constructing measure spaces). Any measurable function f can be written uniquely as $f = f^+ - f^-$, where $f^+ \geq 0$ and $f^- \geq 0$ are measurable function. We say that f is **integrable** on $A \subset S$ provided both f^+ and f^- are, and we set $\int_A f\, d\mu = \int_A f^+\, d\mu - \int_A f^-\, d\mu$. This integral inherits from its predecessors linearity in the function.

So, the subject of integration (of arbitrary measurable functions over arbitrary measure spaces) is a remarkably simple one. We progress in turn from integration of step functions over finite measure spaces (the obvious formula) to integration of bounded functions over finite measure spaces (as limits of step-function integrals) to integration of nonnegative measurable functions over arbitrary measure spaces (restricting the region of integration to achieve convergence of the integral) to integration of arbitrary measurable functions over arbitrary measure spaces (writing such a function as the difference of nonnegative functions). It is true that some work has to be invested at the beginning, to get the notion of a measure in the first place. But, once the ground is prepared, things go easily and smoothly. General integration theory is much simpler, much more general, and much more useful than the theory of the Riemann integral.