Definition 1 : Let X be a non-empty set. A collection $\mathcal{A}$ of subsets of X is called an ALGEBRA on X if

(a) $X \in \mathcal{A}$.
(b) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$ ($\Rightarrow \emptyset \in \mathcal{A}$, in light of (a).)
(c) If $A_1, \dots, A_n \in \mathcal{A}$, then $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$.

(note that (b),(c) together $\Rightarrow$ closure for finite intersections.)

Definition 2 : Let X be a nonempty set. A collection $\mathcal{B}$ of subsets of X is called a σ- ALGEBRA on X if

(a) $X \in \mathcal{A}$
(b) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$
(c) If $A_1, A_2, A_3, \dots \in \mathcal{A}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Note: every σ-algebra on X is automatically an algebra on X, but not vice versa.

Examples: 1) $\mathcal{P}(X)$ - the collection of all subsets of a set - is a σ-algebra.
2) $\{\emptyset, X\} =: \mathcal{A}$. Then $\mathcal{A}$ is a (very boring!) σ-algebra on X.
3) Let $\mathcal{A} := \{A \subseteq X : A \text{ or } A^c \text{ is finite}\}$, where X is an infinite set. Then $\mathcal{A}$ is an algebra, but not a σ-algebra.

FACT: If X is a nonempty set, then the intersection of an arbitrary non-empty collection of σ-algebras on X is a σ-algebra on X.

COROLLARY: Let X be a nonempty set and $\mathcal{F}$ a family of subsets of X. Then there exists a smallest σ-algebra on X containing $\mathcal{F}$.

We will call this smallest σ-algebra the σ-algebra GENERATED BY $\mathcal{F}$ and write $\sigma(\mathcal{F})$.

Important example: Let X be a topological space. Then the σ-algebra generated by the collection of OPEN SETS is called the BOREL σ-ALGEBRA. In case $X = \mathbb{R}$, we will write $\mathcal{B}(\mathbb{R})$ and for $X = \mathbb{R}^n$, $n \geq 2$, we will write $\mathcal{B}(\mathbb{R}^n)$. Note that the σ-algebra $\mathcal{B}(\mathbb{R})$ can be generated by various collections of sets, for instance:

(a) all closed sets,
(b) all subintervals of the form $(-\infty, b]$,
(c) all subintervals of the form $(a, b]$, amongst others. (Similarly for $\mathcal{B}(\mathbb{R}^n)$.)

In particular, all G_δ and F_σ sets are Borel sets, where a G_δ set is a countable intersection of open sets and an F_σ set is a countable union of closed sets.

Definition 3: Let X be a set, $\mathcal{B}$ a σ-algebra on X and let $m : \mathcal{B} \to [0, +\infty]$ be a set function on $\mathcal{B}$ satisfying:

(a) $m(\emptyset) = 0$.

(b) $m\left(\dot{\bigcup}_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} m(A_n)$, where "$\dot{\cup}$" indicates a DISJOINT union of sets.

Then m is said to be a MEASURE on X.

Examples:

1) Let X be a non-empty set, $x \in X$ and define $\delta_x : \mathcal{P}(X) \to \{0,1\}$ by setting

$$\delta_x(A) := \begin{cases} 0 & \text{if } x \notin A; \\ 1 & \text{if } x \in A. \end{cases}$$

(DIRAC POINT MASS)

2) Let X be an arbitrary set and $\mathcal{B}$ be a σ-algebra on X. Define $\mu : \mathcal{B} \to [0, +\infty]$ by setting $\mu(A) := n$ if $A \in \mathcal{B}$ contains exactly n elements and letting $\mu(A) := +\infty$ if A is an infinite set. This is a measure (check!), it is called COUNTING MEASURE on $(X, \mathcal{B})$.

Obvious questions / observations: First of all, these measures are a bit stupid (only half true, certainly f. the point masses), secondly, why do we need all this σ-algebra business?

In order to think about the second point and then go on to construct some more intuitively meaningful examples, it helps to pin down exactly what we might want from a measure. If we think of an abstract measure as a mathematical version of getting busy with a metre stick, we most likely will want the following things:

- To be able to measure all subsets of our given set X.
- That the measure of a set should be non-negative.
- Additivity (that is, exactly part (b) of the definition).
- If we think of Euclidean space, measure ought to be translation invariant, i.e., rigid motions applied to a set do not change its size.
- Again in Euclidean space, subintervals of the line, or rectangles in $\mathbb{R}^2$ and so on, ought to get a "reasonable" value assigned to them.

It turns out that this wish-list is unachievable, at least if we want to work in a model of mathematics including the axiom of choice. (Actually, this can be weakened somewhat ...)

Note: ex 1 above fails on both the last two points, ex 2 only fails on the final one.

Vitali's Construction of a non-measurable set. (in $\mathbb{R}$)

Define the relation $\sim$ on $\mathbb{R}$ by setting $x \sim y$ iff $x-y$ is rational. It is easy to check that this is an equivalence relation and the equivalence classes are of the form $\mathbb{Q}+x$, for some $x \in \mathbb{R}$. These equivalence classes are disjoint and each intersects the unit interval $(0,1)$ (since they are dense in $\mathbb{R}$), so by using AC we can form a subset $E \subseteq (0,1)$ s.t. E contains exactly one element from each equivalence class.

Let $(r_n : n \in \mathbb{N})$ be an enumeration of $\mathbb{Q} \cap (-1,1)$ and let $E_n := E + r_n$. Then we have that:

(a) The sets E_n are disjoint. (clear!)

(b) $\bigcup_{n\in\mathbb{N}} E_n \subseteq (-1,2)$.

(c) $(0,1) \subseteq \bigcup_{n\in\mathbb{N}} E_n$.

To check (b), note that $E \subset (0,1)$ and every $r_n \in (-1,1)$. To see assertion (c), let $x \in (0,1)$ and let $e \in E$ be s.t. $x \sim e$. Then, by defn of $\sim$, $x - e \in \mathbb{Q}$ and must, as $x, e \in (0,1)$, belong to the interval $(-1,1)$. So, $x - e = r_n$ for some n, in other words $x \in \bigcup_{n\in\mathbb{N}} E_n$.

Now, BWOC, suppose that E is measurable. Then E_n must be measurable for all n, moreover, $m(E) = m(E_n)$, $n \in \mathbb{N}$. Then, additivity implies that

$$m\Big(\bigcup_{n\in\mathbb{N}} E_n\Big) = \sum_{n\in\mathbb{N}} m(E_n). \quad \Big(= \sum_{n\in\mathbb{N}} m(E)\Big).$$

So, if $m(E) = 0$, we have that $m((0,1)) = 0$ (not reasonable!)
(in fact, can pick E s.t. $\bigcup_{n\in\mathbb{N}} E_n$ contains as large an interval as we like.)

On the other hand, if $m(E) > 0$, we have that $m((-1,2)) = +\infty$. (again, not reasonable!)

Thus, E is not measurable.

NOTE: The main issue here is the translation invariance.

This is why we must restrict to a σ-algebra and not use the whole power set, if we want to define a measure which gives "good" values to each interval of $\mathbb{R}$ and is translation invariant.

OUTER MEASURES.

A standard technique for constructing measures is the following.

Definition 4: Let X be a set. Define $m^*: \mathcal{P}(X) \to [0, +\infty]$ to be a set function on $\mathcal{P}(X)$ satisfying:

(a) $m^*(\emptyset) = 0$.
(b) If $A \subset B \subset X$, then $m^*(A) \leq m^*(B)$.
(c) If $A_1, A_2, A_3, \ldots \subset X$, then $m^*(\bigcup_{n\in\mathbb{N}} A_n) \leq \sum_{n=1}^{\infty} m^*(A_n)$.

Examples:

1) Let X be a set and define m^* on $\mathcal{P}(X)$ by setting $m^*(A) := \begin{cases} 0 & \text{if } A = \emptyset \\ 1 & \text{otherwise} \end{cases}$.

2) Let X be an arbitrary set and define m^* on $\mathcal{P}(X)$ by setting

$$m^*(A) := \begin{cases} 0 & \text{if } A \text{ is countable;} \\ 1 & \text{if } A \text{ is uncountable.} \end{cases}$$

3) LEBESGUE OUTER MEASURE on $\mathbb{R}$.

For $A \subseteq \mathbb{R}$, let $\mathcal{C}_A$ be the collection of all infinite sequences $\{(a_i, b_i)\}_{i\in\mathbb{N}}$ s.t. $A \subset \bigcup_{i=1}^{\infty} (a_i, b_i)$. Then define $\lambda^*: \mathcal{P}(\mathbb{R}) \to [0, +\infty]$ by setting

$$\lambda^*(A) := \inf \left\{ \sum_{i=1}^{\infty} (b_i - a_i) : \{(a_i, b_i)\}_{i\in\mathbb{N}} \in \mathcal{C}_A \right\}.$$

FACTS:

(a) λ^* is an outer measure. (the only slightly tricky thing to check is property (c).)

(b) λ^* assigns the value $b-a$ to each interval (a,b), $[a,b]$, $(a,b]$, $[a,b)$.

(c) the Lebesgue outer measure of an unbounded interval is $+\infty$.

(Similar construction for $\mathbb{R}^n$, $n \geq 2$, defined on rectangles.)

MEASURABLE SETS.

Definition 5:
Let X be a set and m^* be an outer measure on X. Then a subset $B \subseteq X$ is said to be MEASURABLE with respect to m^* if

$$m^*(A) = m^*(A \cap B) + m^*(A \cap B^c) \quad \text{for every set } A \subseteq X.$$

(This criterion for measurability is due to Carathéodory originally.)

Notice that $m^*(A) \leq m^*(A \cap B) + m^*(A \cap B^c)$, by subadditivity, for all sets $A \in \mathcal{P}(X)$.
So, all that is required to be checked is that

$$m^*(A) \geq m^*(A \cap B) + m^*(A \cap B^c), \quad \text{for all } A \in \mathcal{P}(X) \text{ with } m^*(A) < +\infty.$$

It is a fact that if $m^*(B) = 0$ or $m^*(B^c) = 0$, then B is m^*-measurable.
(This is easy to see from the above inequality. S $m^*(B) = 0$. Then, NTS $m^*(A) \geq m^*(A \cap B^c)$, but this follows from property (b) of outer measure.)

Denote by $\mathcal{M}_{m^*}$ the set of all m^*-measurable sets. The following theorem is the fundamental fact about outer measures.

Theorem: The collection $\mathcal{M}_{m^*}$ is a σ-algebra and the restriction of m^* to $\mathcal{M}_{m^*}$ is a measure.

proof: See Theorem 1.3.4 in Cohn, Measure Theory □

LEBESGUE MEASURE.

This is the measure obtained by restricting Lebesgue outer measure to the collection $\mathcal{M}_{\lambda^*}$.

The first easy-to-verify fact about Lebesgue measure is that all Borel sets are contained in $\mathcal{M}_{\lambda^*}$. To see this, it suffices to check that every interval of the form $(-\infty, b]$ satisfies $\lambda^*(A) \geq \lambda^*(A \cap (-\infty, b]) + \lambda^*(A \cap (b, +\infty))$ for all A st. $\lambda^*(A) < +\infty$. Basically this is done by splitting each cover of A into two disjoint covers of $A \cap (-\infty, b]$ and $A \cap (b, +\infty)$.

Thus, in particular, every interval in $\mathbb{R}$ is Lebesgue measurable and is given the value of its length by λ. (We denote the restriction $\lambda^*|_{\mathcal{M}_{\lambda^*}}$ by λ.)

It is the case that there exist non-Borel yet Lebesgue measurable sets.

Proposition: Let $A \subseteq \mathbb{R}$ be Lebesgue measurable. Then

(a) $\lambda(A) = \inf\{\lambda(\mathcal{U}) : \mathcal{U} \text{ open and } A \subset \mathcal{U}\}$, and

(b) $\lambda(A) = \sup\{\lambda(K) : K \text{ compact and } K \subset A\}$.

$\equiv$ closed, bounded

In other words, λ is REGULAR. (This property is shared by all Borel probability measures on metric spaces.)

The most important (for us, for now) facts about Lebesgue measure are these:

- Lebesgue measure is TRANSLATION INVARIANT, i.e., where $A+x := \{y+x : y \in A\}$,

$$\lambda(A) = \lambda(A+x).$$

- If m is a not-identically-zero measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, that is translation invariant and finite on bounded subsets of $\mathbb{R}$, then there exists a constant $c>0$ s.t. $c\lambda(A) = m(A)$, for all $A \in \mathcal{B}(\mathbb{R})$.

"proof:" The first statement follows easily from the definition of λ^*. For the second, let $c = m([0,1)) < \infty$ (by assumption). Note that $c>0$, since otherwise $m(\mathbb{R})=0$, contradicting that m is not the zero measure.
Now define a measure $\nu(A) := \frac{1}{c} m(A)$, $\forall A \in \mathcal{B}(\mathbb{R})$. Then ν is translation invariant and gives $[0,1)$ the measure 1 (i.e., its length).
To finish the proof, it is necessary to show that ~~...~~ Lebesgue measure is the only measure that gives measure $1/2^k$ to every interval of length $1/2^k$ and then see that $[0,1)$ is the union of 2^k translates of such an interval D, so

$$2^k \nu(D) = \nu(C) = \lambda(C) = 2^k \lambda(D) \Rightarrow \nu = \lambda.$$

$\square$

So, Lebesgue measure is translation-invariant and gives "right" size to subintervals
$\Longrightarrow$ NOT ALL SETS CAN BE LEBESGUE MEASURABLE!

TWO USEFUL THEOREMS.

1. Let $(X, \mathcal{B}, m)$ be a measure space. Then

(a) If $(A_k)_{k=1}^{\infty}$ is an increasing sequence of sets from $\mathcal{B}$ (i.e. $A_i \subseteq A_{i+1}$ $\forall i \in \mathbb{N}$), then

$$m\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_{k\to\infty} m(A_k).$$

(b) If $(A_k)_{k=1}^{\infty}$ is a decreasing sequence of sets from $\mathcal{B}$ s.t. $m(A_n) < +\infty$ for some $n \in \mathbb{N}$, then

$$m\left(\bigcap_{k=1}^{\infty} A_k\right) = \lim_{k\to\infty} m(A_k).$$

2. BOREL-CANTELLI LEMMA! Let $A_1, A_2, A_3, \ldots \in \mathcal{B}$, where $(X, \mathcal{B}, m)$ is a probability space. Define $A_\infty := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \limsup_{n\to\infty} A_n = \{x \in X : x \in A_n \text{ for infinitely many } n\}$

Then, if $\sum_{n=1}^{\infty} m(A_n) < \infty$, we have $m(A_\infty) = 0$.

proof: exercise! It's really not too hard... $\square$