Integration Theory

1. The Riemann Integral. Given a bounded function

$$f : [0,1] \mapsto R^1$$ we can define its Riemann integral as follows:

First, we define a step function $S(x)$ to be a function on $[0,1]$ which is piecewise constant on intervals, i.e. there is a partition $x_0 = 0 < x_1 < x_2, \ldots < x_{n-1} < x_n = 1$ so that $S(x) = c_i$ for all $x \in (x_{i-1}, x_i)$ (for some real number c_i)

Next define the integral of such a step function $S(x)$ in the obvious way:

$$\int_0^1 S(x)\,dx \overset{\text{def}}{=} \sum_{i=1}^{n} c_i (x_i - x_{i-1})$$

Given f, define a step function $S(x)$ to be a lower step function for f provided

$$S(x) \leq f(x) \text{ for all } x \in [0,1]$$

and define an upper step function for f in the obvious way.

Then f is defined to be integrable iff there is exactly one number I so that

$$\int_0^1 s(x)\,dx \leqslant I \leqslant \int_0^1 t(x)\,dx$$

for all choices of lower step functions s and upper step functions t.

(Then, of course, we set $\int_0^1 f(x)\,dx = I$)

This way of defining the integral uses the fact that f can be well approximated by step functions. Roughly speaking, this means f has to be nearly constant over small intervals, i.e. f must be continuous "at most x's".

This theory of the integral has its limitations. For example if $f(x)$ is too rough, f will not be integrable. Consider

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}$$

then it is easy to show that if $s(x)$ is a lower step function for $\chi_{\mathbb{Q}}$ then all the values $c_i \leq 0$ and for upper step functions $t(x)$ all $c_i \geqslant 1$. Hence any number I between 0 and 1 satisfies $\int_0^1 s(x)\,dx \leq I \leq \int_0^1 t(x)\,dx$, and $\chi_{\mathbb{Q}}$ is not (Riemann) integrable.

Another limitation of Riemann integrals is that it is at best awkward to pass the limit under the integral sign. For instance, suppose we have a sequence $f_n(x)$ of Riemann integrable functions on $[0,1]$ and suppose that for all $x \in [0,1]$ $\lim_{n\to\infty} f_n(x) = f(x)$. We would like to have

$$\int_0^1 f(x)\,dx = \lim_{n\to\infty} \int_0^1 f_n(x)\,dx.$$

However $f(x)$ may not even have an integral!

(Example: Let $f_n(x) = \begin{cases} 1 & \text{if } x \text{ is a fraction w/ denominator} \leq n \\ 0 & \text{otherwise} \end{cases}$

On $[0,1]$ $\lim_{n\to\infty} f_n(x) = \chi_{\mathbb{Q}}(x)$, for all x)

2. The Lebesgue Integral

To remedy this H. Lebesgue created a new theory of the integral where Riemann step functions are replaced by so-called simple functions. Rather than insisting that simple functions take on their values on disjoint intervals, Lebesgue allowed simple functions to assume finitely many values, c_i, on disjoint <u>sets</u> E_i; so $s(x)$ is a (Lebesgue) simple function iff

$$s(x) = \sum c_i \chi_{E_i}, \quad E_i \text{ disjoint} \subseteq [0,1].$$

Lebesgue then defines the integral of a simple function as $\int_0^1 s(x)\,dx = \sum c_i\, m(E_i)$ where $m(E_i)$ is called the Lebesgue measure of the set E_i (intuitively its "length")

The advantage here is that in Lebesgue theory, simple functions can well approximate just about any function, ~~but~~ so we can Lebesgue integrate even very rough functions. However the theory is complicated by having to define $m(E)$ for E rather arbitrary.

What do we mean by the "length of a set", $m(E)$?

We want certain properties for m:

(1) $m(I) =$ length of I if I is an interval

(2) if $E = \bigcup_{i=1}^{\infty} E_i$, E_i disjoint, then

$$m(E) = \sum_{i=1}^{\infty} m(E_i)$$

(3) if $E + x$ denotes the translate of E by x, then $m(E+x) = m(E)$

Unfortunately, there is no m defined for all subsets $E \subseteq R^1$ that satisfies (1), (2) and (3).

The way around this is to measure by m "most sets", kicking out those that are too wild. The sets we keep are plentiful enough to make our theory work quite well. The sets we shall measure, known as "Lebesgue measurable sets" are closed under the usual operations applied to a sequence of sets $E_1, E_2, \ldots$

This means that if $\mathcal{L}$ denotes the Lebesgue measurable sets then (a) if $E_i \in \mathcal{L}$ for $i=1,2,3\ldots$ then $\bigcup_{i=1}^{\infty} E_i \in \mathcal{L}$

(b) if $E \in \mathcal{L}$ then ${}^cE \in \mathcal{L}$

and (c) $\emptyset \in \mathcal{L}$

A family of sets $\mathcal{F}$ satisfying (a), (b) and (c) above is called a σ algebra.

We can observe that the intersection of any collection of σ-algebras is also a σ-algebra. This implies that for any family of sets whatever $\mathcal{a}$ there is a smallest σ-algebra containing $\mathcal{a}$ (the intersection of all σ-algebras containing $\mathcal{a}$) This is denoted $\sigma(\mathcal{a})$ On R^1, the smallest σ-algebra containing the open intervals is called the Borel σ algebra or the Borel sets. Because intervals are Lebesgue measurable, the Borel sets are all Lebesgue measurable.

Now, given a function $f: R^1 \longrightarrow R^1$ there are certain subsets of the domain of f that are natural to consider; namely $\{x \mid f(x) > c\}$, $c \in R^1$.

We cannot hope to analyze functions unless these sets are all measurable (Borel or Lebesgue). We define a function f to be measurable iff

$$\{x \in R^1 \mid f(x) > c\} \text{ is measurable for every } c \in R^1.$$

Because

$$\{x \mid f(x) \geq c\} = \bigcap_{n=1}^{\infty} \{x \mid f(x) > c - \tfrac{1}{n}\}$$

it follows that $\{x \mid f(x) \geq c\}$ is measurable if f is measurable. Similar comments apply to $\{f(x) < c\} = {}^c\{f(x) \geq c\}$ and $\{x \mid f(x) \leq c\} = {}^c\{x \mid f(x) > c\}$.

Given a Lebesgue measurable function $f : R^1 \mapsto R^1$ (i.e. $\underset{\text{all } c}{\{x \mid f(x) > c\} \in \mathcal{L}}$) we introduce the concept of $\int_{R^1} f \, dm$ through simple functions. Because we also want to integrate over measurable subsets of R^1 and over such other spaces as R^n, $n > 1$, it is most convenient to abstractly give the theory, which then applies to all cases above, as well as many others. This is what we shall do below:

Let X be a set, $\mathcal{F}$ a σ-algebra of subsets of X and let μ assign a non-negative number to each set in $\mathcal{F}$, i.e. $\mu : \mathcal{F} \mapsto [0, \infty]$ so that

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i) \quad \text{whenever}$$

E_i are pairwise disjoint.

$(X, \mathcal{F}, \mu)$ is called a measure space and $\mathcal{F}$ is called the σ-algebra of measurable sets.

Example: $X = R^n$, $\mathcal{F}$ = the Borel sets, i.e. the smallest σ-algebra containing all cubes and $d\mu = m$ Lebesgue measure. This is the unique translation invariant measure on $\mathcal{F}$ which assigns to each cube its n dimensional volume.

Example: $X = \{1, 2, 3, \dots\}$ $\mathcal{F}$ = all subsets of X, μ = counting measure, i.e. $\mu(E)$ = # of elements of E

Definition: $f : X \longmapsto R^1$, $(X, \mathcal{F}, \mu)$ a measure space is $\mathcal{F}$-measurable provided $\{x \mid f(x) > c\} \in \mathcal{F}$ for each $c \in R^1$.

Definition: Let $s : X \longmapsto R^1$, where $(X, \mathcal{F}, \mu)$ is a measure space. $s(x)$ is said to be a simple function if it takes on finitely many values on measurable sets.

We can of course write $s = \sum_{i=1}^{n} c_i \chi_{E_i}$ where

$$\chi_{E_i}(x) = \begin{cases} 1 & \text{if } x \in E_i \\ 0 & \text{if } x \notin E_i \end{cases}$$

Definition: If s is a positive simple function then $\int_X s\,d\mu = \sum_{i=1}^{n} c_i\,\mu(E_i)$ (may be infinite)

Definition: Suppose $(X, \mathcal{F}, \mu)$ is a measure space and $f: X \mapsto [0,\infty]$ is measurable. Define

$$\int_X f\,d\mu = \sup\left\{\int_X s\,d\mu \;\middle|\; s \text{ is simple and } s(x) \le f(x) \text{ all } x \in X\right\}$$

(Note: when ∞ is an allowed value, f is measurable means $\{x \mid f(x) = \infty\}$ is in $\mathcal{F}$ and also $\{x \mid f(x) > c\}$ is in $\mathcal{F}$ for all $c \in R^1$)

The following results yield important information on measurable functions:

Theorem 1: If $f_n : X \mapsto R^1$ are all measurable, and $f(x) = \lim_{n\to\infty} f_n(x)$ exists for every $x \in X$, then f is also measurable.

Proof: First observe that $\sup_n f_n$ and $\inf_n f_n$ are measurable, since for example $\{\sup_n f_n > c\} = \bigcup_{n=1}^{\infty}\{f_n > c\}$

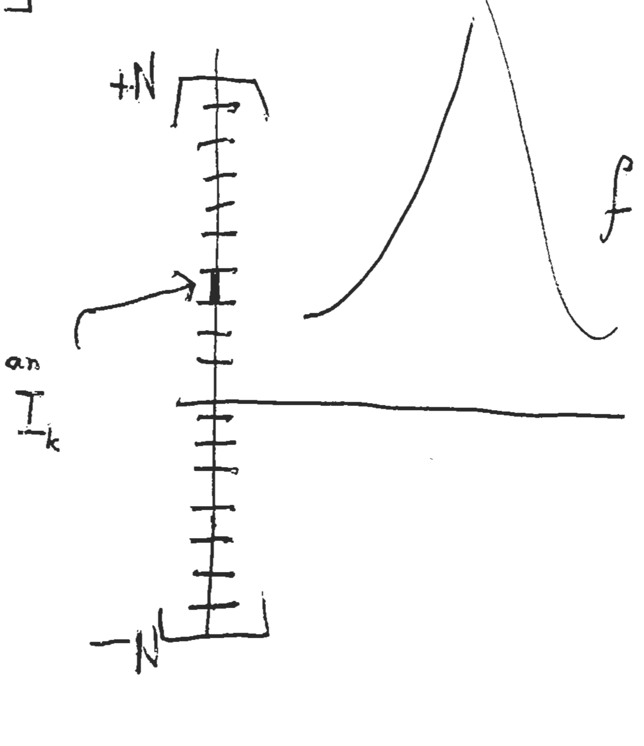
Let $g_N = \sup_{n \geq N} f_n$ then g_N is measurable for each N, and the theorem follows by observing that

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{N\to\infty} g_N(x) = \inf_N g_N \quad \text{and}$$

the last function is measurable, being an inf of a sequence of measurable functions.

Theorem 2: Any measurable $f: X \mapsto R^1$ is a pointwise limit of simple functions, i.e. there are simple functions S_n on X so that for each $x \in X$, $\lim_{n\to\infty} S_n(x) = f(x)$.

Proof: Fix N, a large positive integer, and split the interval $[-N, +N]$ into disjoint intervals of length 2^{-N}, and call these $\{I_k\}_k$



Now for those x so that $f(x) \in I_k$ set $s_N(x) =$ left endpoint of I_k. Do this for each k. Then on such sets of x's in the domain of f, $0 \leq f(x) - s_N(x) \leq \frac{1}{2^N}$. If $f(x) > N$, set $s_N(x) = N$ and if $s_N(x) < -N$, set $s_N(x) = -N$.

Then clearly $s_N(x) \longrightarrow f(x)$ as $N \longrightarrow \infty$.

Corollary 1: $f : X \mapsto R^1$ is measurable if and only if f is the pointwise limit of a sequence of simple functions.

Corollary 2: If $f, g : X \mapsto R^1$ are measurable then so are $f+g$, $f \cdot g$, $|f|$, and more generally $\varphi(f(x))$ where φ is continuous, $\varphi : R \mapsto R$.

Proof: If f, g are measurable, then choose s_n, t_n simple so that $s_n(x) \rightarrow f(x)$, $t_n(x) \longrightarrow g(x)$ as $n \rightarrow \infty$. Then $s_n + t_n$, $s_n \cdot t_n$, $|s_n|$ and $\varphi(s_n(x))$ are simple and converge pointwise to $f+g$, $f \cdot g$, $|f|$ and $\varphi(f(x))$ respectively. Hence all of these are measurable

One additional corollary will be useful very soon:

Corollary 3: If $f : X \mapsto [0, \infty]$ is measurable, then there exist positive simple functions S_n such that not only $\lim_{n\to\infty} S_n(x) = f(x)$ for each $x \in X$, but also

$$S_1(x) \leq S_2(x) \leq S_3(x) \cdots$$

Proof: The construction in Theorem 2 above yields this in case $f \geq 0$.

Now we shall consider an extremely important question.

We would like to have

$$(*) \int_X \lim_{n\to\infty} f_n(x)\, d\mu = \lim_{n\to\infty} \int_X f_n\, d\mu$$

whenever f_n are positive measurable functions. The following example shows this cannot be true in general:

Example: Let $X = [0,1]$ and ~~d~~ $\mu = m$, Lebesgue measure.

Let $f_n(x) = n\chi_{(0,\frac{1}{n})}$. Then $\lim_{n\to\infty} f_n(x) = 0$ for all $x \in [0,1]$

but $\int_0^1 f_n\, dm = 1$, for all n.

In case the f_n satisfy $f_1 \le f_2 \le f_3 \ldots$

then we do have $(*)$.

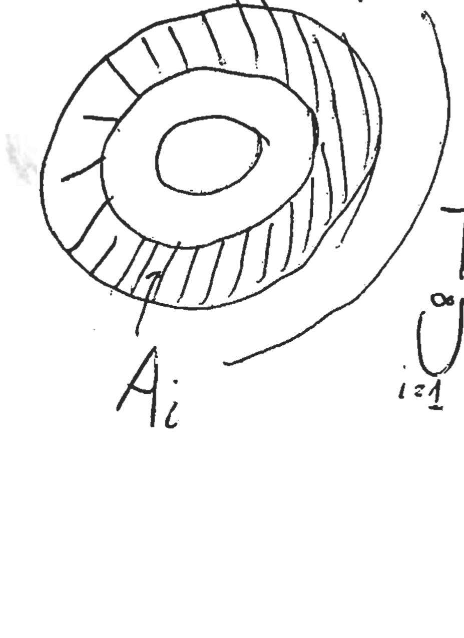
The Monotone Convergence Theorem: Let $f_1(x) \le f_2(x) \le f_3(x) \ldots$ be a sequence of positive measurable functions on a measure space $(X, \mathcal{F}, \mu)$. Let $f(x) = \lim_{n\to\infty} f_n(x)$, $x \in X$ (may be $+\infty$ for some x)

Then $$\int_X f \, d\mu = \lim_{n\to\infty} \int_X f_n \, d\mu .$$

In order to prove this, we require a few simple preliminary results. The first is a simple special case of the Monotone Convergence Thm:

Lemma 1: Suppose $E_1 \subseteq E_2 \subseteq E_3 \ldots$ is a sequence of sets in $\mathcal{F}$. Then $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} \mu(E_n)$

Proof:



Set $A_i = E_i - E_{i-1}$ $(i \ge 2)$ ($A_1 \overset{\text{defin}}{=} E$

Then the A_i are disjoint and $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} E_i$.

Then $$\mu\left(\bigcup_1^\infty E_i\right) = \mu\left(\bigcup_1^\infty A_i\right) = \sum_1^\infty \mu(A_i)$$
$$= \lim_{n\to\infty} \sum_1^n \mu(A_i) = \lim_{n\to\infty} \mu(E_n).$$

It will be convenient to make the fallowing:

Definition: Let $f : X \longmapsto [0,\infty]$ be a measurable function and $E \in \mathcal{F}$. Then we define $\int_E f\,d\mu$ by
$$\int_E f\,d\mu = \int_X \chi_E f\,d\mu .$$

Lemma 2: Suppose s is a positive simple function on X, and set
$$\lambda(E) = \int_E s\,d\mu \quad \text{for each } E \in \mathcal{F}.$$ Then λ is a measure on $\mathcal{F}$

Proof: We need to show that if $E_i \in \mathcal{F}$ are disjoint, then
$$\lambda\left(\bigcup_1^\infty E_i\right) = \sum_1^\infty \lambda(E_i) .$$
Let $s = \sum c_k \chi_{F_k}$, $c_k \neq 0$ Then

$$\cancel{\lambda\left(\bigcup_1^\infty E_i\right) = \int_{\bigcup_1^\infty E_i} \sum c_k \chi_{F_k}\,d\mu = \int_X \chi_{\cup E_i} \sum c_k \chi_{F_k}\,d\mu}$$

$$\lambda(E) = \int_E s\, d\mu = \int_X s \chi_E\, d\mu$$

$$= \int_X \sum c_k \chi_{F_k \cap E}\, d\mu =$$

$$\sum c_k \mu(E \cap F_k).$$

We need only observe that for each k

$\lambda_k(E) = \mu(E \cap F_k)$ is a measure since

$$\lambda_k(\cup E_i) = \mu((\cup E_i) \cap F_k) = \mu(\bigcup_i [E_i \cap F_k])$$

$$= \sum_i \mu(E_i \cap F_k) = \sum \lambda_k(E_i)$$

and λ is just a finite sum of multiples of the λ_k.

Proof of the Monotone Convergence Theorem:

Since $f_1 \le f_2 \ldots$, it is clear that $f_n(x) \le f(x)$ for all x, so that

$$\int_X f_n\, d\mu \le \int_X f\, d\mu \text{ for all } n \text{ and}$$

$$\lim_{n\to\infty} \int_X f_n\, d\mu \le \int_X f\, d\mu.$$ We show now that

$$\lim_{n\to\infty} \int_X f_n\, d\mu \ge \int_X f\, d\mu$$ which will prove the theorem.

Let $\varepsilon > 0$ and let s be a positive simple function so that $s(x) \leq f(x)$ for all $x \in X$. We need only show

$$(\ddagger) \qquad \lim_{n\to\infty} \int_X f_n \, d\mu \geq (1-\varepsilon) \int_X s \, d\mu \text{ ; then by}$$

the definition of $\int_X f \, d\mu$ as $\sup_{s \leq f} \int_X s \, d\mu$ we shall have proven that $\lim_{n\to\infty} \int_X f_n \, d\mu \geq (1-\varepsilon) \int_X f \, d\mu$ for all $\varepsilon > 0$.

How do we show $(\ddagger)$? Let

$$E_n = \{ x \mid f_n(x) \geq (1-\varepsilon) s(x) \}.$$ Then

$$X = \bigcup_{n=1}^{\infty} E_n \quad \text{since } f_n(x) \longrightarrow f(x) \geq s(x) \text{ for all } x \in X,$$

and $E_1 \subseteq E_2 \subseteq E_3 \ldots$ since $f_1 \leq f_2 \leq f_3 \ldots$

Then

$$\int_X f_n \, d\mu \geq \int_{E_n} f_n \, d\mu \geq \int_{E_n} (1-\varepsilon) s(x) \, d\mu = (1-\varepsilon) \int_{E_n} s \, d\mu .$$

By Lemmas 1 and 2 above $\int_{E_n} s \, d\mu = \lambda(E_n) \to \lambda(\cup E_n) = \lambda(X) = \int_X s \, d\mu$.

This shows

$$\lim_{n\to\infty}\int_X f_n \, d\mu \geqslant (1-\varepsilon)\int_X s \, d\mu$$ and completes

the proof of (†) and the theorem.

Corollary 1: Let f and g be positive measurable functions on X, and $c>0$. Then
$$\int_X (f+g)\, d\mu = \int_X f\,d\mu + \int_X g\, d\mu.$$

and
$$\int_X (cf)\, d\mu = c\int_X f\, d\mu.$$

Proof: We know there is a sequence $s_1 \leq s_2 \leq s_3 \ldots$ of positive simple functions converging pointwise to f and a similar sequence t_n for g. Then $s_n+t_n \to f+g$ and $s_n+t_n \leq s_{n+1}+t_{n+1}$

so
$$\int_X (f+g)\, d\mu = \lim_{n\to\infty}\int_X (s_n+t_n)\, d\mu = \lim_{n\to\infty}\int_X s_n\, d\mu + \lim_{n\to\infty}\int_X t_n\, d\mu$$

$$= \int_X f\,d\mu + \int_X g\, d\mu$$

Here we have used Monotone Convergence, and the fact that for simple functions s and t $\int_X (s+t)\, d\mu = \int_X s\, d\mu + \int_X t\, d\mu$ which is easy to prove.

The claim about $\int cf\,d\mu = c\int f\,d\mu$ is proven similarly.

Corollary 2: Let f_n be positive measurable functions on X. Then

$$\int_X \left(\sum_{n=1}^{\infty} f_n\right) d\mu = \sum_{n=1}^{\infty} \int_X f_n\,d\mu .$$

Proof: Let $g_N = \sum_{n=1}^{N} f_n$; then $g_1 \leq g_2 \leq g_3 \cdots$

so by Monotone Convergence,

$$\int_X \sum_1^{\infty} f_n\,d\mu = \int_X \lim_{N\to\infty} g_N\,d\mu \overset{MCT}{=} \lim_{N\to\infty} \int_X g_N\,d\mu$$

$$\overset{(\text{Corollary } 1)}{=} \lim_{N\to\infty} \sum_{n=1}^{N} \int_X f_n\,d\mu$$

$$= \sum_{n=1}^{\infty} \int_X f_n\,d\mu .$$

Corollary 3: Let f be a positive measurable function and set $\lambda(E) = \int_E f\,d\mu$. Then λ is a measure.

Proof. Let $E_k \in \mathcal{F}$ be disjoint.

Then
$$\lambda(\cup E_k) = \int_{\cup E_k} f \, d\mu = \int_X \chi_{\cup E_k} \cdot f \, d\mu$$
$$= \int_X \sum \chi_{E_k} f \, d\mu = \sum \int_X \chi_{E_k} f \, d\mu$$
$$= \sum \int_{E_k} f \, d\mu = \sum \lambda(E_k)$$

Now let us briefly discuss integration of measurable functions $f : X \mapsto R^1$.

Suppose $f : X \mapsto R^1$ is measurable and further assume that $\int_X |f| \, d\mu < \infty$. We then say that f is integrable on X with respect to μ and define

$$\int_X f \, d\mu = \int_X f^+ d\mu - \int_X f^- d\mu$$ where

$$f^{+}(x) = \max\big(f(x), 0\big) \quad \text{and}$$

$$f^{-}(x) = -\min\big(f(x), 0\big)$$

It is easy to verify that

$$f(x) = f^{+}(x) - f^{-}(x), \quad x \in X.$$

By using linearity of the integral for positive f we can prove that $\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$ for f, g integrable. In fact

$$f^{+} - f^{-} + g^{+} - g^{-} = f + g = (f+g)^{+} - (f+g)^{-}$$

so

$$f^{+} + g^{+} + (f+g)^{-} = f^{-} + g^{-} + (f+g)^{+}$$

Integrating both sides with respect to μ and using the additivity of integrals for positive functions,

$$\int_X f^{+}\,d\mu + \int_X g^{+}\,d\mu + \int_X (f+g)^{-}\,d\mu = \int_X f^{-}\,d\mu + \int_X g^{-}\,d\mu + \int_X (f+g)^{+}\,d\mu$$

and so

$$\left(\int_X f^{+}\,d\mu - \int_X f^{-}\,d\mu\right) + \left(\int_X g^{+}\,d\mu - \int_X g^{-}\,d\mu\right) = \int_X (f+g)^{+}\,d\mu - \int_X (f+g)^{-}\,d\mu$$

or $$\int_X f\,d\mu + \int_X g\,d\mu = \int_X (f+g)\,d\mu.$$

Similarly we can easily show that

$$\int_X cf\,d\mu = c\int_X f\,d\mu \text{ for any } c\in R^1$$

and integrable f on X.

Though we shall not prove it here, ~~the~~ we state an extremely important result on passing the limit under the integral sign when the functions are not necessarily positive. It turns out that the proof follows in a rather easy way from the Monotone Convergence Theorem:

Dominated Convergence Theorem: Let $f_1(x), f_2(x), \ldots f_n(x), \ldots$ be integrable functions on X, and suppose

$$f(x) = \lim_{n\to\infty} f_n(x), \quad x\in X.$$

Suppose there exists $\Phi(x) \geq 0$ which is integrable so that $|f_n(x)| \leq \Phi(x)$ for all $x\in X$ and each n.

Then $$\int_X f\,d\mu = \lim_{n\to\infty} \int_X f_n\,d\mu.$$