1 Differentiation

1.1 Derivatives of Measures

Theorem 1. Suppose μ is a complex Borel measure on $\mathbb{R}^1$ and

$$f(x) = \mu((-\infty, x)) \text{ for } x \in \mathbb{R}^1 \tag{1}$$

If $x \in \mathbb{R}^1$ and A is a complex number, TFAE

(a) f is differentiable at x and $f'(x) = A$.

(b) For all $\epsilon > 0$, there exists $\delta > 0$ such that

$$|\frac{\mu(I)}{m(I)} - A| < \epsilon \tag{2}$$

for every open segment I that contains x and whose length is less than δ. Note m is the Lebesgue measure on $\mathbb{R}^1$.

Definition 1. Let us fix a dimension k, denote the open ball with center $x \in \mathbb{R}^k$ and radius $r > 0$ by

$$B(x, r) = \{y \in \mathbb{R}^k \colon |y - x| < r\}$$

We associate to any Borel measure μ on $\mathbb{R}^k$ the quotients:

$$(Q_r\mu)(x) = \frac{\mu(B(x,r))}{m(B(x,r))}$$

Where m is the Lebesgue measure on R^k.

We define the **symmetric derivative** to be

$$(D\mu)(x) = \lim_{r \to 0}(Q_r\mu)(x)$$

Definition 2. Using the same notation as above we define the **maximal function** $M\mu$, for $\mu \geq 0$, to be defined by

$$(M\mu)(x) = \sup_{0<r<\infty}(Q_r\mu)(x)$$

Remark. The maximal function of a complex Borel measure μ is, by definition, its total variation $|\mu|$.

The functions $M\mu \colon R^k \to [0, \infty]$ are lower semicontinuous, hence measurable.

Proof. Assume $\mu \geq 0$, and let $\lambda > 0$ and $E = \{M\mu > \lambda\}$. Fix $x \in E$. Then there is an $r > 0$ such that:

$$\mu(B(x, r)) = tm(B(x, r)) \text{ for some } t > \lambda$$

Indeed since $\sup_{0<r<\infty} \frac{\mu(B(x,r))}{m(B(x,r))} > \lambda$. So for some r, we have $\frac{\mu(B(x,r))}{m(B(x,r))} > \lambda$. Letting $t = \frac{\mu(B(x,r))}{m(B(x,r))}$ gives us the desired result.

Furthermore there is a $\delta > 0$ such that:

$$(r + \delta)^k < \frac{r^k t}{\lambda}$$

If $|y - x| < \delta$, then $B(y, r + \delta) \supseteq B(x, r)$. Therefore

$$\mu(B(y, r + \delta)) \geq \mu(B(x, r)) = tm(B(x, r)) = t[\frac{r}{(r + \delta)^k} m(B(y, r + \delta)) > \lambda m(B(y, r + \delta))]$$

Thus $B(x, \delta) \subseteq E$. So E is open. □

Lemma 1. If W is the union of a finite collection of balls $B(x_i, r_i)$, with $i \leq i \leq N$. Then there is a set $S \subseteq \{1, \dots, N\}$ so that:

(a) the balls $B(x_i, r_i)$ with $i \in S$ are disjoint,

(b)
$$W \subseteq \bigcup_{i \in S} B(x_i, 3r_i),$$

(c)
$$m(W) \leq 3^k \sum_{i \in S} m(B(x_i, r)i).$$

Proof. Order the balls $B_i = B(x_i, r_i)$ such that $r_1 \geq r_2 \geq \cdots \geq r_N$. Put $i_1 = 1$, discard all the B_j that intersect with B_{i_1}. Let B_{i_2} the first of our remaining balls, and discard all B_j with $j > i_2$ that intersect B_{i_2}, and let B_{i_3} be the first of the remaining ones, etc...

This process stops after a finite number of steps, since we only have a finite collection of balls, and we let $S = \{i_1, i_2, \dots\}$. (a) holds by definition and (c) follows from (b) since $m(B(x_i, 3r_i)) = 3^k m(B(x_i, r_i))$.

So we just need to show (b). But notice for every discarded B_j, $B_j \cap B_i \neq \emptyset$ for some $i \in S$, where $r_i > r_j$. Assume that $X \in B_j \cap B_i$. We see that for all $x \in B_j$ we have:

$$\begin{aligned}
|x - x_i| &\leq |x - X| + |X - x_i| \\
&\leq |x - x_j| + |x_j - X| + |X - x_i| \\
&< r_j + r_j + r_i \text{ since } x, X \in B_j \text{ and } X \in B_i \\
&< 3r_i \text{ since } r_j \leq r_i
\end{aligned}$$

So we see that $B_j \subseteq B(x_i, 3r_i)$. This gives us (b). □

The maximal theorem

Theorem 2. If μ is a complex Borel measure on $\mathbb{R}^k$ and λ is a positive number, then

$$m\{M\mu > \lambda\} \leq 3^k \lambda^{-1} ||\mu|| \tag{i}$$

Here $||\mu|| = |\mu|(\mathbb{R}^k)$ and $m\{M\mu > \lambda\}$ is an abbreviation of $m(\{x \in \mathbb{R}^k : (M\mu)(x) > \lambda\})$

Proof. Fix μ and λ. Let K be a compact subset of the open set $\{M\mu > \lambda\}$. Each $x \in K$ is the center of an open ball B for which

$$|\mu|(B) > \lambda m(B)$$

Some finite collection of these B's covers K and Lemma 1 tells us there is a disjoint subcollection $\{B_1, \dots, B_n\}$ such that:

$$m(K) \leq 3^k \sum_1^n m(B_i) \leq 3^k \lambda^{-1} \sum_1^n |\mu|(B_i) \leq 3^l \lambda^{-1} ||\mu||$$

The disjointess of the B_i's was used in the last inequality. So (i) follows by taking the supremum over all compact $K \subseteq \{M\mu > \lambda\}$. □

Weak L^1 If $f \in L^1(\mathbb{R}^k)$ and $\lambda > 0$, then

$$m\{|f| > \lambda\} \leq \lambda^{-1} ||f||_1$$

because, if we let $E = \{|f| > \lambda\}$, we have:

$$\lambda m(E) \leq \int_R |f| dm \leq \int_{\mathbb{R}^k} |f| dm = ||f||_1$$

Definition 3. Any measurable function f for which:

$$\lambda m\{|f| > \lambda\}$$

is a bounnded funtion of λ on $(0,\infty)$ is said to belong to **weak** L^1

So from above we see that the weak L^1 contains L^1. But it is also larger since for example if we let $f = \frac{1}{x}$ on $(0,1)$, then for any $\lambda > 0$, we have

$$\frac{1}{x} > \lambda \iff x < \frac{1}{\lambda}$$

So we have $\lambda \cdot m\{|f| > \lambda\} \leq \lambda \cdot m(0, \frac{1}{\lambda}) = 1 < \infty$. So $\frac{1}{x}$ is weak L^1.

Definition 4. We associate to each $f \in L^1(\mathbb{R}^k)$ its **maximal function** $Mf\colon \mathbb{R}^k \to [0,\infty]$ by setting

$$(Mf)(x) = \sup_{0<r<\infty} \frac{1}{m(B_r)} \int_{B(x,r)} |f|\ dm$$

If we identify f with the measure μ given by $d\mu = f\ dm$, we see that this defintion agrees with the previously defined $M\mu$. So theorem 2 states that the "maximal operator" M sends L^1 to weak L^1, witha bound (namely 3^k) that depends only on the space $\mathbb{R}^k$, i.e: For every $f \in L^1(\mathbb{R}^k)$ and every $\lambda > 0$

$$m\{Mf > \lambda\} \leq 3^k \lambda^{-1} ||f||_1$$

Lebesgue points

Definition 5. If $f \in L^1(\mathbb{R}^k)$, any $x \in \mathbb{R}^k$ for which it is true that

$$\lim_{r\to 0} \frac{1}{m(B_r)} \int_{B(x,r)} |f(y) - f(x)|\ dm(y) = 0$$

is called a **Lebesgue point** of f.

For example this equation holds if f is continuous at the point x. More generally this equation holds, if the averages of $|f - f(x)|$ are not too small on the balls centered at x, i.e. The Lebesgue points of f are thepoints where f doesn't oscillate too much.

Theorem 3. If $f \in L^1(R^k)$, then almost every $x \in \mathbb{R}^k$ is a Lebesgue point of f.

Proof. Let

$$(T, f)(x) = \frac{1}{m(B_r)} \int_{B(x,r)} |f - f(x)| dm \text{ for } x \in \mathbb{R}^k, r > 0$$

Put

$$(Tf)(x) = \limsup_{r\to 0} (T_r f)(x)$$

Pick $y > 0$, let $n \in \mathbb{N}^*$. By a theorem from chap 3, there exists $g \in C(\mathbb{R}^k)$ so that $||f - g||_1 < \frac{1}{n}$. Let $h = f - g$. Since g is continuous, $Tg = 0$, and since:

$$\begin{aligned}(T_r h)(x) &= \frac{1}{B_r} \int_{B(x,r)} |h - h(x)| dm \\ &\leq \frac{1}{B_r} \int_{B(x,r)} (|h| + |h(x)|) dm \\ &= (\frac{1}{B_r} \int_{B(x,r)} |h| dm) + |h(x)|\end{aligned}$$

So we have:

$$Th \leq Mh + |h|$$

But since $T_r f \leq T_r g + T_r h$ it follows that

$$Tf \leq Mh + |h|$$

Therefore

$$\{Tf > 2y\} \subseteq \underbrace{\{Mh > y\} \cup \{|h| > y\}}_{E(y,n)}$$

Since $||h||_1 < \frac{1}{n}$, by theorem 2 we can see that

$$m(E(y,n)) \leq \frac{3^k + 1}{yn}$$

Note $\{Tf > 2y\}$ is independant of n. Hence

$$\{Tf > 2y\} \subseteq \bigcap_{n=1}^{\infty} E(y,n)$$

This intersection has measure zero, so $\{Tf > 2y\}$ is a subset of a set of measure zero. So since Lebesgue measure is complete $\{Tf > 2y\}$ is measurable and has measure zero. This is true for all $y > 0$ so $Tf = 0$ a.e.

So note if $(Tf)(x) = \limsup_{r\to 0}(T_r f)(x) = 0$, then since $(T_r f)(x) \geq 0$ we see that this means that $0 \leq \liminf(T_r f)(x) \leq \limsup(T_r f)(x) = 0$.

So we have $\lim_{r\to 0}(T_r f)$ exists and is equal to zero, so x is a Lebesgue point. So almost every point $x \in \mathbb{R}^k$ is a Lebesgue point of f. □

Definition 6. Recall that by the Radon-Nikodym theorem if μ is a positive σ-finite measure on a σ-algebra $\mathcal{M}$ in a set X, and λ is a complex measure on $\mathcal{M}$ such that $\lambda \ll \mu$:

$$\lambda(E) = \int_E f d\mu$$

For some $f \in L^1(\mu)$

f is called the **Radon-Nikodym derivative**of μ with respect to m and is denoted

$$f = \frac{d\lambda}{d\mu}$$

Theorem 4. Suppose μ is a complex Borel measure on $\mathbb{R}^k$, and $\mu \ll m$. Let f be the Radon-Nikodym derivative of μ with respect to m. Then $D\mu = f$ a.e. $[m]$, and

$$\mu(E) = \int_E (D\mu)\ dm$$

for all Borel sets $E \subseteq \mathbb{R}^k$.

Proof.

$$\mu(E) = \int_E f\ dm$$

For all Borel sets $E \subseteq \mathbb{R}^k$.

Let x be a Lebesgue point and $\Gamma_r = \frac{1}{B_r}\int_{B(x,r)} f\ dm$. Then we have:

$$0 \leq |\Gamma_r - f(x)| = |\frac{1}{B_r}\int_{B(x,r)} (f - f(x))\ dm| \leq \frac{1}{B_r}\int_{B(x,r)} |f - f(x)|\ dm$$

Taking limits we see that

$$f(x) = \lim_{r\to 0}\frac{1}{m(B_r)}\underbrace{\int_{B(x,r)} f\ dm}_{\mu(B(x,r))} = \lim_{r\to 0}\frac{\mu(B(x,r))}{m(B_r)} = (D\mu)(x)$$

Thus $(D\mu)(x)$ exists and equals to $f(x)$ at every Lebesgue point of f, so a.e. □

Nicely shirinking sets

Definition 7. Suppose $x \in \mathbb{R}^k$. A sequence $\{E_i\}$ of Borel sets in $\mathbb{R}^k$ is said to **shrink to x nicely** if there is a number $\alpha > 0$ with the following property:

There is a sequence of balls $B(x, r_i)$ with $\lim r_i = 0$, such that $E_i \subseteq B(x, r_i)$ and:

$$m(E_i) \geq \alpha m(B(x, r_i)) \text{ for } i = 1, 2, 3, \ldots$$

Theorem 5. Associate to each $x \in \mathbb{R}^k$ a sequence $\{E_i(x)\}$ that shrinks to x nicely, and let $f \in L^1(\mathbb{R}^k)$. Then

$$f(x) = \lim_{i \to \infty} \frac{1}{m(E_i(x))} \int_{E_i(x)} f \, dm$$

At every Lebesgue point of f.

Proof. Let x be a Lebesgue point of f and let $\alpha(x)$ and $B(x, r_i)$ be the positive number and the balls associate with $\{E_i(x)\}$. Since $E_i(x) \subseteq B(x, r_i)$ we have:

$$\int_{E_i(x)} |f - f(x)| \, dm \leq \int_{B(x, r_i)} |f - f(x)| \, dm$$

Furthermore, $\alpha m(B(x, r_i)) \leq m(E_i) \iff \frac{\alpha}{m(E_i)} \leq \frac{1}{m(B(x, r_i))}$. Putting this all together we get:

$$\frac{\alpha}{m(E_i)} \int_{E_i(x)} |f - f(x)| \, dm \leq \frac{1}{m(B(x, r_i))} \int_{B(x, r_i)} |f - f(x)| \, dm$$

Since x is a Lebesgue point RHS converges to 0, so the LHS also converges to zero by squeeze. □

Corollary. If $f \in L^1(\mathbb{R}^1)$ and

$$F(x) = \int_{-\infty}^{x} f \, dm, \text{ for } x \in \mathbb{R}$$

then $F'(x) = f(x)$ at every Lebesgue point of f.

Proof. Let x be a Lebesgue point, and $\{\delta_i\}$ be a sequence of positive numbers that converges to 0. Letting $E_i(x) = [x, x + \delta_i]$, the previous theorem tells us that

$$f(x) = \lim_{i \to \infty} \frac{1}{\delta_i} \int_{x}^{x+\delta_i} f \, dm = \lim_{i \to \infty} \frac{1}{\delta_i} \left(\int_{-\infty}^{x+\delta_i} f \, dm - \int_{-\infty}^{x} f \, dm \right) = \lim_{i \to \infty} \frac{F(x + \delta_i) - F(x)}{\delta_i}$$

Since $\{\delta_i\}$ is any sequence of positive numbers converging to zero we have:

$$f(x) = \lim_{h \to 0^+} \frac{F(x + h) - F(x)}{h}$$

Likewise letting $G_i(x) = [x - \delta_i, x]$ we get

$$f(x) = \lim_{i \to \infty} \frac{F(x - \delta_i) - F(x)}{\delta_i} = \lim_{h \to 0^-} \frac{F(x + h) - F(x)}{h}$$

So we have:

$$f(x) = \lim_{h \to 0} \frac{F(x + h) - F(x)}{h} = F'(x) \text{ at every Lebesgue point of } f$$

□

Metric density

Definition 8. Let E be a Lebesgue measurable subset of $\mathbb{R}^k$. The **metric density** of E at a point $x \in \mathbb{R}^k$ is defined to be

$$\lim_{r \to 0} \frac{m(E \cap B(x, r))}{m(B(x, r))} \text{ if this limit exists.}$$

If we let f be the characteristic function of E, and apply Theorem 1, we see that the metric density of E is 1 at almost every point of E and is 0 at almost every point of E^c.

Indeed let x be a Lebesgue point if $\mu(B(x, r)) = \int_{B(x,r)} f \, dm = m(E \cap B(x, r))$, it is clear that $\mu \ll m$ and so we have:

$$\lim_{r\to 0}\frac{m(E\cap B(x,r))}{m(B(x,r))}=\lim_{r\to 0}\frac{\mu(B(x,r))}{m(B(x,r))}=(D\mu)(x)=f(x)=\begin{cases}1 \text{ if } x\in E\\ 0 \text{ if } x\notin E\end{cases}$$

Corollary. If $\epsilon>0$, there is no set $E\subseteq\mathbb{R}^1$ such that

$$\epsilon<\frac{m(E\cap I)}{m(I)}<1-\epsilon$$

For every segment I.

Proof. Let $\epsilon>0$ assume that that such a $E\subseteq\mathbb{R}^1$ exists. Let x be a Lebesgue point, from what we have seen in the definition of metric density we know that there is a R such that:

$$|\frac{m(E\cap(x-R,x+R))}{m(x-R,x+R)}-0|<\epsilon \text{ if } x\notin E$$
$$|\frac{m(E\cap(x-R,x+R))}{m(x-R,x+R)}-1|<\epsilon \text{ if } x\in E$$

I.e.

$$\frac{m(E\cap(x-R,x+R))}{m(x-R,x+R)}<\epsilon \text{ or } 1-\epsilon<\frac{m(E\cap(x-R,x+R))}{m(x-R,x+R)}$$

□

We now look at differentiation of measures that are singular wrt m.

Theorem 6. Associate to each $x\in\mathbb{R}^k$ a sequence $\{E_i(x)\}$ that shrinks to x nicely. If μ is a complex Borel measure and $\mu\perp m$, then

$$\lim_{i\to\infty}\frac{\mu(E_i(x))}{m(E_i(x))}=0 \text{ a.e. } [m]$$

Proof. By the Jordan decomp theorem we just need to show that this is true with $\mu\geq 0$. In that case as we have seen in previous theorems:

$$\frac{\alpha(x)\mu(E_i(x))}{m(E_i(x))}\leq\frac{\mu(E_i(x))}{m(B(x,r_i))}\leq\frac{\mu(B(x,r_i))}{m(B(x,r_i))}$$

So if we can show that $(D\mu)(x)=0$ a.e. [m], we will prove the result by taking limits in the above inequality.

The upper derivative $\bar{D}\mu$ is defined by:

$$(\bar{D}\mu)(x)=\lim_{n\to\infty}\left[\sup_{0<r<1/n}(Q_r\mu)(x)\right] \text{ for } x\in\mathbb{R}^k$$

Is a Borel function.

Choose $\lambda>0$ and $\epsilon>0$. Since $\mu\perp m$, μ is concentrated on a set of Lebesgue measure 0. The regularity of μ shows that there is a compact set Km with $m(K)=0$, and $\mu(K)>||\mu||-\epsilon$.

Define $\mu_1(E)=\mu(K\cap E)$, for any Borel set $E\subseteq\mathbb{R}^k$, and put $\mu_2=\mu-\mu_1$. Then $||\mu_2||<\epsilon$, and for every x outside K,

$$(\bar{D}(\mu))(x)=(\bar{D}(\mu_2))(x)\leq(M\mu_2)(x).$$

Hence

$$\{\bar{D}\mu>\lambda\}\subseteq K\cup\{M\mu_2>\lambda\},$$

And

$$m\{\bar{D}\mu>\lambda\}\leq 3^k\lambda^{-1}||\mu_2||<3^k\lambda^{-1}\epsilon$$

Since this holds for all $\epsilon>0$ and $\lambda>0$, we find that $\bar{D}\mu=0$ a.e. [m], so

$$(D\mu)(x)=0\ a.e.[m]$$

Which gives us our result. $\square$

Corollary. Suppose that to each $x \in \mathbb{R}^k$ is associated to some sequence $\{E_i(x)\}$ that shrinks to x nicely, and that μ is a complex Borel measure on $\mathbb{R}^k$. Let $d\mu = f\ dm + d\mu_s$ be the Lebesgue decomposition of μ wrt m. Then

$$\lim_{i\to\infty} \frac{\mu(E_i(x))}{m(E_i(x))} = f(x)\ a.e.[m]$$

In particular, $\mu \perp m$ if and only if $(D\mu)(x) = 0$ a.e. [m]

Proof. Let $\mu_a(E) = \int_E f\ dm$, then recall that $\mu = \mu_a + \mu_s$, and

$$\begin{cases} \mu_a \ll m \\ \mu_s \perp m \end{cases}$$

Then from theorem 1

$$\lim_{i\to\infty} \frac{\mu_a(E_i(x))}{m(E_i(x))} = f(x)\ a.e.[m]$$

On the other hand from theorem 6

$$\lim_{i\to\infty} \frac{\mu_s(E_i(x))}{m(E_i(x))} = 0\ a.e.[m]$$

So we have

$$\lim_{i\to\infty} \frac{\mu(E_i(x))}{m(E_i(x))} = \lim_{i\to\infty} \frac{\mu_a(E_i(x)) + \mu_s(E_i(x))}{m(E_i(x))} = \lim_{i\to\infty} \frac{\mu_a(E_i(x))}{m(E_i(x))} + \lim_{i\to\infty} \frac{\mu_s(E_i(x))}{m(E_i(x))} = f(x)\ a.e.[m]$$

(Indeed if we let A, B be the sets where the first two equations fail, then our result fails on $A \cup B$, which is the union of two measure zero sets, and so is a measure zero set, so this equality is true almost everywhere). $\square$

Theorem 7. If μ is a positive Borel measure on $\mathbb{R}^k$ and $\mu \perp m$, then

$$(D\mu)(x) = \infty\ a.e.[\mu] \tag{$\dagger$}$$

Proof. There is a Borel set $S \subseteq \mathbb{R}^k$ with $m(S) = 0$ and $\mu(\mathbb{R}^k \setminus S) = 0$, and there are open sets $V_j \supseteq S$ with $m(V_j) < \frac{1}{j}$, for $j = 1, 2, 3, \dots$

For $N = 1, 2, 3, \dots$, let E_N be the set of all $x \in S$ to which correspond radii $r_i = r_i(x)$, with $\lim r_i = 0$ such that

$$\mu(B(x, r_i)) < Nm(B(x, r_i)). \tag{$\dagger\dagger$}$$

Then ($\dagger$) holds for all $s \in S \setminus \bigcup_N E_N$.

Fix N and j, for the moment. Every $x \in E_N$ is in the center of a ball $B_x \subseteq V_j$, that satisfies ($\dagger\dagger$). Let β_x be the open ball with center x whose radius is $\frac{1}{3}$ of that of B_x. The union of the β_x is an open set $W_{j,N}$ such that $E_N \subseteq W_{j,N} \subseteq V_j$

Let $K \subseteq W_{j,N}$ be compact. Finitely many β_x cover K. Lemma 1 shows that there is a finite set $F \subseteq E_N$ such that:

(a) $\{\beta_x : x \in F\}$ is a disjoint collection, and

(b) $K \subseteq \bigcup_{x\in F} B_x$

Therefore

$$\begin{aligned}\mu(K) &\le \sum_{x\in F}\mu(B_x)\\ &< N\sum_{x\in F} m(B_x)\\ &= 3^k N\sum_{x\in F} m(\beta_x)\\ &\le 3^k N m(V_j)\\ &< 3^k N/j\end{aligned}$$

This is true for any compact subset of $W_{j,N}$, since $W_{j,N}$ is open furthermore μ is a positive Borel measure on $\mathbb{R}^k$, so it is regular, therefore we have:

$$\mu(W_{j,N}) = \sup\{\mu(K)\colon K\subseteq W_{j,N} \text{ is compact }\} < 3^k N/j$$

Now let $\Omega_N = \bigcap_j W_{j,N}$, then $E_N \subseteq \Omega_N$, and Ω_N is a G_δ (so is measurable), and $\mu(\Omega_N)=0$, and so:

$$(D\mu)(x) = \infty \text{ for all } x\in S\setminus\bigcup_N \Omega_N$$

Since $\bigcup_N \Omega_N$ is a set of measure zero, we have the desired result. □

1.2 The Fundamental Theorem of Calculus

Problems with the FTC when extending to the Lebesgue integral

(a) Let

$$f(x) = \begin{cases} x^2\sin(x^{-2}) \text{ if } x\neq 0\\ 0 \text{ otherwise}\end{cases}$$

Then f is differentiable at every point, but

$$\int_0^1 |f'(t)|\ dt = \infty$$

So $f'\notin L^1$. But we still have:

$$f(x) = \lim_{\epsilon\to 0}\int_\epsilon^x f'(t)\ dt = \int_0^x f'(t)\ dt$$

(b) Suppose f is continuous on $[a,b]$, f is differentiable at almost every point of $[a,b]$ and $f'\in L^1$ on $[a,b]$. Do these assumptions imply that $f(x)-f(a)=\int_a^x f'$

NO!

Choose $\{\delta_n\}$ so that $1=\delta_0>\delta_1>\cdots$, where $\delta_n\to 0$, we define the sets E_n recursively. Put $E_0=[0,1]$ and if $n\ge 0$ and E_n is constructed so that it is the union of 2^n disjoint closed intervals, each of length $2^{-n}\delta_n$.

Delete a segment in the center of each of the 2^n intervals, so that each 2^{n+1} intervals have length $2^{-(n+1)}\delta_{n+1}$, and let E_{n+1} be the union of these 2^{n+1} intervals. So we have $E_1\supseteq E_2\supseteq\ldots$, and $m(E_n)=\delta_n$ for all n. Now let

$$E = \bigcap_{n=1}^{\infty} E_n$$

Note since each E_n is the finite union of closed sets, they are closed so E is also closed, it is also bounded since it is contained in $[0,1]$. So E is compact and $m(E)=\lim_{n\to\infty} m(E_n) = \lim_{n\to\infty}\delta_n = 0$. Put

$$g_n = \delta_n^{-1}\chi_{E_n} \text{ and } f_n(x) = \int_0^x g_n(t)\ dt \text{ for } n=0,1,2,\ldots$$

So note $f_n(0)=0$ and $f_n(1)=\delta_n^{-1}\int_0^1\chi_{E_n} = \delta_n^{-1}m(E_n)=1$. Each f_n is a monotonic function which is constant on each segment in E_n^c. If I is one of the 2^n interval whose union is E_n, then

$$\int_I g_n(t)\ dt = \int_I g_{n+1}(t)\ dt = 2^{-n}.$$

Therefore we see that

$$f_{n+1}(x) = f_n(x) \text{ for } x \notin E_n$$

Now note that

$$|f_n(x) - f_{n+1}(x)| \leq \int_I |g_n - g_{n+1}| < 2^{-(n-1)} \text{ for } x \in E_n$$

So $\{f_n\}$ converges uniformely to a continuous monotonic function f, with $f(0) = 0$, $f(1) = 1$, and $f'(x) = 0$ for all $x \notin E$. Since $m(E) = 0$, we see that $f' = 0$ almost everywhere. So

$$f(x) \neq \int_0^x f' \text{ in general}$$

Remark. Now we see that if $f' \in L^1$ and that

$$f(x) - f(a) = \int_a^x f' \tag{FTC}$$

Then there is a measure μ defined by $d\mu = f'\ dm$. Since $\mu \ll m$, we know that there corresponds to each $\epsilon > 0$ a $\delta > 0$ such that $|\mu|(E) < \epsilon$, whenever E is a union of disjoint segments whose total length is less than δ. Since $f(y) - f(x) = \mu((x, y))$ if $a \leq x < y \leq b$, it follows that the next definition is necessary for (FTC).

A complex function f, defined on an interval $I = [a, b]$ is said to be **absolutely continuous** on I (or f is AC on I) if for all $\epsilon > 0$ there is a $\delta > 0$ such that

$$\sum_{i=1}^n |f(\beta_i) - f(\alpha_i)| < \epsilon$$

For all nm and any disjoint collection of segments $(\alpha_1, \beta_1), \ldots, (\alpha_n, \beta_n)$ in I whose length satisfy

$$\sum_{i=1}^n (\beta_i - \alpha_i) < \delta$$

Remark. Such an f is continuous since we can choose $n = 1$.

Theorem 8. Let $I = [a, b]$ and $f \colon I \to \mathbb{R}$ be continuous and nondecreasing. TFAE

(a) f is AC on I

(b) f maps sets of measure 0 to sets of measure of 0.

(c) f is differentiable a.e. on I, $f' \in L^1$ and

$$f(x) - f(a) = \int_a^x f'(t)\ dt \text{ for } (a \leq x \leq b)$$

Proof.

- $(a) \Rightarrow (b)$

  Let $\mathcal{M}$ denote the σ-algebra of all Lebesgue measurable subsets of $\mathbb{R}$. Assume f is AC on I, pick $E \subseteq I$ so that $E \in \mathcal{M}$ and $m(E) = 0$. We will show that $f(E) \in \mathcal{M}$ and $m(f(E)) = 0$. WLOG we assume that $E \subseteq (a, b)$. Choose $\epsilon > 0$ and let $\delta > 0$ be as in definition 9. There is an open set V with $m(V) < \delta$, so that $E \subseteq V \subseteq I$. Let (α_i, β) be the disjoint segment whose union is V, then $\sum(\beta_i - \alpha_i) < \delta$ and so

  $$\sum |f(\beta_i) - f(\alpha_i)| < \epsilon \tag{\textit{dagger}}$$

  We know that this holds for every partial sum of this series, so it holds for the whole series, even if † is an infite sum.

  Since $E \subseteq V$, $f(E) \subseteq \bigcup [f(\alpha_i), f(\beta_i)]$. The Lebesgue measure of this union is $\sum |f(\beta_i) - f(\alpha_i)| < \epsilon$. So $f(E)$ is a subset of a borel set of arbitrarily small measure. Since Lebesgue measure is complete, we see that $f(E) \in \mathcal{M}$ and $m(f(E)) = 0$.

- $(b) \Rightarrow (c)$

  Define
  $$g(x) = x + f(x) \text{ for } (a \leq x \leq b).$$
  So note that if the f-image of some segment of length η has length η', then the g-image of this segment has length $\eta + \eta'$ (Indeed $m([x + f(x), y + f(y)]) = x - y + f(x) - f(y) = \eta + \eta'$). So we see that g satisfies condition (b). Now suppose $E \subseteq I$, $E \in \mathcal{M}$. Then $E = E_1 \cup E_0$ where $m(E_0) = 0$ and E_1 is F_σ, by a previous theorem. Thus E_1 is a countable union of compact set and so is $g(E_1)$ since g is continuous. Since $m(g(E_0)) = 0$, we have $g(E) = g(E_1) \cup g(E_0)$ so we conclude that $g(E) \in \mathcal{M}$.

  Therefore we can define
  $$\mu(E) = m(g(E)) \text{ for } E \subseteq I \text{ and } E \in \mathcal{M}$$
  Now let $x < y$, we see that $f(x) \leq f(y)$ therefore $g(x) = x + f(x) < y + f(y) = g(y)$, so this function is 1 to 1. Therefore disjoint sets in I have disjoint g-images. The countable additivity of m shows that μ is a positive bounded measure on $\mathcal{M}$. Furthermore since g satisfies (b) wew see that $\mu \ll m$ so
  $$d\mu = h\ dm$$
  for some $h \in L^1(m)$, by Radon-Nikodym.

  If $E = [a, x]$, then $g(E) = [g(a), g(b)]$ we have
  $$g(x) - g(a) = m(g(E)) = \mu(E) = \int_E h\ dm = \int_a^x h(t)\ dt$$
  So
  $$f(x) - f(a) = (g(x) - g(a)) - (x - a) = \int_a^x (h(t) - 1)\ dt$$
  Thus $f'(x) = h(x) - 1$ a.e. $[m]$, by Theorem .

- $(c) \Rightarrow (a)$ This is shown in the remark from definition 9

□

Theorem 9. Suppose $f \colon I \to \mathbb{R}$ is AC and $I = [a, b]$. Define
$$F(x) = \sup \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})|\ (a \leq x \leq b)$$
where the supremum is taken over all N and over all choices of $\{t_i\}$ such that
$$a = t_0 < t_1 < \cdots < t_N = x$$
The functions F, $F + f$, $F - f$ are then nondecreasing and AC on I.

Proof. If for $\{t_i\}$ with the above property and $x < y \leq b$ then
$$F(y) \geq |f(y) - f(x)| + \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})|$$
So $F(y) \geq |f(y) - f(x)| + F(x)$. In particular
$$F(y) \geq f(y) - f(x) + F(x) \text{ and } F(y) \geq f(x) - f(y) + F(x)$$
So $F, F + f, F - f$. Now we only need to show that F is AC on I since the sum of two AC functions is AC.
If $(\alpha, \beta) \subseteq I$ then
$$F(\beta) - F(\alpha) = \sum_{1}^{n} |f(t_i) - f(t_{i-1})|$$

Note that $\sum(t_i - t_{i-1}) = \beta - \alpha$.

Now let $\epsilon > 0$ and associate $\delta > 0$ to f and ϵ like in 9, choose disjoint segments $(\alpha_j, \beta_j) \subseteq I$ with $\sum(\beta_j - \alpha_j) < 0$, it follows that

$$\sum_j (F(\beta_j) - F(\alpha_j)) \leq \epsilon$$

Thus F is AC on I □

Definition 10. The function F defined in the theorem above is called the **total variation function** of f. If f is any (complex) function on I (AC or not), and $F(b) < \infty$, then f is said to have **bounded variation** on I and $F(b)$ is the **total variation** of f on I.

We have reached our main objective:

Theorem 10. If f is a complex function that is AC on I, then f is differentiable at almost all points of I, $f' \in L^1(m)$ and

$$f(x) - f(a) = \int_a^x f'(t)\ dt$$

Proof. We just need to prove this for real f. Let F be its total variation function and define:

$$f_1 = \frac{1}{2}(F + f) \text{ and } f_2 = \frac{1}{2}(F - f)$$

We apply theorem 8 to f_1 and f_2, and since

$$f = f_1 - f_2$$

We get

$$\begin{aligned} f(x) - f(a) &= (f_1(x) - f_1(a)) - (f_2(x) - f_2(a)) \\ &= \int_a^x f_1'(t)\ dt - \int_a^x f_2'(t)\ dt \\ &= \int_a^x (f_1' - f_2')(t)\ dt \end{aligned}$$

By a previous theorem we $f' = f_1' - f_2'$ a.e. and we get our desired result. □

Theorem 11. If $f\colon [a,b] \to \mathbb{R}$ is differentiable at *every* point of $[a,b]$ and $f' \in L^1$ on $[a,b]$, then

$$f(x) - f(a) = \int_a^x f'(t)\ dt$$

Proof. We just need to prove this for $x = b$. Fix $\epsilon > 0$, by a theorem from chap 2 we know there exists a lower semicontinuous function g on $[a,b]$ such that $g > f'$ and

$$\int_a^b g(t)\ dt < \int_a^b f'(t)\ dt + \epsilon$$

For any $\eta > 0$ we define

$$F_\eta(x) = \int_a^x g(t)\ dt - f(x) + f(a) + \eta(x - a)$$

We keep η fixed. For each $x \in [a,b)$ there corresponds a $\delta_x > 0$ such that

$$g(t) > f'(x) \text{ and } \frac{f(t) - f(x)}{t - x} < f'(x) + \eta$$

For all $t \in (x, x + \delta_x)$. Since g is lower semicontinuous and $g(x) > f'(x)$. For any such t we therefore have

$$F_\eta(t) - F_\eta(x) = \int_x^t g(s)\ ds - [f(t) - f(x)] + \eta(t - x) > (t - x)f'(x) - (t - x)(f'(x) + \eta) + \eta(t - x) = 0$$

Since $F_\eta(a) = 0$ and F_η is continuous there is a last point $x \in [a, b]$ at which $F_\eta(x) = 0$. If $x < b$, the preceding computation implies that $F_\eta(t) > 0$ for $t \in (x, b]$. In any case, $F_\eta(b) \geq 0$. Since this holds for every $\eta > 0$ we see that

$$f(b) - f(a) \leq \int_a^b g(t)\ dt < \int_a^b f'(t)\ dt + \epsilon$$

Since ϵ was arbitrary we conclude that

$$f(b) - f(a) \leq \int_a^b f'(t)\ dt$$

Furthermore $-f$ also satisfies the hypothesis of the theorem so the same inequality holds with $-f$ in the place of f, and these two inequalities give us the desired result. □

1.3 Differentiable Transformations

Definition 11. Suppose V is an open set in $\mathbb{R}^k$, T maps V into $\mathbb{R}^k$, and $x \in V$. If there exists a linear opeator A on $\mathbb{R}^k$ such that

$$\lim_{h \to 0} \frac{|T(x+h) - T(x) - Ah|}{|h|} = 0. \tag{1}$$

Then we say that T is **differentiable** at x, and define

$$T'(x) = A.$$

The linear operator $T'(x)$ is called the **derivative** of T at x. This linear operator is unique, indeed assume that we have linear operators A, B that both satisfy 1 then we have for all h:

$$\frac{|(B-A)(h)|}{|h|} = \frac{|(T(x+h) - T(x) - Ah) - (T(x+h) - T(x) - Bh)|}{|h|} \leq \frac{|T(x+h) - T(x) - Ah|}{|h|} + \frac{|T(x+h) - T(x) - Bh|}{|h|}.$$

Taking limits, we see that

$$\lim_{h \to 0} \frac{|(B-A)(h)|}{|h|} = 0.$$

Now let $x \in \mathbb{R}^k$ be non-zero, then we have

$$0 = \lim_{\to 0} \frac{|(B-A)(\lambda x)|}{|\lambda x|} = \frac{|(B-A)(x)|}{|x|}.$$

So we indeed see that $B - A = 0$. So this definition is indeed well-defined. Furthermore, since every $\alpha \in \mathbb{R}$ gives a linear operator on $\mathbb{R}^k$ where $h\alpha h$, this definition coincides with the definition of derivative in $\mathbb{R}$.

Definition 12. Recall that when $A \colon \mathbb{R}^k \to \mathbb{R}^k$ is linear there is a number $\Delta(A)$ such that

$$m(A(E)) = \Delta(A)m(E) \text{ for all measurable sets } E \subseteq \mathbb{R}^k.$$

Furthermore we know that $\Delta(A) = |\det A|$, so when T is differentiable at x, we call the determinant $T'(x)$ the **Jacobian** of T at x, and we denote it by $J_T(x)$. We thus have

$$\Delta(T'(x)) = |J_T(x)|.$$

Lemma 2. Let $S = \{x \colon |x| = 1\} \subseteq \mathbb{R}^k$, be the sphere that is the boundary of the unit ball $B = B(0, 1)$.
If $F \colon \overline{B} \to \mathbb{R}^k$ is continuous, $0 < \epsilon < 1$, and

$$|F(x) - x| < \epsilon. \tag{2}$$

for all $x \in S$, then $F(B) \supseteq B(0, 1 - \epsilon)$.

Proof. Assume, for a contradiction that some point $a \in B(0, 1-\epsilon)$ is not in $F(B)$. By 2, and the reverse triangle inequality we have

$$||F(x)| - \underbrace{|x|}_{1}| \leq |F(x) - x| < \epsilon \Rightarrow -\epsilon < |F(x)| - 1 < \epsilon \Rightarrow |F(x)| > 1 - \epsilon \text{ if } x \in S.$$

Thus a is not in $F(S)$, and therefore $a \neq f(x)$ for every $x \in B \cup S = \overline{B}$. So we define a continuous map $G\colon \overline{B} \to \overline{B}$ by

$$G(x) = \frac{a - F(x)}{|a - F(x)|}.$$

If $x \in S$, then

$$x \cdot (a - F(x)) = x \cdot a - x \cdot F(x) + 1 - 1 = x \cdot a + x \cdot (x - F(x))1 < |a| + \epsilon - 1 < 0..$$

So $x \cdot G(x) < 0$, so $x \neq G(x)$. On the other hand if $x \in B$, then since $G(x) \in S$ we see that $x \neq G(x)$. So G has no fix point, but this is impossible by Brouwer's theorem. $\square$

Index