Mathematical Analysis II

Felix Chen

Contents

Returning to the proof of **??**, we can assmue g is continuous with compact support,

$$\frac{1}{m(B)}\int_B |f(y)-g(y)|\,\mathrm{d}y \le M(f-g)(x)$$

Hence by taking B sufficiently small s.t. $|g(y)-g(x)| \le \varepsilon_0$ for all $x, y \in B$,

$$\frac{1}{m(B)}\int_B f(y)\,\mathrm{d}y \ge 3\varepsilon_0$$

$$\Longleftarrow \ |f(x)-g(x)| + M(f-g)(x) \ge 2\varepsilon_0.$$

But

$$m\{|f(x)-g(x)| \ge \varepsilon_0\} + m\{M(f-g) > \varepsilon_0\} \le \frac{\|f-g\|_{\mathcal{L}^1}}{\varepsilon_0} + \frac{3^d}{\varepsilon_0}\|f-g\|_{\mathcal{L}^1} \le \frac{3^d+1}{\varepsilon_0}\varepsilon.$$

This completes the proof.

Definition 0.1 (Lebesgue points)**.** Let $|f(x)| < \infty$, f is *locally integrable*. If x satisfies

$$\lim_{|B|\to 0, B\ni x} \frac{1}{|B|}\int_B |f(y)-f(x)|\,\mathrm{d}y = 0,$$

we say x is a **Lebesgue point** of f.

Remark 0.2 — Here "locally integrable" means for all bounded measurable sets E, $f\chi_E \in \mathcal{L}^1$. This is denoted by $f \in \mathcal{L}^1_{loc}$.

Let E be a measurable set, χ_E locally integrable, If point x is called a **density point** of E if it's a Lebesgue point of χ_E.

Theorem 0.3

Let E be a measurable set, then almost all the points in E are density points of E, almost all the points outside of E are not density points of E.

Proof. This is a direct corollary of **??**. □

The differentiation theorem has some applications in convolution:

$$\begin{aligned}\frac{1}{|B|}\int_B f(y)\,\mathrm{d}y &= c_d^{-1}\varepsilon^{-d}\int_{B(x,\varepsilon)} f(y)\,\mathrm{d}y\\ &= \int f(x-y)\cdot c_d^{-1}\varepsilon^{-d}\chi_{B(0,\varepsilon)}(y)\,\mathrm{d}y\\ &= f * K_\varepsilon.\end{aligned}$$

where $K_\varepsilon = c_d^{-1}\varepsilon^{-d}\chi_{B(0,\varepsilon)}$, c_d is the measure of a unit sphere in $\mathbb{R}^d$.

By differentiation theorem, $\lim_{\varepsilon\to 0} f * K_\varepsilon = f(x), a.e.$. In the homework we proved that there doesn't exist a function I s.t. $f * I = f$ for all $f \in \mathcal{L}^1$, but the functions K_ε is approximating this "convolution identity".

Definition 0.4. In general, if $\int K_\varepsilon = 1$, $|K_\varepsilon| \le A\min\{\varepsilon^{-d}, \varepsilon|x|^{-d-1}\}$ for some constant A, we say K_ε is an **approximation to the identity**.

"convolution kernel"

Let φ be a smooth function whose support is in $\{|x| \le 1\}$, and $\int\varphi = 1$. The function $K_\varepsilon := \varepsilon^{-d}\varphi(\varepsilon^{-1}x)$ is called the Friedrichs smoothing kernel.

Theorem 0.5

If K_ε is an approximation to the identity, f integrable,

$$\lim_{\varepsilon\to 0}\|f * K_\varepsilon - f\|_{\mathcal{L}^1} = 0.$$

Proof.

$$\begin{aligned}|(f*K_\varepsilon)(x) - f(x)| &= \left|\int f(x-y)K_\varepsilon(y)\,\mathrm{d}y - f(x)\right|\\ &\le \int |f(x-y)-f(x)||K_\varepsilon(y)|\,\mathrm{d}y\\ &\le \int_{|y|\le R}|f(x-y)-f(x)|A\varepsilon^{-d}\,\mathrm{d}y + \int_{|y|>R}|f(x-y)-f(x)|A\varepsilon|y|^{-d-1}\,\mathrm{d}y.\end{aligned}$$

Taking the integral over $\mathbb{R}^d$:

$$\begin{aligned}&\|K_\varepsilon * f - f\|_{\mathcal{L}^1}\\ &\le A\varepsilon^{-d}\int\int_{|y|\le R}|f(x-y)-f(x)|\,\mathrm{d}y\,\mathrm{d}x + A\varepsilon\int\int_{|y|>R}|f(x-y)-f(x)||y|^{-d-1}\,\mathrm{d}y\,\mathrm{d}x\\ &\le A\varepsilon^{-d}\int\int_{|y|\le R}|\tau_{-y}f(x)-f(x)|\,\mathrm{d}y\,\mathrm{d}x + A\varepsilon\int_{|y|>R}|y|^{-d-1}\int|\tau_{-y}f(x)|+|f(x)|\,\mathrm{d}x\,\mathrm{d}y\\ &\le A\varepsilon^{-d}\int_{|y|\le R}\|\tau_{-y}f - f\|_{\mathcal{L}^1}\,\mathrm{d}y + A\varepsilon\int_{|y|>R}|y|^{-d-1}2\|f\|_{\mathcal{L}^1}\,\mathrm{d}y.\end{aligned}$$

By the continuity of translation, $\forall\varepsilon_0$, let R be sufficiently small we have

$$\|\tau_{-y}f - f\|_{\mathcal{L}^1} < \varepsilon_0, \forall|y| < R.$$

$$\|K_\varepsilon * f - f\|_{\mathcal{L}^1} \le A\varepsilon^{-d}R^d c_d \varepsilon_0 + \varepsilon R^{-1} C$$

where C is a constant.

Take suitable $\varepsilon, \varepsilon_0$ s.t. $R = \varepsilon\varepsilon_0^{-\frac{1}{2d}}$, then $LHS \le C_1 \varepsilon_0^{\frac{1}{2}} + C_2 \varepsilon_0^{\frac{1}{2d}} \to 0$. □

Theorem 0.6

Let K_ε be an approximation to the identity, f integrable,

$$\lim_{\varepsilon \to 0} f * K_\varepsilon = f(x)$$

holds for Lebesgue points x of f.

Proof. WLOG $x = 0$, let

$$\omega(r) = \frac{1}{r^d} \int_{B(0,r)} |f(y) - f(0)| \, \mathrm{d}y,$$

we have $\lim_{r\to 0} \omega(r) = 0$, and ω is continuous.

$$\omega(r) \le \frac{\|f\|_{\mathcal{L}^1}}{r^d} + |f(0)|,$$

Thus ω is bounded.

Therefore we can compute

$$\begin{aligned}
|K_\varepsilon * f(x) - f(x)| &\le \int |K_\varepsilon(y)||f(x-y) - f(x)| \, \mathrm{d}y \\
&\le \int_{B(0,r)} |K_\varepsilon(y)||f(x-y) - f(x)| \, \mathrm{d}y + \sum_{k\ge 0} \int_{2^k r \le |y| < 2^{k+1} r} |K_\varepsilon(y)||f(x-y) - f(x)| \, \mathrm{d}y \\
&\le A\varepsilon^{-d} \cdot r^d \omega(r) + \sum_{k \ge 0} \int_{2^k r \le |y| < 2^{k+1} r} A\varepsilon |y|^{-d-1} |f(x-y) - f(x)| \, \mathrm{d}y \\
&\le A\varepsilon^{-d} r^d \omega(r) + \sum_{k\ge 0} A\varepsilon 2^{-k(d+1)} r^{-d-1} \cdot 2^{(k+1)d} r^d \omega(2^{k+1} r) \\
&= A\varepsilon^{-d} r^d \omega(r) + A 2^d \varepsilon r^{-1} \sum_{k \ge 0} 2^{-k} \omega(2^{k+1} r).
\end{aligned}$$

Let $r = \varepsilon$, since $\omega(r)$ is continuous and bounded, we're done. □

§0.1 Lebesgue Differentiation theorem for monotone functions

Recall that Fundamental theorem of Calculus has a second part, Given any differentiable function $F(x)$, if $F'(x)$ Riemann integrable, then

$$F(b) - F(a) = \int_a^b F'(x) \, \mathrm{d}x.$$

When F has nice properties (smooth), then it has a generalization to higher dimensions (Stokes' theorem).

But in Lebesgue integration theory, the derivative of a function is hard to define. Hence we'll find an alternative for $F'(x)$.

Example 0.7

Consider Heaviside function $H(t) = \begin{cases} 1, & t > 0 \\ 0, & t \leq 0 \end{cases}$

Then H is differentiable almost everywhere, but $\int_{-1}^{1} H'(t)\,\mathrm{d}t = 0 \neq H(1) - H(-1)$.

Example 0.8

Consider Cantor-Lebesgue function F, similarly we have $F'(x) = 0, a.e.$, but $\int_0^1 F'(x)\,\mathrm{d}x = 0 \neq F(1) - F(0)$.

Definition 0.9 (Dini derivatives)**.** Let $f(x)$ be a measurable function, define

$$D^+(f)(x) = \limsup_{h>0, h\to 0} \frac{f(x+h) - f(x)}{h}, \quad D^-(f)(x) = \limsup_{h<0, h\to 0} \frac{f(x+h) - f(x)}{h}.$$

$$D_+(f)(x) = \liminf_{h>0, h\to 0} \frac{f(x+h) - f(x)}{h}, \quad D_-(f)(x) = \liminf_{h<0, h\to 0} \frac{f(x+h) - f(x)}{h}.$$

Theorem 0.10 (Lebesgue Differentiation theorem for increasing functions)

Let f be an increasing function on $[a, b]$, then $F'(x)$ exists almost everywhere, and

$$\int_a^b F'(x)\,\mathrm{d}x \leq F(b) - F(a).$$

Sketch of the proof. The outline of the proof is as follows:
Step 1: Decompose $F = F_c + J$, where F_c is continuous, J is a jump function.
Step 2: Prove F_c increasing and $J' = 0, a.e.$.
Step 3: Prove $D^+(F) < +\infty, a.e.$, $D^+(F) \leq D_-(F), a.e.$, and $D^-(F) \leq D_+(F), a.e.$. □

We proceed step by step.

Step 1 Denote $F(x+0) = \lim_{h\to 0^+} F(x+h), F(x-0) = \lim_{h\to 0^-} F(x+h)$.
Since F increasing, let $\{x_n\}$ be all the discontinuous points of F. Define:

$$j_n(x) = \begin{cases} 0, & x < x_n \\ \beta_n, & x = x_n \\ \alpha_n, & x > x_n \end{cases}$$

where $\alpha_n = F(x_n + 0) - F(x_n - 0), \beta_n = F(x_n) - F(x_n - 0)$.
Hence the jump function

$$J_F(x) = \sum_{n=1}^{+\infty} j_n(x) \leq \sum_{n=1}^{+\infty} \alpha_n = \sum_{n=1}^{+\infty} (F(x_n + 0) - F(x_n - 0)) \leq F(b) - F(a)$$

is well-defined and increasing.

Lemma 0.11
$F - J_F$ is continuous and increasing.

Proof. First note that

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

This can be derived from the definition of J_F: If F is continuous at x, the equality is obvious;

If $x = x_n$ for some n,

$$\lim_{h\to 0^+} J_F(x+h) = \sum_{x_k \le x_n} \alpha_k + \lim_{h\to 0^+} \sum_{x_n < x_k \le x_n + h} j_k(x+h) = \sum_{x_k \le x_n} \alpha_k$$

$$\lim_{h\to 0^+} J_F(x-h) = \lim_{h\to 0^+} \sum_{x_k < x_n - h} \alpha_k + \lim_{j\to 0^+} \sum_{x_k = x_n - h} \beta_k = \sum_{x_k < x_n} \alpha_k$$

Note that $\alpha_n = F(x_n+0) - F(x_n-0)$, thus $F - J_F$ is continuous.

Secondly,

$$F(x) - J_F(x) \le F(y) - J_F(y), \quad \forall a \le x \le y \le b.$$

Because

$$J_F(y) - J_F(x) = \sum_{x < x_j < y} \alpha_j + \sum_{x_k = y} \beta_k - \sum_{x_k = x} \beta_k \le \sum_{x < x_j < y} \alpha_j + F(y) - F(y-0) \le F(y) - F(x).$$

which means $F - J_F$ is increasing. □

Step 2

Proposition 0.12
The jump function $J(x)$ is differentiable almost everywhere, and $J'(x) = 0, a.e.$.

Proof. The Dini derivatives of $J(x)$ exist and are non-negative (since J is increasing).

$$\overline{D}(J)(x) = \max\{D^+(J)(x), D^-(J)(x)\}.$$

Let $E_\varepsilon = \{\overline{D}(J)(x) > \varepsilon > 0\}$. We'll prove E_ε is null for all ε. If $x \in E_\varepsilon$, $\exists h$ s.t.

$$\frac{J(x+h) - J(x)}{h} > \varepsilon \implies J(x+h) - J(x-h) > \varepsilon h.$$

Let $N \in \mathbb{N}$ s.t. $\sum_{n>N} \alpha_n < \frac{\varepsilon\delta}{10}$. Define $J_N(x) = \sum_{n>N} j_n(x)$.

$$E_{\varepsilon,N} = \{\overline{D}(J_N)(x) > \varepsilon\}, \quad E_\varepsilon \subset E_{\varepsilon,N} \cup \{x_1, \dots, x_N\},$$

Since for $x \ne x_i$,

$$\begin{aligned}\overline{D}(J)(x) &= \limsup_{h\to 0} \frac{J(x+h) - J(x)}{h} \\ &= \limsup_{h\to 0} \left(\frac{J_N(x+h) - J_N(x)}{h} + \frac{1}{h}\sum_{n=1}^{N} (j_n(x+h) - j_n(x)) \right) = \overline{D}(J_N)(x).\end{aligned}$$

Next we need to control the measure of $E_{\varepsilon,N}$.

For all $y \in E_{\varepsilon,N}$, there exists sufficiently small h s.t. $J_N(y+h) - J_N(y) > h\varepsilon$. So the intervals $(y-h, y+h)$ is a covering of $E_{\varepsilon,N}$, and it can be controled using the value of J_N. Therefore we hope to find some *disjoint* intervals which cover certain ratio of $E_{\varepsilon,N}$.

Lemma 0.13

Let $\mathcal{B}$ be a collection of balls with bounded radius in $\mathbb{R}^d$. There exists countably many disjoint balls B_i s.t.

$$\bigcup_{B\in\mathcal{B}} B \subset \bigcup_{i=1}^{\infty} 5B_i.$$

Proof. Let $r(B)$ denote the radius of B. Take B_1 s.t. $r(B_1) > \frac{1}{2}\sup_{b\in\mathcal{B}} r(B)$.
The rest is the same as before. □

By lemma, there exists countably many disjoint intervals $(x_i + h_i, x_i - h_i)$ s.t.

$$\begin{aligned} m^*(E_{\varepsilon,N}) &\le 5\sum_{i=1}^{\infty} 2h_i \\ &\le 10\sum_{i=1}^{\infty} \varepsilon^{-1}(J_N(x_i+h_i) - J_N(x_i-h_i)) \\ &\le 10\varepsilon^{-1}(J_N(b) - J_N(a)) < \delta. \end{aligned}$$

Hence $m^*(E_\varepsilon) \le m^*(E_{\varepsilon,N}) < \delta \implies m^*(E_\varepsilon) = 0$, which gives $\overline{D}(J) = 0, a.e.$. □

Step 3 First we prove $D^+(F) < \infty, a.e.$.
Let $E_\gamma = \{x : D^+(F)(x) > \gamma\}$.
When $h \in [\frac{1}{n+1}, \frac{1}{n}]$:

$$\begin{aligned} \frac{F(x+h) - F(x)}{h} &\le \frac{n+1}{n}\frac{F(x+\frac{1}{n}) - F(x)}{\frac{1}{n}}, \\ &\ge \frac{n}{n+1}\frac{F(x+\frac{1}{n+1}) - F(x)}{\frac{1}{n+1}}. \end{aligned}$$

Thus

$$D^+(F)(x) = \limsup_{h\to 0}\frac{F(x+h) - F(x)}{h} = \limsup_{n\to\infty}\frac{F(x+\frac{1}{n}) - F(x)}{\frac{1}{n}}$$

is measurable.

Lemma 0.14 (Riesz sunrise lemma)

Let $G(x)$ be a continuous function on $\mathbb{R}$. Define

$$E = \{x : \exists h > 0, s.t.\ G(x+h) > G(x)\}.$$

Then E is open and $E = \bigcup_{i=1}^{\infty}(a_i, b_i)$, where (a_i, b_i) are disjoint finite intervals s.t. $G(a_i) = G(b_i)$.

When G is defined on finite interval $[a, b]$, we also have $G(a) \le G(b_1)$.

Proof. Note that E is open since G is continuous.
Take a maximum open interval $(a, b) \subset E$, i.e. $a, b \notin E$, so $G(a) \ge G(b)$.
Since $b \notin E, G(x) \le G(b), \forall x > b$. If $G(a) > G(b)$, Let $G(a+\varepsilon) > G(b)$, as $a+\varepsilon \in E$, exists $h > 0$ s.t. $G(a+\varepsilon+h) > G(a+\varepsilon)$.

But G has a maximum on $[a+\varepsilon, b]$, say $G(c)$, we must have $c \neq a+\varepsilon, b$. This leads to a contradiction. $\square$

> **Remark 0.15** — This lemma provides a better estimation than previous covering lemmas, since it directly claims that E can be broken into disjoint intervals.

For $x \in E_\gamma$, $\exists h > 0$ s.t. $F(x+h) - F(x) > \gamma h$, by Lemma 0.14 on $F(x) - \gamma x$,

$$m(E_\gamma) \leq \sum_{k=1}^{\infty}(b_k - a_k) \leq \gamma^{-1}\sum_{k=1}^{\infty}(F(b_k) - F(a_k)) \leq \gamma^{-1}(F(b) - F(a)).$$

Therefore when $\gamma \to \infty$, $m(E_\gamma) \to 0$.

The last part is $D^+(F) \leq D_-(F), a.e.$.

Similarly it's sufficient to prove the following set is null for all rational numbers $r < R$:

$$E_{r,R} = \{D^+(F)(x) > R, D_-(F)(x) < r\}.$$

Since $D^+(F)$ is measurable, $E_{r,R}$ is measurable. If $m(E_{r,R}) > 0$, we can restrict it to a smaller interval $[c,d] \subset [a,b]$ such that $d - c < \frac{R}{r}m(E_{r,R})$.

Let $G(x) = F(-x) + rx$, by Lemma 0.14 on $[-d, -c]$,

$$\{s : \exists h > 0, G(x+h) > G(x)\} = \bigcup_k(-b_k, -a_k).$$

Note that $-E_{r,R}$ is contained in the above set, and $G(-b_k) \leq G(-a_k) \iff F(b_k) - F(a_k) \leq r(b_k - a_k)$,

We use Lemma 0.14 again on each (a_k, b_k) and $F(x) - Rx$,

$$E_{r,R} \cap (a_k, b_k) \subset \{x : \exists h > 0, F(x+h) - F(x) \geq Rh\} = \bigcup_{l=1}^{\infty}(a_{k,l}, b_{k,l}).$$

Hence

$$\begin{aligned} m(E_{r,R}) &\leq \sum_{k,l=1}^{\infty}(b_{k,l} - a_{k,l}) \\ &\leq R^{-1}\sum_{k,l=1}^{\infty}(F(b_{k,l}) - F(a_{k,l})) \leq R^{-1}\sum_{k=1}^{\infty}(F(b_k) - F(a_k)) \\ &\leq R^{-1}r\sum_{k=1}^{\infty}(b_k - a_k) \leq R^{-1}r(d-c), \end{aligned}$$

which gives a contradiction! So $m(E_{r,R}) = 0$ for all rationals $r < R$. Therefore we're done by

$$m(\{D^+(F) > D_-(F)\}) \leq \sum_{r,R} m(E_{r,R}) = 0$$

Now we can complete the proof of Theorem 0.10. Here we state the theorem again:

Let F be an increasing function on $[a,b]$, then F is differentiable almost everywhere, and

$$\int_a^b F'(x)\,\mathrm{d}x \leq F(b) - F(a).$$

Let $F_n(x) = n(F(x+\frac{1}{n}) - F(x))$, where $F(x) = F(b)$ for $x > b$. Since $F_n \geq 0$, by Fatou's Lemma, (we've already proved F is differentiable almost everywhere and $F' \geq 0$)

$$\begin{aligned}
&\int_a^b \liminf_{n\to\infty} F_n \,\mathrm{d}x \leq \liminf_{n\to\infty} \int_a^b F_n \,\mathrm{d}x \\
\implies &\int_a^b F'(x) \,\mathrm{d}x \leq \liminf_{n\to\infty} \int_a^b n\left(F\left(x+\frac{1}{n}\right) - F(x)\right) \mathrm{d}x \\
&= \liminf_{n\to\infty} n\left(\int_{a+\frac{1}{n}}^{b+\frac{1}{n}} F(x) - \int_a^b F(x)\right) \mathrm{d}x \\
&= \liminf_{n\to\infty} \left(F(b) - n\int_a^{a+\frac{1}{n}} F(x)\,\mathrm{d}x\right) \\
&\leq F(b) - F(a)
\end{aligned}$$

§0.2 Absolute continuous functions

Definition 0.16 (Absolute continuity)**.** We say a function $F(x)$ is **absolutely continuous** on interval $[a,b]$, if $\forall \varepsilon > 0, \exists \delta > 0$, such that for all disjoint intervals $(a_k, b_k), k = 1, \ldots, N$ with

$$\sum_{k=1}^{N} (b_k - a_k) < \delta,$$

we must have

$$\sum_{k=1}^{N} |F(b_k) - F(a_k)| < \varepsilon.$$

The space consisting of all the absolutely continuous functions on $[a,b]$ is denoted by $Ac([a,b])$.

Example 0.17

A C^1 function with bounded derivative or a Lipschtiz function is absolutely continuous.

Some obvious properties of absolutely continuous function F:

- F is continuous;
- F has bounded variation, i.e. $F \in BV$.
- F is differentiable almost everywhere, since $F = F_1 - F_2$, where F_1, F_2 are increasing. In fact we have
$$T_F([a,b]) = \int_a^b |F'(x)| \,\mathrm{d}x.$$
- If N is a null set, then $F(N)$ is also null. In particular F maps measurable sets to measurable sets.

Proof of the last property. Take intervals (a_k, b_k) s.t. $N \subset \bigcup_{k=1}^{\infty}(a_k, b_k)$. Since $F(N) \subset F(\bigcup(a_k, b_k))$,

$$|F(N)| \leq \sum_{k=1}^{\infty} |F([a_k, b_k])| \leq \sum_{k=1}^{\infty} |F(\tilde{b}_k) - F(\tilde{a}_k)| < \varepsilon.$$

□

Proposition 0.18

The space $Ac([a,b]) \subset BV([a,b])$, moreover it's an algebra, and it's a separable Banach space under the norm induced from BV.

Finally we come to the full generalization of Newton-Lebniz formula:

Theorem 0.19 (Fundamental theorem of Calculus)

A function $F \in Ac([a,b]) \implies F$ is differentiable almost everywhere, F' is integrable, and

$$F(x) - F(a) = \int_a^x F'(\tilde{x})\,\mathrm{d}\tilde{x}, \quad \forall x \in [a,b].$$

Proof. Let $\tilde{F}(x) = F(a) + \int_a^b F'(y)\,\mathrm{d}y \in Ac([a,b])$ (by the absolute continuity of integrals).

We have $F - \tilde{F} \in Ac([a,b])$ and $(F - \tilde{F})' = 0, a.e.$.

Thus it suffices to prove the following theorem:

Theorem 0.20

Let $F \in Ac([a,b])$, and $F' = 0, a.e.$, then $F(a) = F(b)$, i.e. F is constant on $[a,b]$.

□

To prove this, we'll need Vitali covering theorem:

Definition 0.21 (Vitali covering)**.** Let $\mathcal{B} = \{B_\alpha\}$, where B_α is closed balls in $\mathbb{R}^d$. We say $\mathcal{B}$ is a **Vitali covering** of a set E, if $\forall x \in E, \forall \eta > 0$, exists $B_\alpha \in \mathcal{B}$ s.t. $m(B_\alpha) < \eta$, $x \in B_\alpha$.

Theorem 0.22 (Vitali)

Let $E \subset \mathbb{R}^d$ with $m^*(E) < \infty$, for any Vitali covering $\mathcal{B}$ of E and $\delta > 0$, exists disjoint balls $B_1, \ldots, B_n \in \mathcal{B}$, such that

$$m^*\left(E \backslash \bigcup_{i=1}^n B_i\right) < \delta.$$

Proof. For all $\varepsilon > 0$, exists an open set A s.t. $E \subset A$ and $m(A) < m^*(E) + \varepsilon < +\infty$.

Remove all the balls in $\mathcal{B}$ with radius greater than 1. Each time we take a ball B_i with radius greater than $\frac{1}{2}\sup_{B \in \mathcal{B}'} r(B)$, where $\mathcal{B}'$ are the remaining balls, and remove all the balls which intersect with B_i.

If we end up with finitely many balls $B_1, \ldots, B_n$, we must have $E \subset \bigcup_{i=1}^n B_i$, otherwise $x \notin B_i \implies \exists B \ni x$, $B \cap B_i = \emptyset$, contradiction!

If we take out countably many balls $B_1, B_2, \cdots \subset A$, since $\sum_{i=1}^\infty m(B_i) \le m(A) < \infty$, there exists N s.t. $\sum_{i=N+1}^\infty m(B_i) < 5^{-d}\delta$.

Now we only need to prove

$$E \backslash \bigcup_{i=1}^N B_i \subset \bigcup_{i>N} 5B_i.$$

□

Let $E = \{x : F'(x) = 0\}$, $\forall x \in E$, $\exists \delta(x) > 0$, s.t.

$$|F(y) - F(x)| < \varepsilon |y - x|, \forall |y - x| < \delta(x).$$

Hence $[x - h, x + h], 0 < h < \delta(x)$ is a Vitali covering of E. By Theorem 0.22, there exists finitely many disjoint intervals $[x_k - h_k, x_k + h_k] = I_k$ s.t.

$$m^*(E \setminus \bigcup_{k=1}^{N} I_k) < \varepsilon.$$

Assmue $a \leq a_1 < b_1 < \cdots < a_N < b_N \leq b$.

$$F(b) - F(a) \leq \sum_{k=1}^{N} |F(b_k) - F(a_k)| + \sum_{k=0}^{N} |F(a_{k+1}) - F(b_k)| \leq \varepsilon(b - a) + \delta.$$