Mathematical Analysis II

Felix Chen

Contents

Theorem 0.1 (Fubini's Theorem)
Let $f(x,y) : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$, and f is integrable on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$.

1. $f(x,y)$ as a function of y is integrable on $\mathbb{R}^{d_2}$ for $x \in \mathbb{R}^{d_1} \backslash Z$ with $m(Z) = 0$.
2. Let $g(x) = \int_{\mathbb{R}^{d_2}} f(x,y)\,\mathrm{d}y$, for $x \in \mathbb{R}^{d_1} \backslash Z$, where Z is a null set. We have g is integrable on $\mathbb{R}^{d_1}$.
3. $$\int_{\mathbb{R}^{d_1+d_2}} f(x,y) = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x.$$

Proof. Let $\mathscr{F}$ be the space consisting of all the integrable functions that satisfy Fubini's theorem.

Lemma 0.2
$\mathscr{F}$ is a vector space. Furthermore, for non-negative monotone sequence $f_n \in \mathscr{F}$, if $\lim f_n$ is integrable, then $\lim f_n \in \mathscr{F}$ as well.

Proof of the lemma. First notice that $f \in \mathscr{F} \implies cf \in \mathscr{F}$.

If $f, g \in \mathscr{F}$, consider $f + g$:

By our conditions, there exists $X_f, X_g \subset \mathbb{R}^{d_1}$, s.t. $f(x,y)$ integrable on $\mathbb{R}^{d_2}$, $\forall x \notin X_f$, and $g(x,y)$ integrable on $\mathbb{R}^{d_2}$, $\forall x \notin X_g$.

This implies $f(x,y) + g(x,y)$ integrable on $\mathbb{R}^{d_2}$ for $x \notin X_f \cup X_g$, which proves (1).

$$\int_{\mathbb{R}^{d_2}} f(x,y) + g(x,y)\,\mathrm{d}y = \int_{\mathbb{R}^{d_2}} f(x,y)\,\mathrm{d}y + \int_{\mathbb{R}^{d_2}} g(x,y)\,\mathrm{d}y.$$

So the LHS is integrable on $\mathbb{R}^{d_1}$ (this is (2)), taking the integral we get

$$\int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x + \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} g(x,y)\,\mathrm{d}y \right) \mathrm{d}x = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x,y) + g(x,y)\,\mathrm{d}y \right) \mathrm{d}x.$$

Therefore $\mathscr{F}$ is a vector space.

For a monotone non-negative sequence f_n, $\exists X_n \subset \mathbb{R}^{d_1}$ s.t. f_n is integrable with respect to y for $x \notin X_n$.

Similarly, when $x \notin \bigcup_{n=1}^{\infty} X_n$, as a function of y, by Beppo-Levi (or Dominated convergence),

$$\int_{\mathbb{R}^{d_2}} f(x,y)\,\mathrm{d}y = \lim_{n\to\infty} \int_{\mathbb{R}^{d_2}} f_n(x,y)\,\mathrm{d}y.$$

This equation holds when $\int f(x,y)\,\mathrm{d}y$ is finite, so we need to prove it is finite almost everywhere. For $x \notin \bigcup X_n$, we have:

$$\int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f_n(x,y)\,\mathrm{d}y \right) \mathrm{d}x \to \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$

$$\int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f_n(x,y)\,\mathrm{d}y \right) \mathrm{d}x = \int_{\mathbb{R}^{d_1+d_2}} f_n \to \int_{\mathbb{R}^{d_1+d_2}} f$$

Compare these relations we deduce

$$\int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x = \int_{\mathbb{R}^{d_1+d_2}} f < +\infty.$$

so $\int_{\mathbb{R}^{d_2}} f(x,y)\,\mathrm{d}y$ is finite almost everywhere. This gives (1), and (2), (3) follows immediatedly. □

Back to the proof of the original theorem, we want to prove $\mathscr{F} = \mathcal{L}^1$.
We prove the indicator function of following sets are in $\mathscr{F}$:

- Cuboids;
- Finite open sets;
- G_δ sets;
- Null sets;
- General measurable sets.

Let I be a cuboid, $I = I_x \times I_y$, so $\chi_I = \chi_{I_x}\chi_{I_y}$.

$$\int \chi_I = |I| = |I_x||I_y| = \int \chi_{I_x}|I_y|\,\mathrm{d}x = \int \int (\chi_{I_x}\chi_{I_y}\,\mathrm{d}y)\,\mathrm{d}x.$$

Let O be a finite open set, $O = \bigcup_{n=1}^{\infty} I_n$, where I_n are pairwise disjoint cuboids.

$$\chi_O = \lim_{n\to\infty} \chi_{\bigcup_{k=1}^{n} I_k} \in \mathscr{F},$$

as it's an inceasing sequence.
For $G_\delta = \bigcap_{n=1}^{\infty} O_n$, $\chi_{O_n} \searrow \chi_{G_\delta}$. $\implies \chi_{G_\delta} \in \mathscr{F}$.
For null set E, if $\chi_E \in \mathscr{F}$, $\forall A \subset E$,

$$0 = \int \chi_E = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} \chi_E\,\mathrm{d}y \right) \mathrm{d}x.$$

hence $\int_{\mathbb{R}^{d_2}} \chi_E\,\mathrm{d}y = 0$, for $x, a.e.$ $\implies \int_{\mathbb{R}^{d_2}} \chi_A\,\mathrm{d}y = 0$ for $x, a.e.$.
Taking the integral with respect to x, we have $\chi_A \in \mathscr{F}$.
Therefore if E is a null set, by taking its equi-measure hull we deduce $\chi_E \in \mathscr{F}$.
Finally, for a general measurable set E, let O be its equi-measure hull, and $E = O \backslash A$. since $\mathscr{F}$ is a vector space, $\chi_E \in \mathscr{F}$.
The rest is trival now: Because all the simple functions are in $\mathscr{F}$, and any measurable functions can be expressed as limits of increasing simple functions, so $\mathscr{F} = \mathcal{L}^1(\mathbb{R}^{d_1+d_2})$. □

Theorem 0.3 (Tonelli's theorem)
Let f be a non-negative measurable function on $\mathbb{R}^d$.

- $f(x, y)$ is measurable on $\mathbb{R}^{d_2}$ for x almost everywhere;
- $\int_{\mathbb{R}^{d_2}} f(x, y)\,\mathrm{d}y$ as a function of x is measurable;
- The integral satisfies:
$$\int_{\mathbb{R}^d} f = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x, y)\,\mathrm{d}y \right) \mathrm{d}x.$$

Proof. Consider the truncation function $f(x, y)\chi_{|x|+|y|<k}\chi_{f<k}$. □

Proposition 0.4
Let E be a measurable set on $\mathbb{R}^d$. For x almost everywhere, $E^x = \{y \mid (x, y) \in E\}$ is measurable on $\mathbb{R}^{d_2}$.

As a function of x, $m(E^x)$ satisfies
$$m(E) = \int_{\mathbb{R}^{d_1}} m(E^x).$$

Proof. Consider $f = \chi_E$ and use Tonelli's theorem. □

§0.1 Applications of Fubini's theorem

Definition 0.5 (Product measure)**.** Let $(X, \mathscr{F}, m)$ and $(Y, \mathscr{G}, m)$ be measure spaces, define a measure on $X \times Y$: The measure m induces an outer measure on $X \times Y$, and complete it to a normal measure by using Caratheodory conditions. This measure is called the **product measure** on $X \times Y$.

Theorem 0.6
Let $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, E_1, E_2 are subsets of $\mathbb{R}^{d_1}, \mathbb{R}^{d_2}$, respectively.

- If E_1, E_2 are measurable, then E is measurable as well, and $m(E) = m(E_1)m(E_2)$.
- If E is measurable, then E_1, E_2 are measurable, and $m(E) = m(E_1)m(E_2)$, unless one of E_1, E_2 is null set, which means E is null as well.

Proof. First it's easy to note that
$$m^*(E) \le m^*(E_1)m^*(E_2).$$
So we directly conclude that if one of E_1, E_2 is null set, E must be null.

Thus we may assume below that E_1, E_2 have finite nonzero measure. By taking the equimeasure hull of E_1, E_2 (denoted by F_1, F_2), let $Z_1 = F_1 \backslash E_1, Z_2 = F_2 \backslash E_2$, we have
$$(F_1 \times F_2) \backslash (Z_1 \times F_2 \cup F_1 \backslash Z_2) \subset E \subset F_1 \times F_2,$$

so E is measurable.

Conversely, if E is measurable, consider the measurable function χ_E, by definition $\chi_E = \chi_{E_1}\chi_{E_2}$, hence by Tonelli's theorem, for x almost everywhere, $\chi_{E_1}(x)\chi_{E_2}$ is measurable on $\mathbb{R}^{d_2}$ $\implies$ E_2 is measurable.

Therefore we have the equation

$$m(E) = \int_{\mathbb{R}^d} \chi_E = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} \chi_{E_1}\chi_{E_2} \right) = m(E_1)m(E_2).$$

This proves the theorem. □

Corollary 0.7

Let $f(x)$ be a measurable function on $\mathbb{R}^{d_1}$, we have $g(x,y) = f(x)$ is measurable on $\mathbb{R}^{d_2}$.

Proof. It's sufficient to prove that $\{(x,y)|f(x) > t\}$ is measurable in $\mathbb{R}^d$. This follows from the fact that

$$\{(x,y)|f(x) > t\} = \{x|f(x) > t\} \times \mathbb{R}^{d_2},$$

and the previous theorem. □

Proposition 0.8

Let L be a linear map $\mathbb{R}^d \to \mathbb{R}^d$, $E \subset \mathbb{R}^d$ a measurable set, then $L(E)$ is measurable, and

$$m(L(E)) = |\det L| m(E).$$

Proof. In fact we only need to prove it for cuboids E and elementary linear transformation L.

Now we only need to look at the case where $L = \begin{pmatrix} 1 & c & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$ since the other cases are trivial or similar to this case.

Thus by Fubini's theorem, WLOG E is the unit cube,

$$m(L(E)) = \int \chi_{L(E)} = \int_{\mathbb{R}^{d-1}} \left(\int_{\mathbb{R}} \chi_{L(E)}\,\mathrm{d}x_1 \right) = \int_{\mathbb{R}^{d-1}} \chi_{E'} \cdot 1 = 1 = |\det L| m(E),$$

where $E' = \{(x_2, \dots, x_n)|0 \leq x_i \leq 1\}$. □

From this transformation formula we deduce the integral version:

Let f be an integrable function on $\mathbb{R}^d$, then $f(L(x))$ is also integrable, and

$$\int f(L(x)) = \frac{1}{|\det L|} \int f(x).$$

Here we require $L \in \mathrm{GL}(n)$, since if $\det L = 0$, the function $f(L(x))$ need not be measurable.

At last we take a look at Fubini's theorem with the convolution product.

Definition 0.9 (Convolution)**.** Let f, g be smooth functions with compact support, define their **convolution** to be

$$f * g = \int f(x-y)g(y)\,\mathrm{d}y.$$

Then $f * g$ is also a smooth function with compact support.

In fact we can generalize this definition for $f, g \in L^1$.

First note that $f(x-y), g(y)$ are measurable functions on $\mathbb{R}^{2d}$, by Tonelli's theorem,

$$\int_{\mathbb{R}^{2d}} |f(x-y)||g(y)|\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f(x-y)||g(y)|\,\mathrm{d}x \right) \mathrm{d}y = \|f\|_{L^1}\|g\|_{L^1} < +\infty.$$

This shows that $f(x-y)g(y)$ is integrable on $\mathbb{R}^{2d}$. Hence by Fubini's theorem $f(x-y)g(y)$ is integrable as a function of y, and $f * g$ is integrable on $\mathbb{R}^d$.

Moreover we have

$$\|f * g\|_{L^1} \le \|f\|_{L^1}\|g\|_{L^1}.$$

The equality holds when both f and g are non-negative.

Fubini's theorem is also useful when computing integrals.

Example 0.10 (Gauss integral)

Recall the Gauss integral:

$$\int_0^{\infty} e^{-x^2}\,\mathrm{d}x = \frac{\sqrt{\pi}}{2}.$$

Here we give a different proof:

$$\begin{aligned} \int e^{-x^2}\,\mathrm{d}x \int e^{-y^2}\,\mathrm{d}y &= \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_0^{+\infty} e^{-r^2}\,\mathrm{d}r^2 \cdot \pi = \pi. \end{aligned}$$

§1 Lebesgue differentiation

The most important theorem in calculus is no doubt the Fundamental theorem of Calculus (which is also called Newton-Lebniz formula). Since we generalized the integrals, there must be a generalized version of this theorem:

Theorem 1.1 (Lebesgue differentiation theorem, part 1)

If f is integrable on $\mathbb{R}^d$, for any ball $B \subset \mathbb{R}^d$, we have

$$\lim_{x \in B, |B| \to 0} \frac{1}{m(B)} \int_B f(y)\,\mathrm{d}y = f(x), a.e.$$

This theorem clearly holds for continuous points of f.

Our basic idea is to take a continuous g, such that $\|g - f\|_{\mathcal{L}^1} < \varepsilon$.

and to prove

$$\left\{ x : \limsup_{x \in B, |B| \to 0} \frac{1}{m(B)} \int_B |f(y) - f(x)|\,\mathrm{d}y \ge \varepsilon_0 \right\}$$

is a null set.

Now we estimate

$$\begin{aligned}\frac{1}{m(B)}\int_B |f(y)-f(x)|\,\mathrm{d}y &\le \frac{1}{m(B)}\int_B \left(|f(y)-g(y)|+|g(y)-g(x)|+|g(x)-f(x)|\right)\mathrm{d}y \\ &= |f(x)-g(x)|+\varepsilon+\frac{1}{m(B)}\int_B |f(y)-g(y)|\,\mathrm{d}y\end{aligned}$$

We find that the last term is pretty hard to deal with, so we'll introduce some new tools:

Definition 1.2 (Hardy-Littlewood maximal function)**.** Let f be an integrable function on $\mathbb{R}^d$. Define

$$Mf(x)=\sup_{x\in B}\frac{1}{m(B)}\int_B |f(y)|\,\mathrm{d}y.$$

to be the **maximal function** of f.

Theorem 1.3 (Hardy-Littlewood)

The maximal function Mf satisfies:

- Mf is measurable;
- For x almost everywhere, $|f(x)|\le Mf(x)<+\infty$.
- There exists a constant C s.t.

$$|\{x: Mf>\alpha\}|\le \frac{C}{\alpha}\|f\|_{\mathcal{L}^1}.$$

Proof. First we prove $\{Mf>\alpha\}$ is measurable. If $Mf(x_0)>\alpha$, then exists an open ball $B\ni x_0$,

$$\int_B |f(y)|\,\mathrm{d}y>\alpha m(B).$$

This implies that $B\subset\{Mf>\alpha\}\implies\{Mf>\alpha\}$ is an open set.

For the second part, we'll prove for $\forall\varepsilon_0>0, N>0$,

$$m(\{x: Mf(x)+\varepsilon_0<|f(x)|\le N\})=0.$$

Otherwise denote the above set as E, for $\forall 0<\lambda<1$, $\exists B$ s.t. $|E\cap B|>\lambda|B|$.

Thus for $x\in E$,

$$\begin{aligned}Mf(x)&\ge\frac{1}{m(B)}\int_B |f(y)|\,\mathrm{d}y\\ &\ge\frac{1}{m(B)}\int_{E\cap B}|f(y)|\,\mathrm{d}y\\ &\ge\frac{1}{m(B)}\int_{E\cap B}\varepsilon_0+Mf(y)\,\mathrm{d}y\\ &=\frac{m(E\cap B)}{m(B)}\varepsilon_0+\frac{1}{|B|}\int_{E\cap B}Mf(y)\,\mathrm{d}y.\end{aligned}$$

Taking the integral with respect to x:

$$\left(1-\frac{|E\cap B|}{|B|}\right)\int_{E\cap B}Mf\ge\frac{|E\cap B|^2}{|B|}\varepsilon_0.$$

This implies $(1-\lambda)N \geq \lambda\varepsilon_0$, which is impossible as $\lambda \to 1$.

Now for the last part, since $\{Mf > \alpha\}$ is open, $\forall x \in \{Mf > \alpha\}$, $\exists B$ s.t.

$$\int_B |f(y)| \, \mathrm{d}y > \alpha m(B).$$

Hence for disjoint balls B_{i_k},

$$\|f\|_{\mathcal{L}^1} \geq \sum_{l=1}^{k} \int_{B_{i_l}} |f(y)| \, \mathrm{d}y > \alpha \sum_{l=1}^{k} |B_{i_l}|.$$

If we could select B_{i_l}'s such that their measure achieves say 1% of E, then we're done.

Lemma 1.4

Let $B_1, \ldots, B_n$ be open balls in $\mathbb{R}^d$. There exists $i_1, \ldots, i_k$ such that B_{i_j}'s are pairwise disjoint, and

$$\bigcup_{i=1}^{n} B_i \subset \bigcup_{j=1}^{k} 3B_{i_j}.$$

Here $3B$ means to multiply the radius of the ball by 3.

Proof of lemma. Trivial, just take the largest ball first and using greedy algorithm. □

Remark 1.5 — For countable many balls, the conculusion holds with 3 replaced by 5.

In particular, for all compact sets $K \subset \{Mf > \alpha\}$, there exists a finite open cover $B_1, B_2, \ldots, B_n$ of K. By lemma we can select B_{i_j}'s satisfying

$$\sum_{i=1}^{k} m(B_{i_j}) \geq \frac{1}{3^d} m\left(\bigcup_{i=1}^{n} B_i\right) \geq \frac{1}{3^d} m(K).$$

Combining with our previous discussion we get $\|f\|_{\mathcal{L}^1} \geq \frac{\alpha}{3^d} m(K)$.

□

Returning to the proof of Theorem 1.1, we can assmue g is continuous with compact support,

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

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

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

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

But

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

This completes the proof.

Definition 1.6 (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 1.7 — 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 1.8

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 Theorem 1.1. □

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 1.9. 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 1.10

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 1.11

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

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

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