Convolutions and Approximate Identities

Subhasish Mukherjee

April 5, 2022

Contents

Motivating Convolution

Suppose G is an Abelian group equipped with the counting measure and consider the vector space $\mathbb{C}[G] \cong L^1(G;\mathbb{C})$. As we have seen, this vector space is useful in representation theory of finite groups. We identify a natural basis in this space.

Proposition 1. *The set $\{x^g \mid g \in G\}$ is a basis for $\mathbb{C}[G]$.*

We also know $\mathbb{C}[G]$ is a ring (and in fact an algebra with the natural scalar multiplication), so given $r, s \in \mathbb{C}[G]$ it's a reasonable inquiry to wonder how multiplication interacts with this basis. Writing $r = \sum_{g\in G} r_g x^g$ and $s = \sum_{g\in G} s_g x^g$, we see that

$$r * s = \left(\sum_{g\in G} r_g x^g\right)\left(\sum_{h\in G} s_h x^h\right) = \left(\sum_{g\in G}\sum_{h\in G} r_g s_h x^{g+h}\right)$$

Thus we get the identity

$$(r*s)_h = \sum_{gk=h} r_g s_k = \sum_{g\in G} r_g s_{h-g}$$

Now we shift this perspective to $L^1(G;\mathbb{C})$ with the counting measure μ. Note originally this space is just a normed vector space and we do not have a natural product on this space. However we can use our analysis in $\mathbb{C}[G]$ to lift to a product on L^1. Given $r, s \in L^1(G;\mathbb{C})$ and $h \in G$, we can naturally define a new function $r * s$ given by

$$(r*s)(h) = \int_G r(g)s(h-g)\mathrm{d}\mu.$$

We also know representations of finite groups correspond exactly to $\mathbb{C}[G]$ modules. This provides us a starting point to begin the study of convolution.

This idea holds in much greater generality. A similar construction holds for any locally compact Hausdorff topological group. There is an analog for Lebesgue measure on any locally compact Hausdorff group called the Haar measure and we can define a similar algebra on the continuous compactly supported functions. In fact, there is a nice bijection between the unitary representations of G and unitary representations of $C_c(G;\mathbb{C})$ much in the same way $\mathbb{C}[G]$ modules are exactly representations of G.

Convolution on $\mathbb{R}^n$

Defining Convolution

With some of our motivation in hand we define convolution on $\mathbb{R}^n$.

Definition 2. Suppose we have f, g which are measurable functions from $\mathbb{R}^n \to \mathbb{R}$. Then define $f * g : \mathbb{R}^n \to \mathbb{R}$ via

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-t)g(t)\, \mathrm{d}t$$

Of course, we still don't know if or when this is well-defined. The next proposition gives us our first foothold in analyzing convolution further.

Proposition 3. *If $f \in L^1(\mathbb{R}^n)$ and $g \in L^1(\mathbb{R}^n)$, then $f * g$ is defined almost everywhere and $f * g \in L^1(\mathbb{R}^n)$ with the inequality*

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1$$

Before we prove this, we need a quick lemma to show that $f * g$ is even measurable.

Lemma 4. *If $f : \mathbb{R}^n \to \mathbb{R}$ is measurable, then the function $F : \mathbb{R}^{2n} \to \mathbb{R}$ given by $F(x,t) = f(x-t)$ is measurable as well.*

Proof. We let $F_1(x,t) = f(x)$. Since f is measurable, we easily see that F_1 is measurable as well as we have $\{F_1 > \alpha\} = \{f > \alpha\} \times \mathbb{R}^n$. Then we note that the transformation $(x,t) \mapsto (x-t, x)$ is linear and thus the function $(x,t) \mapsto F_1(x-t, x)$ is measurable and $F_1(x-t,x) = f(x-t) = F(x,t)$. ■

Proof (Proposition 3). We prove the result for non-negative f and g, with the general case following by noting that $|f * g| \leq |f| * |g|$. We see that $f(x-t)g(t)$ is measurable in $\mathbb{R}^{2n}$ since it is the product of two measurable functions. Thus the desired double integral is well-defined, and by Tonelli's theorem, we have

$$\begin{aligned}\iint f(x-t)g(t)\, \mathrm{d}t\, \mathrm{d}x &= \int g(t) \left(\int f(x-t) \mathrm{d}x \right) \mathrm{d}t \\ &= \int g(t) \|f\|_1 \,\mathrm{d}t = \|f\|_1 \|g\|_1\end{aligned}$$

■

Young's Convolution Inequality

The main result of this section will be the following theorem.

Theorem 5 (Young's Convolution Inequality). *Suppose we have $p, q, r \in [1, \infty]$ satisfying the relation $\frac{1}{p} + \frac{1}{q} - \frac{1}{r} = 1$. Then given $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$, the convolution $f * g \in L^r(\mathbb{R}^n)$ and we have the inequality*

$$\|f * g\|_r \leq \|f\|_p \|g\|_q$$

Remark 6. In fact, this bound can be improved to a bound $\|f * g\|_r \leq C_{p,q} \|f\|_p \cdot \|g\|_q$ with the constant $C_{p,q}$ strictly less than 1 for $p > 1$.

The theorem is not particularly difficult, but it will require a few steps. First we prove the generalized Hölder's inequality.

Proposition 7 (Generalized Hölder's Inequality)**.** *Suppose we have* $p_1, \dots, p_n \in [1, \infty]$ *satisfying* $\sum_{i=1}^n \frac{1}{p_i} = 1$ *and for* $1 \le i \le n$ *we have* $f_i \in L^{p_i}(\mathbb{R}^n)$ *satisfying* $f_i : \mathbb{R}^n \to \mathbb{R}$*. Then* $\prod_{i=1}^n f_i \in L^1(\mathbb{R}^n)$ *and*

$$\left\| \prod_{i=1}^n f_i \right\|_1 \le \prod_{i=1}^n \|f_i\|_{p_i}$$

Proof. We proceed by induction, with the $n = 2$ case already happily being handled. Suppose we have shown the result holds for $n-1$ functions. Define $q = \frac{p_n}{p_n - 1}$ and for $i < n$ we scale the p_i-s in a form suitable for induction and define $r_i = \frac{p_i}{q}$. Then we have $\frac{1}{p_n} + \frac{1}{q} = 1$ and $\sum_{i=1}^{n-1} \frac{1}{r_i} = 1$. Applying Hölder's inequality to $\prod_{i=1}^{n-1} f_i$ and f_n, we get that

$$\begin{aligned} \int \left| \prod_{i=1}^n f_i \right| &\le \|f_n\|_{p_n} \left\| \prod_{i=1}^{n-1} f_i \right\|_q \\ &\le \|f_n\|_{p_n} \left(\prod_{i=1}^{n-1} \left\| f_i^q \right\|_{r_i} \right)^{\frac{1}{q}} \\ &= \prod_{i=1}^n \|f_i\|_{p_i} \end{aligned}$$

■

We now proceed with our proof for Young's inequality.

Proof (Theorem 5). We start by bounding $|(f * g)(x)|$. We see we have by the generalized Hölder's inequality:

$$\begin{aligned} |(f*g)(x)| &\le \int |f(x-t)|\,|g(t)|\,\mathrm{d}t \\ &= \int |f(x-t)|^{p/r}\,|g(t)|^{q/r}\,|f(x-t)|^{1-p/r}\,|g(t)|^{1-q/r}\,\mathrm{d}t \\ &= \int \left(|f(x-t)|^p\,|g(t)|^q\right)^{\frac{1}{r}} |f(x-t)|^{\frac{r-p}{r}}\,|g(t)|^{\frac{r-q}{r}}\,\mathrm{d}t \\ &\le \underbrace{\left\| \left(f(x-\cdot)^p g(\cdot)^q\right)^{\frac{1}{r}} \right\|_r}_{(1)} \underbrace{\left\| f(x-\cdot)^{\frac{r-p}{r}} \right\|_{\frac{pr}{r-p}}}_{(2)} \underbrace{\left\| g(\cdot)^{\frac{r-q}{r}} \right\|_{\frac{qr}{r-q}}}_{(3)} \end{aligned}$$

We evaluate each separately and justify the use of Hölder's inequality. For (1) we see

$$\begin{aligned} \left\| \left(f(x-\cdot)^p g(\cdot)^q\right)^{\frac{1}{r}} \right\|_r &= \left(\int \left(\left| f(x-t)^p g(t)^q \right|^{\frac{1}{r}} \right)^r \mathrm{d}t \right)^{1/r} \\ &= \left(\int \left| f(x-t)^p g(t)^q \right| \mathrm{d}t \right)^{\frac{1}{r}} \end{aligned}$$

This is finite by Proposition 3. For (2) we have

$$\left\| f(x-\cdot)^{\frac{r-p}{r}} \right\|_{\frac{pr}{r-p}} = \left(\int \left(|f(x-t)|^{\frac{r-p}{r}} \right)^{\frac{pr}{r-p}} \mathrm{d}t \right)^{\frac{r-p}{pr}} = \|f\|_p^{\frac{r-p}{r}}$$

This is finite because $f \in L^p$. And finally for (3) we have

$$\left\| g(\cdot)^{\frac{r-q}{r}} \right\|_{\frac{qr}{r-q}} = \left(\int \left(|g(t)|^{\frac{r-q}{r}} \right)^{\frac{qr}{r-q}} \mathrm{d}t \right)^{\frac{r-q}{qr}} = \|g\|_q^{\frac{r-q}{r}}$$

and this is finite because $g \in L^q$.

With our bounds and simplifications in hand, we move to our main proof:

$$\begin{aligned}
\|f*g\|_r^r &= \int |(f*g)(x)|^r \,\mathrm{d}x \\
&\le \int \left(\|f\|_p^{\frac{r-p}{r}} \|g\|_q^{\frac{r-q}{r}} \left(\int \left| f(x-t)^p g(t)^q \right| \mathrm{d}t \right)^{\frac{1}{r}} \right)^r \mathrm{d}x \\
&= \|f\|_p^{r-p} \|g\|_q^{r-q} \iint \left| f(x-t)^p g(t)^q \right| \mathrm{d}t\,\mathrm{d}x \\
&= \|f\|_p^{r-p} \|g\|_q^{r-q} \int |g(t)|^q \int |f(x-t)|^p \,\mathrm{d}x\,\mathrm{d}t \\
&= \|f\|_p^{r-p} \|g\|_q^{r-q} \int |g(t)|^q \|f\|_p^p \,\mathrm{d}t \\
&= \|f\|_p^{r-p} \|g\|_q^{r-q} \|f\|_p^p \|g\|_q^q = \|f\|_p^r \|g\|_q^r
\end{aligned}$$

■

Derivatives of Convolutions

Definition 8. Given a fixed function $K : \mathbb{R}^n \to \mathbb{R}$, we define the operator $T_K : f \mapsto f * K$. In this case, T_K is called the *convolution operator with kernel K*.

Setting $q = 1$ in Theorem 5 shows that for $K \in L^1(\mathbb{R}^n)$, T_K is a linear operator on $L^p(\mathbb{R}^n)$. In the following result, we recall that $C_c^m(\mathbb{R}^n)$ is the space of C^m functions with compact support on $\mathbb{R}^n$. We note that $C_c^m(\mathbb{R}^n) \subseteq L^1(\mathbb{R}^n)$.

Proposition 9. *Suppose $p \in [1, \infty]$ and $m \in \mathbb{N} \cup \{\infty\}$ with $f \in L^p(\mathbb{R}^n)$ and $K \in C_c^m(\mathbb{R}^n)$. Then we have that $f * K \in C^m$ and the derivatives are given by*

$$D^\alpha(f*K)(x) = (f * D^\alpha K)(x)$$

for $\alpha \in \mathbb{N}^n$ with $\sum_{i=1}^n \alpha_i \le m$.

Proof. We first show the case where $m = 0$. For $h \in \mathbb{R}^n$, we compute directly:

$$\begin{aligned}
|(f*K)(x+h) - (f*K)(x)| &= \left| \int f(t)K(x+h-t)\,\mathrm{d}t - \int f(t)K(x-t)\,\mathrm{d}t \right| \\
&= \left| \int f(x-t)\left(K(t+h) - K(t)\right)\mathrm{d}t \right| \\
&\le \|f(x-\cdot)\|_p \|K(\cdot+h) - K(\cdot)\|_{p'} \\
&= \|f\|_p \|K(\cdot+h) - K(\cdot)\|_{p'}
\end{aligned}$$

which indeed goes to 0 as $h \to 0$ since K is uniformly continuous and compactly supported.

Remark 10. Essentially the same technique can be used to show that if $K \in UC_b(\mathbb{R}^n)$ then $f * K \in UC_b(\mathbb{R}^n)$.

Now suppose $K \in C_c^1$. Fix $i \in \{1, \ldots, n\}$ and let $h \in \mathbb{R}$. Then we see that

$$\begin{aligned}
\frac{(f*K)(x+he_i) - (f*K)(x)}{h} &= \int f(t)\left(\frac{K(x-t+he_i) - K(x-t)}{h}\right)\mathrm{d}t \\
&= \int f(t)\partial_i K(x-t+h'(t)e_i)\mathrm{d}t
\end{aligned}$$

for some $h'(t)$ between 0 and h by the mean value theorem. Then using the dominated convergence theorem, we get the desired result and taking limits on both sides we have the desired result. Higher order partials hold by an inductive argument. ■

Approximations of the Identity

As we discussed, convolution turns $L^1(\mathbb{R}^n)$ into a Banach algebra. However, we have a problem - there's no identity element! We turn our attention back to the discrete group from before. There we see the element x^0 functions as an identity element since $x^0 * r = r$ for all $r \in \mathbb{C}[G]$. Shifting perspective to $L^1(G)$, this corresponds to the function $r : G \to \mathbb{C}$ such that $r(0) = 1$ and $r(g) = 0$ for all $g \neq 0$ - i.e. a function that looks like a "Dirac delta". There is no immediately obvious analog of this in $\mathbb{R}^n$, however what we do get are *approximate identities*, i.e. families of functions which look like identities in the limit.

K_ε as approximations of the identity

Definition 11. Given a kernel $K : \mathbb{R}^n \to \mathbb{R}$ and $\varepsilon > 0$, we define $K_\varepsilon : \mathbb{R}^n \to \mathbb{R}$ via

$$K_\varepsilon(x) = \frac{1}{\varepsilon^n} K\left(\frac{x}{\varepsilon}\right)$$

Lemma 12. *Given $K \in L^1(\mathbb{R}^n)$ and $\varepsilon > 0$ we have $\int K_\varepsilon = \int K$. Furthermore, for fixed $\delta > 0$, we have*

$$\int_{|x|>\delta} |K_\varepsilon| \to 0 \text{ as } \varepsilon \to 0$$

Remark 13. The first condition in this lemma shows that the areas under the graph of K_ε and K are the same. The second shows that as ε goes to 0, most of the mass of the function is concentrated around 0, thus we are indeed getting something like a Dirac delta. Note we also have $K_\varepsilon \in L^p \Leftrightarrow K \in L^p$.

Proof. The first equality follows immediately by the change of variables $x \mapsto \frac{x}{\varepsilon}$. To show we have the desired limit, fix $\delta > 0$ and let $y = \frac{x}{\varepsilon}$. Then we have

$$\int_{|x|>\delta} |K_\varepsilon| = \int_{|x|>\delta} \left| K\left(\frac{x}{\varepsilon}\right)\right| \frac{1}{\varepsilon^n} \mathrm{d}x = \int_{|y|>\frac{\delta}{\varepsilon}} |K(y)| \, \mathrm{d}y$$

We know $K \in L^1(\mathbb{R}^n)$ and $\frac{\delta}{\varepsilon} \to \infty$ as $\varepsilon \to 0$, thus the last integral indeed tends to 0, completing the proof. ∎

The next few results show that for any $K \in L^1(\mathbb{R}^n)$, the family $\{K_\varepsilon\}$ indeed behaves like an approximate identity and we will have $f * K_\varepsilon \to f$ as $\varepsilon \to 0$ in various senses. First we establish that $\{K_\varepsilon\}$ is an approximate identity in L^p. To do this, we first need a helpful lemma.

Lemma 14. *If $f \in L^p(\mathbb{R}^n)$ for $p \in [1, \infty)$, then $\lim_{h\to 0} \|f(\cdot + h) - f(\cdot)\|_p = 0$.*

Remark 15. Note that this does not hold for $p = \infty$ - consider $f = \chi_{(0,1)} \in L^\infty(\mathbb{R})$. Then for every $0 < h < 1$, $\|f(\cdot + h) - f(\cdot)\|_\infty = 1$.

Proof. Consider the set

$$C_p = \left\{ f \in L^p(\mathbb{R}^n) \mid \lim_{h\to 0} \|f(\cdot + h) - f(\cdot)\|_p = 0 \right\}$$

We claim that (a) C_p is a vector space and (b) C_p is closed in L^p. (a) follows easily from Minkowski's inequality, as if we have $f, g \in C_p$, then

$$\begin{aligned} &\limsup_{h\to 0} \|(\alpha f + \beta g)(\cdot + h) - (\alpha f + \beta g)(\cdot)\|_p \\ \leq\ &\limsup_{h\to 0} |\alpha| \, \|f(\cdot + h) - f(\cdot)\|_p + |\beta| \, \|g(\cdot + h) - g(\cdot)\|_p \end{aligned}$$

(b) also follows easily from Minkowski's inequality. Suppose we have a sequence $\{f_k\}_{k=0}^{\infty} \subseteq C_p$ with $f_k \to f$ in $L^p(\mathbb{R}^n)$. Then

$$\begin{aligned}\|f(\cdot+h) - f(\cdot)\|_p &\leq \|f(\cdot+h) - f_k(\cdot+h)\|_p + \|f_k(\cdot+h) - f_k(\cdot)\|_p + \|f_k(\cdot) - f(\cdot)\|_p \\ &= \|f_k(\cdot+h) - f_k(\cdot)\|_p + 2\|f_k - f\|_p\end{aligned}$$

Taking $\limsup_{h\to 0}$ of both sides and letting $k \to \infty$ gives the desired result. Now we note that the characteristic function of any cube belongs to C_p. We know step functions are dense in $L^p(\mathbb{R}^n)$, thus it follows by (b) that C_p contains the step functions and is dense in $L^p(\mathbb{R}^n)$ and thus equals $L^p(\mathbb{R}^n)$ since it is closed. ■

Theorem 16 (Approximate Identity in L^p)**.** *Suppose $K \in L^1(\mathbb{R}^n)$ and $\int K = 1$. If $f \in L^p(\mathbb{R}^n)$ for $p \in [1, \infty)$, then*

$$\|(f * K_\varepsilon) - f\|_p \to 0 \text{ as } \varepsilon \to 0$$

Proof. By Lemma 12, we note that we have for all $x \in \mathbb{R}^n$ that

$$f(x) = \int f(x) K_\varepsilon(t) \mathrm{d}t$$

Thus we have

$$\begin{aligned}|(f * K_\varepsilon)(x) - f(x)| &= \left| \int (f(x-t) - f(x))\, K_\varepsilon(t)\, \mathrm{d}t \right| \\ &\leq \int |f(x-t) - f(x)|\, |K_\varepsilon(t)|^{\frac{1}{p}}\, |K_\varepsilon(t)|^{\frac{p-1}{p}}\, \mathrm{d}t \\ &\leq \left\| (f(x-\cdot) - f(x)) K_\varepsilon(\cdot)^{\frac{1}{p}} \right\|_p \left\| K_\varepsilon(\cdot)^{\frac{p-1}{p}} \right\|_{\frac{p}{p-1}}\end{aligned}$$

Where the last line followed by Hölder's inequality. Then exponentiating, expanding, and integrating, we have

$$\begin{aligned}\int |(f * K_\varepsilon)(x) - f(x)|^p\, \mathrm{d}x &\leq \int \left(\int |f(x-t) - f(x)|^p\, |K_\varepsilon(t)|\, \mathrm{d}t \right) \left(\int |K_\varepsilon(t)|\, \mathrm{d}t \right)^{p-1} \mathrm{d}x \\ &= \|K\|_1^{p-1} \iint |f(x-t) - f(x)|^p\, |K_\varepsilon(t)|\, \mathrm{d}t\, \mathrm{d}x \\ &= \|K\|_1^{p-1} \int \|f(\cdot - t) - f(\cdot)\|_p^p\, |K_\varepsilon(t)|\, \mathrm{d}t\end{aligned}$$

We define $\varphi(t) = \|f(\cdot - t) - f(\cdot)\|_p^p$. By Minkowski's inequality, we can note that $|\varphi(t)| \leq 2^p \|f\|_p^p$, thus φ is bounded. By Lemma 14, we have that $\varphi(t) \to 0$ as $t \to 0$, so for fixed $\eta > 0$ we can pick some $\delta > 0$ such that $\varphi(t) < \eta$ if $|t| < \delta$. Thus we have

$$\begin{aligned}\|(f * K_\varepsilon) - f\|_p^p &\leq \|K\|_1^{p-1} \left(\int_{|t|<\delta} \varphi(t)\, |K_\varepsilon(t)|\, \mathrm{d}t + \int_{|t|\geq\delta} \varphi(t)\, |K_\varepsilon(t)|\, \mathrm{d}t \right) \\ &\leq \|K\|_1^{p-1} \left(\eta \|K\|_1 + \|\varphi\|_\infty \int_{|t|\geq\delta} |K_\varepsilon(t)|\, \mathrm{d}t \right)\end{aligned}$$

As $\varepsilon \to 0$, we can make the right-hand side arbitrarily small, so we are done! ■

The previous theorem does not hold when $p = \infty$, so this next theorem is an analog for that result for L^∞.

Theorem 17 (Approximate Identity in L^∞)**.** *Again suppose $K \in L^1(\mathbb{R}^n)$ with $\int K = 1$ and let $f \in L^\infty(\mathbb{R}^n)$. Then $(f * K_\varepsilon) \to f$ as $\varepsilon \to 0$ at every point of continuity of f, and the convergence is uniform on any set where f is uniformly continuous.*

Proof. Note that $(f * K_\varepsilon)(x)$ converges absolutely for all $x \in \mathbb{R}^n$ as $f \in L^\infty$ and $K \in L^1$. As in the previous theorem, we have

$$|(f * K_\varepsilon)(x) - f(x)| \leq \int |f(x-t) - f(x)| \, |K_\varepsilon(t)| \, \mathrm{d}t$$

Then if f is continuous at x, there exists some $\delta > 0$ such that $|f(x-t) - f(x)| < \eta$ if $|t| < \delta$. Thus we get the similar bound

$$\int |f(x-t) - f(x)| \, |K_\varepsilon(t)| \, \mathrm{d}t \leq \eta \, \|K\|_1 + 2 \, \|f\|_\infty \int_{|t| \geq \delta} |K_\varepsilon(t)| \, \mathrm{d}t$$

and again we see we can make the right-hand side arbitrarily small. If f is uniformly continuous, δ can be chosen independently of f so we get uniform convergence as desired. ∎

We care about the convergence of $(f * K_\varepsilon)(x)$ here since we are dealing with the function itself, and we can't get away with just thinking of the integral as in the last theorem.

If we have more regularity on K, we can get the previous result for generic L^p as well.

Theorem 18. *Suppose $f \in L^p(\mathbb{R}^n)$ for $p \in [1, \infty)$ and $K \in L^1(\mathbb{R}^n) \cap L^{\frac{p}{p-1}}(\mathbb{R}^n)$ with $\int K = 1$ and $K(x) = o\left(|x|^{-n}\right)$ as $|x| \to \infty$.*[2] *Then $(f * K_\varepsilon) \to f$ as $\varepsilon \to 0$ at each point of continuity of f.*

[2] Recall we say $f(x) = o(g(x))$ as $x \to x_0$ if $\lim_{x \to x_0} f(x)/g(x) = 0$.

Proof. First we note by Hölder's inequality that $(f * K_\varepsilon)(x)$ converges absolutely for each x because $f \in L^p$ and $K_\varepsilon \in L^{\frac{p}{p-1}}$. Using a bound analogous to that in Theorem 16, we have

$$|f_\varepsilon(x) - f(x)|^p \leq \left(\int |f(x-t) - f(x)|^p \, |K_\varepsilon(t)| \, \mathrm{d}t \right) \|K_\varepsilon\|_1^{p-1}$$

Then if f is continuous at x, for fixed $\eta > 0$ we can pick some $\delta > 0$ such that $|f(x-t) - f(x)| < \eta$ if $|t| < \delta$. Thus we have

$$\begin{aligned} |f_\varepsilon(x) - f(x)|^p &\leq \|K_\varepsilon\|_1^{p-1} \left(\eta^p \int_{|t| < \delta} |K_\varepsilon(t)| \, \mathrm{d}t + \int_{|t| \geq \delta} |f(x-t) - f(x)| \, |K_\varepsilon(t)| \, \mathrm{d}t \right) \\ &\leq \|K_\varepsilon\|_1^{p-1} \left(\eta^p \, \|K_\varepsilon\|_1 + \int_{|t| \geq \delta} |f(x-t) K_\varepsilon(t)| \, \mathrm{d}t + |f(x)| \int_{|t| \geq \delta} |K_\varepsilon(t)| \, \mathrm{d}t \right) \end{aligned}$$

The first and third terms we can bound as before. For the second term, we first write $K(t) = s(t) \, |t|^{-n}$ where $s(t) \to 0$ as $t \to \infty$. Then

$$\begin{aligned} \int_{|t| \geq \delta} |f(x-t) K_\varepsilon(t)| \, \mathrm{d}t &= \int_{|t| \geq \delta} |f(x-t)| \, s\left(\frac{t}{\varepsilon}\right) |t|^{-n} \, \mathrm{d}t \\ &\leq \delta^{-n} \left(\sup_{|t| \geq \delta} s\left(\frac{t}{\varepsilon}\right) \right) \int_{|t| \geq \delta} |f(x-t)| \, \mathrm{d}t \\ &\leq \delta^{-n} \left(\sup_{|t| \geq \delta} s\left(\frac{t}{\varepsilon}\right) \right) \|f\|_1 \end{aligned}$$

This also goes to 0 as $\varepsilon \to 0$, thus the theorem follows. ∎

Remark 19. In fact, we can get something even more spectacular - if K is bounded as well and we have $K(x) = O(|x|^{-n-\lambda})$ for some $\lambda > 0$, then $(f * K_\varepsilon) \to f$ at each Lebesgue point of f, i.e. almost everywhere!

In summary, we have turned L^1 into a Banach algebra and shown many ways how convolution products interact in L^p spaces. We've also shown how these families K_ε approximate the identity in convolutions. Neat!

References

The main reference for these notes was *Measure and Integral* by Wheeden and Zygmund. The motivating convolution section was inspired by Professor Tice's notes on Applied Harmonic Analysis.