Bernstein’s inequality and Nikolsky’s inequality for $\mathbb{R}^d$

Jordan Bell
ejordan.bell@gmail.com
Department of Mathematics, University of Toronto
February 16, 2015

1 Complex Borel measures and the Fourier transform

Let $M(\mathbb{R}^d) = rca(\mathbb{R}^d)$ be the set of complex Borel measures on $\mathbb{R}^d$. This is a Banach algebra with the total variation norm, with convolution as multiplication; for $\mu \in M(\mathbb{R}^d)$, we denote by $|\mu|$ the total variation of $\mu$, which itself belongs to $M(\mathbb{R}^d)$, and the total variation norm of $\mu$ is $||\mu|| = |\mu|(\mathbb{R}^d)$.

For $\mu \in M(\mathbb{R}^d)$, it is a fact that the union $O$ of all open sets $U \subset \mathbb{R}^d$ such that $|\mu|(U) = 0$ itself satisfies $|\mu|(O) = 0$. We define $\text{supp} \mu = \mathbb{R}^d \setminus O$, called the support of $\mu$.

For $\mu \in M(\mathbb{R}^d)$, we define $\hat{\mu}: \mathbb{R}^d \to \mathbb{C}$ by

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i \xi \cdot x} d\mu(x), \quad \xi \in \mathbb{R}^d.$$ 

It is a fact that $\hat{\mu}$ belongs to $C_0(\mathbb{R})$, the collection of bounded uniformly continuous functions $\mathbb{R}^d \to \mathbb{C}$. For $\xi \in \mathbb{R}^d$,

$$|\hat{\mu}(\xi)| \leq \int_{\mathbb{R}^d} |e^{-2\pi i \xi \cdot x}| d|\mu|(x) = |\mu|(\mathbb{R}^d) = ||\mu||. \quad (1)$$

Let $m_d$ be Lebesgue measure on $\mathbb{R}^d$. For $f \in L^1(\mathbb{R}^d)$, let

$$\Lambda f = f m_d,$$

which belongs to $M(\mathbb{R}^d)$. We define $\hat{f}: \mathbb{R}^d \to \mathbb{C}$ by

$$\hat{f}(\xi) = \Lambda \hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i \xi \cdot x} d\Lambda f(x) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i \xi \cdot x} dm_d(x), \quad \xi \in \mathbb{R}^d.$$

The following theorem establishes properties of the Fourier transform of a complex Borel measure with compact support.\footnote{Thomas H. Wolff, Lectures on Harmonic Analysis, p. 3, Proposition 1.3.}
**Theorem 1.** If \( \mu \in \mathcal{M}(\mathbb{R}^d) \) and \( \text{supp} \mu \) is compact, then \( \hat{\mu} \in C^\infty(\mathbb{R}^d) \) and for any multi-index \( \alpha \),
\[
D^\alpha \hat{\mu} = \mathcal{F}((-2\pi ix)^\alpha \mu).
\]

For \( R > 0 \), if \( \text{supp} \mu \subset B(0,R) \), then
\[
\|D^\alpha \hat{\mu}\|_{\infty} \leq (2\pi R)^{|\alpha|} \|\mu\|.
\]

**Proof.** For \( j = 1, \ldots, d \), let \( e_j \) be the \( j \)th coordinate vector in \( \mathbb{R}^d \), with length 1. Let \( \xi \in \mathbb{R}^d \), and define
\[
\Delta(h) = \hat{\mu}(\xi + he_j) - \hat{\mu}(\xi), \quad h \neq 0.
\]

We can write this as
\[
\Delta(h) = \int_{\mathbb{R}^d} \frac{e^{-2\pi ihx_j} - 1}{h} e^{-2\pi ix_j} d\mu(x).
\]

For any \( x \in \mathbb{R}^d \),
\[
\left| \frac{e^{-2\pi ihx_j} - 1}{h} \right| = \left| \frac{e^{-2\pi ihx_j} - 1}{|h|} \right| \leq \frac{|-2\pi ihx_j|}{|h|} = 2\pi |x_j|.
\]

Because \( \mu \) has compact support, \( 2\pi|x_j| \in L^1(\mu) \). Furthermore, for each \( x \in \mathbb{R}^d \) we have
\[
\frac{e^{-2\pi ihx_j} - 1}{h} \to -2\pi ix_j, \quad h \to 0.
\]

Therefore, the dominated convergence theorem tells us that
\[
\lim_{h \to 0} \Delta(h) = \int_{\mathbb{R}^d} -2\pi ix_j e^{-2\pi ix_j} d\mu(x).
\]

On the other hand, for \( \alpha_k = 1 \) for \( k = j \) and \( \alpha_k = 0 \) otherwise,
\[
(D^\alpha \hat{\mu})(\xi) = \lim_{h \to 0} \Delta(h),
\]
so
\[
(D^\alpha \hat{\mu})(\xi) = \int_{\mathbb{R}^d} (-2\pi ix)^\alpha e^{-2\pi ix} d\mu(x) = \mathcal{F}((-2\pi ix)^\alpha \mu)(\xi),
\]
and in particular, \( \hat{\mu} \in C^1(\mathbb{R}^d) \). (The Fourier transform of a regular complex Borel measure on a locally compact abelian group is bounded and uniformly continuous.)

Because \( \mu \) has compact support so does \((-2\pi ix)^\alpha \mu\), hence we can play the above game with \((-2\pi ix)^\alpha \mu\), and by induction it follows that for any \( \alpha \),
\[
D^\alpha \hat{\mu} = \mathcal{F}((-2\pi ix)^\alpha \mu),
\]

---

\(^2\)Walter Rudin, *Fourier Analysis on Groups*, p. 15, Theorem 1.3.3.
and in particular, $\hat{\mu} \in C^\infty(\mathbb{R}^d)$.

Suppose that $\text{supp } \mu \subset B(0, R)$. The total variation of the complex measure $(-2\pi ix)^\alpha \mu$ is the positive measure

$$(2\pi)^{|\alpha|} |x_1|^{\alpha_1} \cdots |x_d|^{\alpha_d} |\mu|,$$

hence

$$\|(-2\pi ix)^\alpha \mu\| = (2\pi)^{|\alpha|} \int_{\mathbb{R}^d} |x_1|^{\alpha_1} \cdots |x_d|^{\alpha_d} d|\mu|(x)$$

$$\leq (2\pi)^{|\alpha|} \int_{B(0, R)} R^{\alpha_1} \cdots R^{\alpha_d} d|\mu|(x)$$

$$= (2\pi R)^{|\alpha|} \int_{B(0, R)} d|\mu|(x)$$

$$= (2\pi R)^{|\alpha|} \int_{B(0, R)} d|\mu|(x)$$

$$= (2\pi R)^{|\alpha|} \|\mu\|.$$

Then using (1),

$$\|\mathcal{F}((-2\pi ix)^\alpha \mu)\|_\infty \leq \|(-2\pi ix)^\alpha \mu\| \leq (2\pi R)^{|\alpha|} \|\mu\|.$$
Lemma 2. Suppose that \( f \in C^N(\mathbb{R}^d) \) and \( D^\alpha f \in L^1(\mathbb{R}^d) \) for each \(|\alpha| \leq N\). Then for each \(|\alpha| \leq N\), \( D^\alpha(\phi_k f) \to D^\alpha f \) in \( L^1(\mathbb{R}^d) \) as \( k \to \infty \).

Proof. Let \(|\alpha| \leq N\). In the case \( \alpha = 0 \),

\[
\|\phi_k f - f\|_1 = \int_{\mathbb{R}^d} |\phi_k(x)f(x) - f(x)|dx
\]

\[
= \int_{|x| \geq k} |\phi_k(x)f(x) - f(x)|dx
\]

\[
\leq \int_{|x| \geq k} |f(x)|dx.
\]

Because \( f \in L^1(\mathbb{R}^d) \), this tends to 0 as \( k \to \infty \).

Suppose that \( \alpha > 0 \). The Leibniz rule tells us that with \( c_\beta = \binom{\alpha}{\beta} \), we have, for each \( k \),

\[
D^\alpha(\phi_k f) = \phi_k D^\alpha f + \sum_{0 < \beta \leq \alpha} c_\beta D^{\alpha-\beta} f D^\beta \phi_k.
\]

For \( C_1 = \max_\beta |c_\beta| \),

\[
\|D^\alpha(\phi_k f) - \phi_k D^\alpha f\|_1 \leq \sum_{0 < \beta \leq \alpha} \|c_\beta D^{\alpha-\beta} f D^\beta \phi_k\|_1
\]

\[
\leq C_1 \sum_{0 < \beta \leq \alpha} \|D^\beta \phi_k\|_\infty \|D^{\alpha-\beta} f\|_1.
\]

Let \( C_2 = \max_{0 < \beta \leq \alpha} \|D^\beta \phi\|_\infty \). By (2), for \( 0 < \beta \leq \alpha \) we have

\[
\|D^\beta \phi_k\|_\infty = k^{-|\beta|_1} \|D^\beta \phi\|_\infty \leq C_2 k^{-|\beta|_1} \leq C_2 k^{-1}.
\]

Thus

\[
\|D^\alpha(\phi_k f) - \phi_k D^\alpha f\|_1 \leq C_1 C_2 k^{-1} \sum_{0 < \beta \leq \alpha} \|D^{\alpha-\beta} f\|_1,
\]

which tends to 0 as \( k \to \infty \). For any \( k \),

\[
\|\phi_k D^\alpha f - D^\alpha f\|_1 = \int_{\mathbb{R}^d} |\phi_k(x)(D^\alpha f)(x) - (D^\alpha f)(x)|dx
\]

\[
= \int_{|x| \geq k} |\phi_k(x)(D^\alpha f)(x) - (D^\alpha f)(x)|dx
\]

\[
\leq \int_{|x| \geq k} |(D^\alpha f)(x)|dx,
\]

and because \( D^\alpha f \in L^1(\mathbb{R}^d) \), this tends to 0 as \( k \to \infty \). But

\[
\|D^\alpha(\phi_k f) - D^\alpha f\|_1 \leq \|D^\alpha(\phi_k f) - \phi_k D^\alpha f\|_1 + \|\phi_k D^\alpha f - D^\alpha f\|_1,
\]

which completes the proof. \( \square \)
Now we calculate the Fourier transform of the derivative of a function, and show that the smoother a function is the faster its Fourier transform decays.\(^4\)

**Theorem 3.** If \( f \in C^N(\mathbb{R}^d) \) and \( D^\alpha f \in L^1(\mathbb{R}^d) \) for each \(|\alpha| \leq N\), then for each \(|\alpha| \leq N\),

\[
\hat{D^\alpha f}(\xi) = (2\pi i)^\alpha \hat{f}(\xi), \quad \xi \in \mathbb{R}^d.
\]  

There is a constant \( C = C(f, N) \) such that

\[
|\hat{f}(\xi)| \leq C(1 + |\xi|)^{-N}, \quad \xi \in \mathbb{R}^d.
\]

**Proof.** If \( g \in C^1_c(\mathbb{R}^d) \), then for any \( 1 \leq j \leq d \), integrating by parts,

\[
\int_{\mathbb{R}^d} (\partial_j g)(x)e^{-2\pi i \xi \cdot x} dx = 2\pi i \xi_j \int_{\mathbb{R}^d} g(x)e^{-2\pi i \xi \cdot x} dx.
\]

It follows by induction that if \( g \in C^N_c(\mathbb{R}^d) \), then for each \(|\alpha| \leq N\),

\[
\hat{D^\alpha g}(\xi) = (2\pi i)^\alpha \hat{g}(\xi), \quad \xi \in \mathbb{R}^d.
\]

Let \(|\alpha| \leq N\). For \( k = 1, 2, \ldots \), let \( f_k = \phi_k f \). For each \( k \) we have \( f_k \in C^N(\mathbb{R}^d) \), hence

\[
\hat{D^\alpha f_k}(\xi) = (2\pi i)^\alpha \hat{f_k}(\xi), \quad \xi \in \mathbb{R}^d.
\]

On the one hand,

\[
\left\| \hat{D^\alpha f_k} - \hat{D^\alpha f} \right\|_\infty = \| \mathcal{F}(D^\alpha f_k - D^\alpha f) \|_\infty \leq \| D^\alpha f_k - D^\alpha f \|_1,
\]

and Lemma 2 tells us that this tends to 0 as \( k \to \infty \). On the other hand, for \( \xi \in \mathbb{R}^d \),

\[
|\hat{D^\alpha f_k}(\xi) - (2\pi i)^\alpha \hat{f}(\xi)| = |(2\pi i)^\alpha \hat{f_k}(\xi) - (2\pi i)^\alpha \hat{f}(\xi)|
\]

\[
= |(2\pi i)^\alpha \| \mathcal{F}(f_k - f)(\xi) \|_1
\]

\[
\leq |(2\pi i)^\alpha \| f_k - f \|_1,
\]

which by Lemma 2 tends to 0 as \( k \to \infty \). Therefore, for \( \xi \in \mathbb{R}^d \),

\[
|\hat{D^\alpha f}(\xi) - (2\pi i)^\alpha \hat{f}(\xi)| \leq \left\| \hat{D^\alpha f_k} - \hat{D^\alpha f} \right\|_\infty + \left| \hat{D^\alpha f_k}(\xi) - (2\pi i)^\alpha \hat{f}(\xi) \right|,
\]

and because the right-hand side tends to 0 as \( k \to \infty \), we get

\[
\hat{D^\alpha f}(\xi) = (2\pi i)^\alpha \hat{f}(\xi).
\]

If \( y \in S^{d-1} \) then there is at least one \( 1 \leq j \leq d \) with \( y_j \neq 0 \), from which we get

\[
\sum_{|\beta|=N} |y^\beta| > 0.
\]

\(^4\)Thomas H. Wolff, Lectures on Harmonic Analysis, p. 4, Proposition 1.4.
The function $y \mapsto \sum_{|\beta|_1=N} |y^\beta|$ is continuous $S^{d-1} \to \mathbb{R}$, so there is some $C_N > 0$ such that
\[
\frac{1}{C_N} \leq \sum_{|\beta|_1=N} |y^\beta|, \quad y \in S^{d-1}.
\]
For nonzero $x \in \mathbb{R}^d$, write $x = |x|y$, with which
\[
\sum_{|\beta|_1=N} |x^\beta| = |x|^N \sum_{|\beta|_1=N} |y^\beta|.
\]
Therefore
\[
|x|^N \leq C_N \sum_{|\beta|_1=N} |x^\beta|, \quad x \in \mathbb{R}^d.
\]
For $|\alpha| \leq N$, because the Fourier transform of an element of $L^1$ belongs to $C_0$, we have by (3) that $\xi \mapsto \xi^\alpha \hat{f}(\xi)$ belongs to $C_0(\mathbb{R}^d)$, and in particular is bounded. Then for $\xi \in \mathbb{R}^d$,
\[
|\xi|^N|\hat{f}(\xi)| \leq C_N \sum_{|\beta|_1=N} |\xi^\beta||\hat{f}(\xi)|
= C_N \sum_{|\beta|_1=N} |\xi^\beta\hat{f}(\xi)|
\leq C_N \sum_{|\beta|_1=N} \|\xi^\beta\hat{f}(\xi)\|_{\infty}
= C'.
\]
On the one hand, for $|\xi| \geq 1$ we have
\[
1 + |\xi| \leq 2|\xi|,
\]
hence
\[
|\xi|^{-N} \leq \left(\frac{1 + |\xi|}{2}\right)^{-N} = 2^N (1 + |\xi|)^{-N},
\]
giving
\[
|\hat{f}(\xi)| \leq C'|\xi|^{-N} \leq C'2^N (1 + |\xi|)^{-N}.
\]
On the other hand, for $|\xi| \leq 1$ we have
\[
1 + |\xi| \leq 2,
\]
and so
\[
|\hat{f}(\xi)| \leq \|\hat{f}\|_{\infty} 2^N 2^{-N} \leq \|\hat{f}\|_{\infty} 2^N (1 + |\xi|)^{-N}.
\]
Thus, for
\[
C = \max \left\{2^N C', 2^N \|\hat{f}\|_{\infty}\right\}
\]
we have
\[
|\hat{f}(\xi)| \leq C (1 + |\xi|)^{-N}, \quad \xi \in \mathbb{R}^d,
\]
completing the proof.
3 Bernstein’s inequality for $L^2$

For a Borel measurable function $f : \mathbb{R}^d \to \mathbb{C}$, let $O$ be the union of those open subsets $U$ of $\mathbb{R}^d$ such that $f(x) = 0$ for almost all $x \in U$. In other words, $O$ is the largest open set on which $f = 0$ almost everywhere. The essential support of $f$ is the set

$$\text{ess supp } f = \mathbb{R}^d \setminus O.$$ 

The following is Bernstein’s inequality for $L^2(\mathbb{R}^d)$.

**Theorem 4.** If $f \in L^2(\mathbb{R}^d)$, $R > 0$, and

$$\text{ess supp } \hat{f} \subset B(0, R),$$

then there is some $f_0 \in C^\infty(\mathbb{R}^d)$ such that $f(x) = f_0(x)$ for almost all $x \in \mathbb{R}^d$, and for any multi-index $\alpha$,

$$\|D^\alpha f_0\|_2 \leq (2\pi R)^{|\alpha|_1} \|f\|_2.$$

**Proof.** Let $\chi_R$ be the indicator function for $B(0, R)$. By (4), the Cauchy-Schwarz inequality, and the Parseval identity,

$$\left\|\hat{f}\right\|_1 = \left\|\chi_R \hat{f}\right\|_1 \leq \left\|\chi_R\right\|_2 \left\|\hat{f}\right\|_2 = m_d(B(0, R))^{1/2} \|f\|_2 < \infty,$$

so $\hat{f} \in L^1(\mathbb{R}^d)$. The Plancherel theorem tells us that if $g \in L^2(\mathbb{R}^d)$ and $\hat{g} \in L^1(\mathbb{R}^d)$, then

$$g(x) = \int_{\mathbb{R}^d} \hat{g}(\xi) e^{2\pi i x \cdot \xi} d\xi$$

for almost all $x \in \mathbb{R}^d$. Thus, for $f_0 : \mathbb{R}^d \to \mathbb{C}$ defined by

$$f_0(x) = \int_{\mathbb{R}^d} \hat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi = \mathcal{F}(\hat{f})(-x), \quad x \in \mathbb{R}^d,$$

we have $f(x) = f_0(x)$ for almost all $x \in \mathbb{R}^d$. Because $f = f_0$ almost everywhere,

$$\hat{f}_0 = \hat{f}.$$

Applying Theorem 1 to $d\mu(\xi) = f_0(-\xi) d\xi$, we have $f_0 \in C^\infty(\mathbb{R}^d)$ and for any multi-index $\alpha$,

$$D^\alpha f_0 = \mathcal{F}((-2\pi i \xi)^\alpha \hat{f}(-\xi)).$$

---

By Parseval’s identity,
\[
\|D^\alpha f_0\|_2 = \left\| (-2\pi i \xi)^\alpha \hat{f}(-\xi) \right\|_2 \\
= \left\| (2\pi i \xi)^\alpha \chi_R(\xi) \hat{f}(\xi) \right\|_2 \\
\leq \|(2\pi i \xi)^\alpha \chi_R(\xi)\|_\infty \left\| \hat{f} \right\|_2 \\
\leq (2\pi R)^{\alpha |\alpha|} \left\| \hat{f} \right\|_2 \\
= (2\pi R)^{\alpha |\alpha|} \| f \|_2,
\]
proving the claim. \(\square\)

4 Nikolsky’s inequality

Nikolsky’s inequality tells us that if the Fourier transform of a function is supported on a ball centered at the origin, then for \(1 \leq p \leq q \leq \infty\), the \(L^q\) norm of the function is bounded above in terms of its \(L^p\) norm.\(^7\)

**Theorem 5.** There is a constant \(C_d\) such that if \(f \in \mathcal{S}(\mathbb{R}^d), R > 0,\)
\[
supp \hat{f} \subset \overline{B(0,R)},
\]
and \(1 \leq p \leq q \leq \infty\), then
\[
\|f\|_q \leq C_d R^d \left( \frac{1}{p} - \frac{1}{q} \right) \|f\|_p.
\]

**Proof.** Let \(g = f_R\), i.e.
\[
g(x) = R^{-d} f(R^{-1} x), \quad x \in \mathbb{R}^d.
\]
Then for \(\xi \in \mathbb{R}^d,\)
\[
\hat{g}(\xi) = \int_{\mathbb{R}^d} g(x) e^{-2\pi i \xi \cdot x} dx = \int_{\mathbb{R}^d} R^{-d} f(R^{-1} x) e^{-2\pi i \xi \cdot x} dx = \hat{f}(R\xi),
\]
showing that \(supp \hat{g} = R^{-1} supp \hat{f} \subset \overline{B(0,1)}\). Let \(\chi \in \mathcal{S}(\mathbb{R}^d)\) with \(\chi(\xi) = 1\) for \(|\xi| \leq 1\), with which
\[
\hat{\chi} = \chi \hat{\chi}.
\]
Then \(g = (\mathcal{F}^{-1}\chi) * g\), and using Young’s inequality, with \(1 + \frac{1}{q} = \frac{1}{p} + \frac{1}{p},\)
\[
\|g\|_q \leq \|\mathcal{F}^{-1}\chi\|_r \|g\|_r = \|\hat{\chi}\|_r \|g\|_q. \tag{5}
\]

Moreover, 
\[ \|g\|_a = \left( \int_{\mathbb{R}^d} |R^{-d}f(R^{-1}x)|^a dx \right)^{1/a} \]
\[ = \left( \int_{\mathbb{R}^d} R^{-da+d}|f(y)|^a dy \right)^{1/a} \]
\[ = R^{d\left(\frac{1}{a} - 1\right)} \|f\|_a, \]
so (5) tells us 
\[ R^{d\left(\frac{1}{a} - 1\right)} \|f\|_q \leq \|\hat{\chi}\|_r R^{d\left(\frac{1}{r} - \frac{1}{q}\right)} \|f\|_p, \]
i.e.
\[ \|f\|_q \leq \|\hat{\chi}\|_r R^{d\left(\frac{1}{r} - \frac{1}{q}\right)} \|f\|_p. \]
Now, \( \frac{1}{r} = 1 + \frac{1}{q} - \frac{1}{p} \), so \( 0 \leq \frac{1}{r} \leq 1 \) because \( 1 \leq p \leq q \leq \infty \), namely, \( 1 \leq r \leq \infty \).
By the log-convexity of \( L^r \) norms, for \( \frac{1}{r} = 1 - \theta \) we have
\[ \|\hat{\chi}\|_r \leq \|\hat{\chi}\|_{1-\theta} \|\hat{\chi}\|_\theta^\theta. \]
Thus with 
\[ C_d = \max\{\|\hat{\chi}\|_1, \|\hat{\chi}\|_\infty\} \]
we have proved the claim. \( \square \)

5 The Dirichlet kernel and Fejér kernel for \( \mathbb{R} \)

The function \( D_M \in C_0(\mathbb{R}) \) defined by
\[ D_M(x) = \frac{\sin 2\pi Mx}{\pi x}, \quad x \neq 0 \]
and \( D_M(0) = 2M \), is called the Dirichlet kernel. Let \( \chi_M \) be the indicator function for the set \([-M,M]\). We have, for \( x \neq 0 \),
\[ \hat{\chi}_R(x) = \int_{\mathbb{R}} \chi_R(\xi)e^{-2\pi i\xi x} d\xi \]
\[ = \int_{-M}^{M} e^{-2\pi i\xi x} d\xi \]
\[ = \frac{e^{-2\pi iMx} - 2\pi iMx}{-2\pi i x} \]
\[ = 1 - \frac{e^{2\pi iMx} - e^{-2\pi iMx}}{2\pi i}. \]
For $x = 0$, $\hat{\chi}_R(0) = 2M = D_M(0)$. Thus,

$$D_M = \hat{\chi}_R.$$ 

For $f \in L^1(\mathbb{R})$ and $M > 0$, we define

$$(S_M f)(x) = \int_{-M}^{M} f(\xi) e^{2\pi i \xi x} d\xi, \quad x \in \mathbb{R}.$$ 

It is straightforward to check that

$$(S_M f)(x) = \int_{\mathbb{R}} \sin 2\pi Mt \frac{f(x-t)}{\pi t} dt = (D_M * f)(x), \quad x \in \mathbb{R}.$$ 

For $f \in L^1(\mathbb{R})$, $M > 0$, and $x \in \mathbb{R}$,

$$\frac{1}{M} \int_{0}^{M} (S_m f)(x) dm = \frac{1}{M} \int_{0}^{M} \left( \int_{-m}^{m} f(\xi) e^{2\pi i \xi x} d\xi \right) dm$$

$$= \frac{1}{M} \int_{0}^{M} \left( \int_{-m}^{m} \left( \int_{\mathbb{R}} f(y) e^{-2\pi i \xi y} dy \right) e^{2\pi i \xi x} d\xi \right) dm$$

$$= \frac{1}{M} \int_{\mathbb{R}} f(y) \left( \int_{0}^{M} \left( \int_{-m}^{m} e^{-2\pi i \xi (y-x)} d\xi \right) dm \right) dy$$

$$= \frac{1}{M} \int_{\mathbb{R}} f(y) \left( \int_{0}^{M} D_m(y-x) dm \right) dy$$

$$= \frac{1}{M} \int_{\mathbb{R}} f(y) \left( \int_{0}^{M} \sin 2\pi m(y-x) \pi(y-x) dm \right) dy$$

$$= \frac{1}{M} \int_{\mathbb{R}} f(y) \left( \int_{0}^{M} \frac{\cos 2\pi m(y-x)}{2\pi^2(y-x)^2} \right) dy$$

$$= \frac{1}{M} \int_{\mathbb{R}} f(y) \left( \frac{1}{2\pi^2(y-x)^2} - \frac{\cos 2\pi M(y-x)}{2\pi^2(y-x)^2} \right) dy.$$ 

We define the Fejér kernel $K_M \in C_0(\mathbb{R})$ by

$$K_M(x) = \frac{1 - \cos 2\pi Mx}{2\pi^2 x^2}, \quad x \neq 0,$$

and $K_M(0) = M$. Thus, because $K_M$ is an even function,

$$\frac{1}{M} \int_{0}^{M} (S_m f)(x) dm = (K_M * f)(x).$$

One proves that $K_M$ is an approximate identity: $K_M \geq 0$,

$$\int_{\mathbb{R}} K_M(x) dx = 1,$$

10
and for any $\delta > 0$,
\[
\lim_{M \to \infty} \int_{|x| > \delta} K_M(x) dx = 0.
\]
The fact that $K_M$ is an approximate identity implies that for any $f \in L^1(\mathbb{R})$, $K_M \ast f \to f$ in $L^1(\mathbb{R})$ as $M \to \infty$.

We shall use the Fejér kernel to prove Bernstein’s inequality for $\mathbb{R}$.

**Theorem 6.** If $\mu \in M(\mathbb{R})$, $M > 0$, and
\[
\text{supp } \mu \subset [-M,M],
\]
then
\[
\|\hat{\mu}'\|_{\infty} \leq 4\pi M \|\hat{\mu}\|_{\infty}.
\]

**Proof.** For $x_0 \in \mathbb{R}$, let $d\mu_{x_0}(t) = e^{-2\pi i x_0 t}d\mu(t)$. $\mu_{x_0}$ has the same support has $\mu$, and
\[
\hat{\mu}_{x_0}(x) = \int_{\mathbb{R}} e^{-2\pi i t x} d\mu_{x_0}(t) = \int_{\mathbb{R}} e^{-2\pi i t x} e^{-2\pi i x_0 t} d\mu(t) = \hat{\mu}(x + x_0).
\]

It follows that to prove the claim it suffices to prove that $|\hat{\mu}'(0)| \leq 4\pi M \|\hat{\mu}\|_{\infty}$.

Write $f = \hat{\mu} \in C_c(\mathbb{R})$. Define $\Delta_M \in C_c(\mathbb{R})$ by
\[
\Delta_M(t) = \begin{cases} M - |t| & |t| < M \\ 0 & |t| \geq M \end{cases}, \quad t \in \mathbb{R}.
\]

We calculate, for $x \neq 0$,
\[
\int_{\mathbb{R}} \Delta_M(t) e^{-2\pi i x t} dt = -\frac{e^{-2\pi i M x}(-1 + e^{2\pi i M x})^2}{4\pi^2 x^2} = \frac{(\sin \pi M x)^2}{\pi^2 x^2} = 1 - \cos 2\pi M x = \frac{2\pi^2 M x}{2\pi^2 x^2}.
\]

so
\[
\hat{\Delta}_M(x) = M K_M(x).
\]

Then for $t \in [-M,M]$,
\[
\int_{\mathbb{R}} (e^{2\pi i M \xi} - e^{-2\pi i M \xi}) K_M(\xi) e^{-2\pi i \xi t} d\xi = \frac{\hat{K}_M(t - M) - \hat{K}_M(t + M)}{M} = \Delta_M(-t + M) - \Delta_M(-t - M) = \frac{t}{M}.
\]

Mark A. Pinsky, *Introduction to Fourier Analysis and Wavelets*, p. 122, Theorem 2.3.17.
On the one hand, the integral of the left-hand side with respect to \( \mu \) is
\[
\int_{\mathbb{R}} \int_{\mathbb{R}} (e^{2\pi i M \xi} - e^{-2\pi i M \xi}) K_M(\xi) e^{-2\pi i \xi t} \, d\xi d\mu(t) = \int_{\mathbb{R}} (e^{2\pi i M \xi} - e^{-2\pi i M \xi}) K_M(\xi) f(\xi) \, d\xi.
\]

On the other hand, the integral of the right-hand side with respect to \( \mu \) is
\[
\int_{\mathbb{R}} \frac{t}{M} \, d\mu(t) = \frac{1}{-2\pi i M} \int_{\mathbb{R}} -2\pi i t d\mu(t) = \frac{1}{-2\pi i M} \mathcal{F}((-2\pi i t)\mu)(0) = \frac{1}{-2\pi i M} f'(0).
\]

Hence
\[
\frac{1}{-2\pi i M} f'(0) = \int_{\mathbb{R}} (e^{2\pi i M \xi} - e^{-2\pi i M \xi}) K_M(\xi) f(\xi) \, d\xi,
\]
giving
\[
|f'(0)| \leq 4\pi M \|f\|_{\infty} \|K_M\|_1 = 4\pi M \|f\|_{\infty},
\]
proving the claim. \( \square \)