7. Wiener and Besicovich lemmas; differentiation on $\mathbb{R}^d$

Having analyzed the behavior of functions of bounded variation on $\mathbb{R}$ in great detail, we now turn attention to the similar problem in $\mathbb{R}^d$ for $d > 1$. As it turns out, many of the ideas developed for the one-dimensional case carry over, albeit not without additional innovations. We start by a covering lemma due to N. Wiener that is based on a combinatorial argument and thus works in arbitrary metric spaces:

**Lemma 7.1** (Wiener’s covering lemma) Let $\{B(x_i, r_i): i = 1, \ldots, N\}$ be open balls in a metric space. Then there exists $M \in \{1, \ldots, N\}$ and a strictly increasing sequence $k_1, \ldots, k_M$ from $\{1, \ldots, N\}$ such that

$$\forall 1 \leq i < j \leq M: \quad B(x_i, r_i) \cap B(x_j, r_j) = \emptyset \quad (7.1)$$

and

$$\bigcup_{i=1}^n B(x_i, r_i) \subseteq \bigcup_{j=1}^m B(x_{k_j}, 3r_{k_j}). \quad (7.2)$$

**Proof.** Assume without loss of generality that the balls are ordered such that

$$r_1 \geq r_2 \geq \cdots \geq r_N. \quad (7.3)$$

and let us abbreviate $B_i := B(x_i, r_i)$. Then set $k_1 := 1$ and, for each $m \geq 1$ for which $k_m$ has already been defined, set

$$k_{m+1} := \inf\{i > k_m: B_i \cap \bigcup_{j=1}^m B_{k_j} = \emptyset\} \quad (7.4)$$

unless the set on the right is empty, in which case we set $M := m$ and terminate. This process generates a sequence $k_1 < \cdots < k_M$ such that (7.1) holds.

It remains to show (7.2). Let $x \in \bigcup_{i=1}^n B_i$, set

$$i_0 := \min\{i = 1, \ldots, N: x \in B_i\} \quad (7.5)$$

and denote

$$j_0 := \max\{j = 1, \ldots, M: k_j \leq i_0\}. \quad (7.6)$$

Then $B_{i_0} \cap B_{j_0} \neq \emptyset$, for if not then $i_0 \geq k_{j_0} + 1$ in contradiction with definition of $j_0$. Picking any $z \in B_{i_0} \cap B_{j_0}$ and writing $d$ for the distance in the metric space, we get

$$d(x, x_{k_{i_0}}) \leq d(x, x_{i_0}) + d(x_{i_0}, x_{k_{i_0}}) \leq d(x, x_{i_0}) + d(x_{i_0}, z) + d(x_{k_{i_0}}, z) < 2r_{i_0} + r_{k_{i_0}}. \quad (7.7)$$

In light of (7.3) and $i_0 \geq k_{j_0}$, this gives $d(x, x_{k_{i_0}}) < 3r_{k_{i_0}}$ and so $x \in B(x_{k_{i_0}}, 3r_{k_{i_0}})$. Hereby (7.2) follows. \qed

Using Wiener’s lemma, we now show:

**Theorem 7.2** (Strong maximal inequality for Lebesgue measure on $\mathbb{R}^d$) Let $d \geq 1$, let $\| \cdot \|$ be a norm on $\mathbb{R}^d$ and denote $B(x, r) := \{y \in \mathbb{R}^d: \|x - y\| < r\}$. Given a function
\[ f \in L^1(\mathbb{R}^d, \lambda), \text{ consider its uncentered maximal function} \]

\[
f^{**}(x) := \sup_{r > 0} \sup_{y \in B(x,r)} \frac{1}{\lambda(B(y,r))} \int_{B(y,r)} |f|d\lambda.
\] (7.8)

Then

\[
\forall t > 0: \quad \lambda(f^{**} > t) \leq \frac{3^d}{t} \int_{\{f^{**} > t\}} |f|d\lambda.
\] (7.9)

**Proof.** Pick \( t > 0 \) and assume \( \{f^{**} > t\} \neq \emptyset \). Pick \( K \subseteq \mathbb{R}^d \) compact such that

\[ K \subseteq \{f^{**} > t\}. \] (7.10)

Abbreviating \( B := \{B(x,r) : x \in \mathbb{R}^d \land r > 0\} \), consider the collection

\[ W := \left\{ B \in B : \int_B |f|d\lambda > t\lambda(B) \right\}. \] (7.11)

Since \( f^{**}(x) > t \) implies existence of a ball \( B \) such that \( x \in B \) and \( \int_B |f|d\lambda > t\lambda(B) \), the family \( W \) covers \( K \). The Heine-Borel Theorem implies existence of a finite collection \( \bar{B}_1, \ldots, \bar{B}_N \in W \) such that \( K \subseteq \bigcup_{i=1}^N \bar{B}_i \). Since our definition of \( f^{**} \) ensures that every \( B \in W \) is entirely contained in \( \{f^{**} > t\} \), Wiener’s covering lemma gives

\[
\exists M \geq 1 \exists B_1, \ldots, B_M \in W: \quad \text{disjoint} \quad \bigcup_{i=1}^M B_i \subseteq \{f^{**} > t\} \quad \land \quad K \subseteq \bigcup_{i=1}^M B_i',
\] (7.12)

where \( B_i' \) is the ball with the same center at \( B_i \) but 3-times its radius.

Since the scaling property of the Lebesgue measure implies

\[ \lambda(B_i') = 3^d \lambda(B_i) \] (7.13)

with the help of the definition of \( W \) we now compute

\[
t\lambda(K) \leq t \sum_{i=1}^M \lambda(B_i') \leq 3^d t \sum_{i=1}^M \lambda(B_i) \leq 3^d \sum_{i=1}^M \int_{B_i} |f|d\lambda
\] (7.14)

\[ = 3^d \int_{\bigcup_{i=1}^M B_i} |f|d\lambda \leq 3^d \int_{\{f^{**} > t\}} |f|d\lambda. \]

Taking supremum over compact sets \( K \) subject to (7.10), the claim follows from the inner regularity of the Lebesgue measure.

\[ \square \]

Notice that, compared to the one-dimensional version of the above statement, we pick up an additional factor \( 3^d \). On the other hand, the proof does not use anything specific about \( \mathbb{R}^d \) except for the scaling property of the Lebesgue measure in (7.13). Using the same argument as in one dimension, we then conclude:
Theorem 7.3 (Lebesgue differentiation for Lebesgue measure in $\mathbb{R}^d$)  For each function $f \in L^1,\text{loc}(\mathbb{R}^d, \lambda)$ and for $B(x, r)$ denoting the open ball of radius $r$ centered at $x$ relative to any norm-metric on $\mathbb{R}^d$,
\[
\lim_{r \downarrow 0} \frac{1}{\lambda(B(x, r))} \int_{B(x, r)} f \, d\lambda = f(x)
\tag{7.15}
\]
and, in fact, even
\[
\lim_{r \downarrow 0} \frac{1}{\lambda(B(x, r))} \int_{B(x, r)} |f - f(x)| \, d\lambda = 0
\tag{7.16}
\]
holds for $\lambda$-a.e. $x \in \mathbb{R}^d$.

We leave the corresponding generalizations (which amount to proving a $d$-dimensional version of Lemma 4.4) to the reader.

Although inequalities of the type (7.9) are often referred to as “weak” because, technically, they mean that the map $f \mapsto f^{**}$ takes $L^1(\lambda)$ into the weak-$L^1$ space $L^{1,w}(\lambda)$, we refer to it as “strong” because of the event $\{f^{**} > t\}$ under the integral sign. With this improvement, the inequality readily implies the “strong” version of the maximal inequality which states that $f \mapsto f^{**}$ is continuous as a map $L^p(\lambda) \to L^p(\lambda)$ for any $p > 1$.

This can be proved even without reference to a maximal function:

Lemma 7.4  Let $\mu$ be a Borel measure on $\mathbb{R}^d$ and let $f, g : \mathbb{R}^d \to [0, \infty)$ be Borel measurable functions such that either $\mu$ is finite or $f \in L^1(\mu)$ and
\[
\forall t > 0: \quad \mu(g > t) \leq \int_{\{g > t\}} f \, d\mu. \tag{7.17}
\]

Then for all $p \in (1, \infty)$, $f \in L^p(\mu)$ implies $g \in L^p(\mu)$ and, in fact,
\[
\|g\|_{L^p(\mu)} \leq \frac{p}{p-1}\|f\|_{L^p(\mu)}. \tag{7.18}
\]

Proof. Let $M > 1$ and $\epsilon \in (0, 1)$ and note that $\mu(g > \epsilon)$ is finite by $f \in L^1(\mu)$. We then have $(g \wedge M)1_{\{g \geq \epsilon\}} \in L^p(\mu)$. Tonelli’s theorem then shows
\[
\int_{\{g \geq \epsilon\}} (g \wedge M)^p \, d\mu = \int_{\{g \geq \epsilon\}} (g \wedge M)^p \, d\mu
\leq \int_{\{g \geq \epsilon\}} \left( \int_0^M pt^{p-1} \, dt \right) \, d\mu
\leq p \int_0^M t^{p-1} \mu(g \geq t \vee \epsilon) \, dt
\leq p \int_0^M t^{p-1} \frac{1}{t \vee \epsilon} \left( \int_{\{s \geq t \vee \epsilon\}} f \, d\mu \right) \, dt. \tag{7.19}
\]
where we plugged (7.17) in the last step. Applying Tonelli one more time along with
\[
p \int_0^M t^{p-1} \frac{1}{t} \vee e 1_{\{g \geq t \vee e\}}(x) \, dt \\
= 1_{\{g \geq e\}}(x) \left( \int_0^e M pt^{p-1} \, dt + \int_e^g pt^{p-2} \, dt \right) \\
= 1_{\{g \geq e\}}(x) \left( \int_0^e (g(x) \wedge M) pt^{p-2} \, dt \right) \\
\leq \frac{p}{p-1} 1_{\{g \geq e\}}(x) (g(x) \wedge M)^{p-1} 
\]
(7.20)
which uses \( \frac{p}{p-1} \geq 1 \) for \( p > 1 \), we get
\[
\int_{\{g \geq e\}} (g \wedge M)^p d\mu \leq \frac{p}{p-1} \int_{\{g \geq e\}} (g(x) \wedge M)^p f \, d\mu \\
\leq \frac{p}{p-1} \left( \int_{\{g \geq e\}} (g \wedge M)^p d\mu \right)^{1/q} \left( \int f^p d\mu \right)^{1/p},
\]
(7.21)
where \( q \) is the conjugate index to \( p \) and where we used Hölder’s inequality along with the fact that \( q(p-1) = p \) in the last step. This gives (7.18) with \((g \wedge M)1_{\{g \geq e\}}\) instead of \( g \). Taking \( M \to \infty \) and \( e \downarrow 0 \) with the help of the Monotone Convergence Theorem, we get (7.18) for \( g \) as well.

\[\text{Remark 7.5} \quad \text{When } g \text{ is the maximal function } f^{**}, \text{the truncations of of } g \text{ can be circumvented by assuming that } f \text{ is bounded and } \text{supp}(f) \text{ is bounded. The maximal function is then bounded and } \mu \text{ can effectively be treated as finite (i.e., restricted to supp}(f)\text{). The proof then becomes an elementary calculation:}
\]
which by the above reasoning implies
\[
\|f^*\|_{L^p(\lambda)} \leq 2^{d(p-1)} 3^d \frac{p}{p-1} \|f\|_{L_p(\lambda)}.
\] (7.25)

This is of course worse than the bound we get for \(f^{**}\), because \(f^{**} \leq f^*\).

As already noted, a deficiency of the derivation underlying Theorem 7.3 is that it relies on the scaling property of the Lebesgue measure. This fails, even as a bound, for most other measures and so, in order to address general measures, we need to come up with a different covering argument. This was achieved by A.S. Besicovich. Let \(B_\infty(x,r)\) denote the \(\ell^\infty\)-ball of radius \(r\) centered at \(x\), i.e.,
\[
B_\infty(x,r) := \{ y \in \mathbb{R}^d : \|x - y\|_\infty < r \}. \tag{7.26}
\]

We then have:

**Proposition 7.6** (Besicovich covering lemma)  
For each \(d \geq 1\) there is \(c(d) \in (0, \infty)\) such that for all bounded sets \(E \subseteq \mathbb{R}^d\), all \(R > 0\) and any collection \(B\) of \(\ell^\infty\)-balls of radii in \((0, R)\) with the property
\[
\forall x \in E \exists r \in (0, R) : B_\infty(x,r) \in B \tag{7.27}
\]
there exists \(\{B_n\}_{n \geq 1} \subseteq B \cup \{\emptyset\}\) such that
\[
E \subseteq \bigcup_{n \geq 1} B_n \quad \text{and} \quad \sum_{n \geq 1} 1_{B_n} \leq c(d). \tag{7.28}
\]

Similarly to Wiener’s covering lemma, the proof hinges on the following combinatorial/geometric observation:

**Lemma 7.7**  
Let \(d \geq 1\) and let \(\{B_\infty(x_i,r_i) : i = 1, \ldots, N\}\) be \(\ell^\infty\)-balls in \(\mathbb{R}^d\) such that
\[
\forall 1 \leq i < j \leq N : \ r_j \leq 2r_i \land x_j \not\in B_\infty(x_i,r_i). \tag{7.29}
\]

Then
\[
\sum_{i = 1}^{N} 1_{B_\infty(x_i,r_i)} \leq 4^d. \tag{7.30}
\]

**Proof.** Let \([0, \infty)^d\) be the positive orthant in \(\mathbb{R}^d\). Thanks to translation and reflection invariance of the \(\ell^\infty\)-metric, it suffices to show that if (in addition to above conditions) \(x_i \in [0, \infty)^d\) holds for all \(i = 1, \ldots, N\), then
\[
\sum_{i = 1}^{N} 1_{B_\infty(x_i,r_i)}(0) \leq 2^d, \tag{7.31}
\]
where \(0\) labels the origin in \(\mathbb{R}^d\). (Indeed, it suffices to prove the bound at a generic point, for which the origin suffices. Then divide \(\mathbb{R}^d\) into \(2^d\) orthants and then select just the balls with centers in one of the orthants.)

Assume thus that \(x_i \in [0, \infty)^d\) for all \(i = 1, \ldots, N\). We may also assume that \(0 \in B(x_i,r_i)\) for all \(i = 1, \ldots, N\) because the other points need not be considered. In particular, we have \(\|x_1\|_\infty < r_1\) and so, in light of the positivity of all coordinates of \(x_1,\)
\[
[0,r_1)^d \subseteq B_\infty(x_1,r_1). \tag{7.32}
\]
It follows that
\[ \forall j > 1: \ x_j \not\in [0, r_1)^d. \]  
(7.33)

On the other hand, since \( r_j \leq 2r_1 \), from \( 0 \in B(x_j, r_j) \) we get
\[ \forall j > 1: \ x_j \in [0, 2r_1)^d \quad \text{and} \quad r_j \geq r_1. \]  
(7.34)

Let \( S \) be a translate of \( [0, r_1) \) by \( r_1 \) in some of, and at least one of, the coordinate directions. Then \( S \) contains at most one point from \( \{x_2, \ldots, x_N\} \) because if \( x_j \in S \), then \( r_j \geq r_1 \) implies \( S \subseteq B(x_j, r_j) \) which by (7.29) rules out \( x_k \in S \) for any \( k > j \). Since \([0, 2r_1)^d \) contains \( 2^d \) distinct translates \( S \) of \([0, r_1)^d \) as above, we get (7.31).

We now give:

**Proof of Lemma 7.6.** Define
\[ \alpha_1 := \sup \{ r > 0 : (\exists x \in E : B(x, r) \in \mathcal{B}) \} \]  
(7.35)

By assumption, \( \alpha_1 \in (0, R] \) and so there exists \( B_1 := B(x_1, r_1) \in \mathcal{B} \) such that \( r_1 > \alpha_1 / 2 \). Proceeding recursively, given that \( B_1, \ldots, B_n \) have been defined, if \( E \sim \bigcup_{k=1}^n B_k = \emptyset \) then set \( \alpha_{n+1} := 0 \) and \( B_{n+1} := \emptyset \); otherwise, let
\[ \alpha_{n+1} := \sup \left\{ r > 0 : (\exists x \in E \setminus \bigcup_{k=1}^n B_k : B(x, r) \in \mathcal{B}) \right\}. \]  
(7.36)

By assumption that each point of \( E \) is the center of some ball in \( \mathcal{B} \), there is \( B_{n+1} := B(x_{n+1}, r_{n+1}) \) such that \( x_{n+1} \in E \setminus \bigcup_{k=1}^n B_k \) and \( r_{n+1} > \alpha_{n+1} / 2 \).

Notice that
\[ \forall 1 \leq i < j: \ r_j \leq \alpha_j \leq \alpha_i < 2r_i \wedge x_j \not\in B_i \]  
(7.37)

and so the balls \( \{B_i : i \geq 1\} \) obey (7.29). We thus have
\[ \sum_{i \geq 1} 1_{B_i} \leq 4^d 1_{\bigcup_{i \geq 1} B_i} \]  
(7.38)

It remains to show that \( \bigcup_{i \geq 1} B_i \) covers \( E \). For note that \( \bigcup_{i \geq 1} B_i \subseteq E \subseteq [-R, R]^d \) implies that \( \bigcup_{i \geq 1} B_i \) is bounded and so, by (7.38) and \( \lambda(B_i) \geq 2^d (\alpha_i / 2)^d = \alpha_i^d \),
\[ \sum_{i \geq 1} \alpha_i^d \leq \sum_{i \geq 1} \lambda(B_i) \leq 4^d \lambda \left( \bigcup_{i \geq 1} B_i \right) < \infty \]  
(7.39)

thus implying
\[ \lim_{n \to \infty} \alpha_n = 0. \]  
(7.40)

For each \( x \in E \), the assumptions ensure existence of \( r > 0 \) such that \( B(x, r) \in \mathcal{B} \). Thus, if \( x \not\in \bigcup_{k=1}^n B_k \) for some \( n \geq 1 \) with \( \alpha_n / 2 < r \), then \( B(x, r) \) will have to be picked as \( B_{n+1} \). Such an \( n \) always exists by (7.40) and so, in particular, so \( x \in \bigcup_{k \geq 1} B_k \) as desired. The claim follows with \( c(d) := 4^d \).

Using Proposition 7.6, we now show:
Theorem 7.8 (Weak maximal inequality for Radon measures on \(\mathbb{R}^d\)) Suppose \(\mu\) is a Radon measure on \(\mathbb{R}^d\) such that
\[
\forall r > 0: \quad \mu(B_\infty(x,r)) > 0. \tag{7.41}
\]
Given a function \(f \in L^1(\mu)\), denote
\[
f^*(x) := \sup_{r>0} \frac{1}{\mu(B_\infty(x,r))} \int_{B_\infty(x,r)} |f|d\mu. \tag{7.42}
\]
Then
\[
\forall t > 0: \quad \mu(f^* > t) \leq \frac{c(d)}{t} \int |f|d\mu, \tag{7.43}
\]
where \(c(d)\) is the constant from Proposition 7.6.

Proof. We follow closely the proof of Theorem 7.2 albeit with some important modifications. (For instance, unlike what I erroneously claimed in class, we cannot use the Heine-Borel theorem in conjunction with the Besicovich covering lemma as that would not necessarily lead to a cover of the whole set.)

Fix \(R > 0\) and denote by \(f^*_R\) the supremum in (7.42) restricted to \(r \in (0,R)\). Consider the collection of \(\ell^\infty\)-balls
\[
B := \left\{ B_\infty(x,r) : x \in \mathbb{R}^d \land r \in (0,R) \land \int_{B_\infty(x,r)} |f|d\mu > \mu(B_\infty(x,r)) \right\}. \tag{7.44}
\]
Note that \(B\) contains an \(\ell^\infty\)-ball centered at every point of \(\{f^*_R > t\}\) and the radii of these balls are bounded. Proposition 7.6 then ensures the existence of \(\{B_i\}_{i \geq 1} \subset B \cup \{\emptyset\}\) with
\[
\{f^*_R > t\} \subset \bigcup_{j \geq 1} B_j \land \sum_{j \geq 1} 1_{B_j} \leq c(d) 1_{\bigcup_{j \geq 1} B_j}. \tag{7.45}
\]
The calculation (7.14) then becomes
\[
t \mu(\{f^*_R > t\}) \leq t \sum_{j \geq 1} \mu(B_j) \leq \sum_{j \geq 1} \int_{B_j} |f|d\mu = \int \left(\sum_{j \geq 1} 1_{B_j}\right) |f|d\lambda \leq c(d) \int \sum_{j \geq 1} 1_{B_j} |f|d\mu \leq c(d) \int |f|d\mu. \tag{7.46}
\]
Since \(f^*_R \uparrow f^*\) as \(R \to \infty\), the claim follows by the Monotone Convergence Theorem for measures.

Remark 7.9 Note that the condition (7.41) can be omitted if we interpret the quantity under supremum in (7.42) as 0 whenever \(\mu(B_\infty(x,r)) = 0\). (This prevents such a ball from showing up in \(B\) and has no effect on \(f^*\) which is strictly positive everywhere whenever \(f\) does not vanish \(\mu\)-a.e. — which also rules out that \(\mu\) is identically zero.) This observation will be useful when we consider differentiation of measures where having to deal with (7.41) is somewhat inconvenient.
Using Theorem 7.8 we now get a result whose detailed proof we again leave to the reader:

**Theorem 7.10** (Differentiation for Radon measures in \( \mathbb{R}^d \)) Let \( \mu \) be a Radon measure on \( \mathbb{R}^d \). For each \( f \in L^{1,\text{loc}}(\mu) \) and for \( B_\infty(x,r) \) denoting the open \( \ell^\infty \)-ball of radius \( r \) centered at \( x \),

\[
\lim_{r \downarrow 0} \frac{1}{\mu(B_\infty(x,r))} \int_{B_\infty(x,r)} f \, d\lambda = f(x)
\]  

(7.47)

and, in fact, even

\[
\lim_{r \downarrow 0} \frac{1}{\mu(B_\infty(x,r))} \int_{B_\infty(x,r)} |f - f(x)| \, d\lambda = 0
\]  

(7.48)

holds for \( \mu \)-a.e. \( x \in \mathbb{R}^d \).

Since the inequality (7.43) does not have \( \{ f^* > t \} \) under the integral sign, one might wonder what happens with the conclusion in Lemma 7.4 for general \( \mu \). First note that, by using (7.43) for \( |f|^p \) instead of \( f \), for all \( p \geq 1 \) we have

\[
\forall t > 0: \quad \mu(f^* > t) \leq \frac{c(d)}{t^p} \| f \|^p_{L^p(\mu)}.
\]  

(7.49)

and so \( f \in L^p(\mu) \) implies \( f^* \in L^{p,w}(\mu) \) — where \( L^{p,w}(\mu) \) is the weak-\( L^p \)-space. (We also get \( f^* \in L^{\tilde{p}}(\mu) \) for all \( \tilde{p} < p \).) However, the stronger conclusion is still in effect:

**Theorem 7.11** (Hardy-Littlewood strong maximal inequality in \( \mathbb{R}^d \)) For all \( d \geq 1 \) and all \( p \in (1, \infty) \) there is \( c(d, p) \in (0, \infty) \) such that for any Radon measure \( \mu \) on \( \mathbb{R}^d \) satisfying (7.41), any \( f \in L^1(\mu) \) and \( f^* \) defined by (7.42),

\[
\| f^* \|_{L^p(\mu)} \leq c(d, p) \| f \|_{L^p(\mu)}.
\]  

(7.50)

The proof is, however, quite different from that in Lemma 7.4 as it is based on interpolation inequalities due to Marcinkiewicz. We will prove this theorem when we discuss interpolation in detail.