THE MAXIMAL FUNCTION AND CONDITIONAL SQUARE FUNCTION
CONTROL THE VARIATION: AN ELEMENTARY PROOF

KEVIN HUGHES, BEN KRAUSE, AND BARTOSZ TROJAN

Abstract. In this note we prove the following good-\(\lambda\) inequality, for \(r > 2\), all \(\lambda > 0\), \(\delta \in (0, \frac{1}{2})\)
\[
\nu\{V_r(f) > 3\lambda; M(f) \leq \delta \lambda\} \leq 4\nu\{s(f) > \delta \lambda\} + \delta^2 \left(1 + \frac{16}{r-2}\right)^2 \cdot \nu\{V_r(f) > \lambda\},
\]
where \(M(f)\) is the martingale maximal function, \(s(f)\) is the conditional martingale square function.
This immediately proves that \(V_r(f)\) is bounded on \(L^p\), \(1 < p < \infty\) and moreover is integrable when the maximal function is.

1. Introduction

Let \((X, \mathcal{F}, \nu)\) be a \(\sigma\)-finite measure space and let \((\mathcal{F}_n : n \in \mathbb{Z})\) be a fixed filtration, i.e \(\mathcal{F}_n\) is a \(\sigma\)-field such that \(\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}\). For a martingale \((f_n : n \in \mathbb{Z})\) we define the maximal function, the square function
\[
M(f) = \sup_{n \in \mathbb{Z}} |f_n|, \quad S(f) = \left(\sum_{n \in \mathbb{Z}} |f_n - f_{n-1}|^2\right)^{1/2},
\]
and the conditional square function
\[
s(f) = \left(\sum_{n \in \mathbb{Z}} \mathbb{E}[|f_n - f_{n-1}|^2|\mathcal{F}_{n-1}]\right)^{1/2}.
\]
For the dyadic filtration on \(\mathbb{R}^d\) we have \(S(f) = s(f)\) since the square of a martingale difference \(|f_n - f_{n-1}|^2\) is \(\mathcal{F}_{n-1}\)-measurable. It is well known (see \cite{1}, Theorem 9 and \cite{3}, Theorem 1) that for all \(p \in [1, \infty)\) there exist \(C_p > 0\) such that
\[
C_p^{-1}\|M(f)\|_{L^p(\nu)} \leq \|S(f)\|_{L^p(\nu)} \leq C_p\|M(f)\|_{L^p(\nu)}.
\]
Also, by the Convexity Lemma (see \cite{2}, Theorem 3.2), there is another constant \(C_p > 0\) satisfying
\[
\|s(f)\|_{L^p(\nu)} \leq C_p\|S(f)\|_{L^p(\nu)}.
\]
Another family of operators which measure oscillation are the \(r\)-variation operators defined for \(r \geq 1\) by
\[
V_r(f) = \sup_{n_0 < n_1 < \cdots < n_j} \left(\sum_{j=1}^{j} |f_{n_j} - f_{n_{j+1}}|^r\right)^{1/r},
\]
where the supremum is over all possible finite, increasing subsequences of \(\mathbb{N}\). These variation operators are more difficult to control than the maximal function \(M(f)\). For any \(n_0 \in \mathbb{Z}\), one may pointwise dominate
\[
M(f) \leq V_r(f) + |f_{n_0}|,
\]
where \(r \geq 1\) is arbitrary. We further remark that the variation operators become larger, hence more sensitive to oscillation, as \(r\) decreases. The fundamental boundedness result concerning the \(r\)-variation operators is due to Lépingle.
Theorem 1 (6). For each \( p \in [1, \infty) \) there is \( A_p > 0 \) such that for all \( f \in L^p(X, \nu) \) and \( r > 2 \)

\[
\|V_r(f)\|_{L^p(\nu)} \leq A_p \frac{r}{r-2} \|f\|_{L^p(\nu)}, \quad (p > 1)
\]

and for all \( \lambda > 0 \)

\[
\nu\{V_r(f) > \lambda\} \leq A_1 \frac{r}{r-2} \lambda^{-1} \|f\|_{L^1(\nu)}, \quad (p = 1).
\]

We remark that the range of \( r > 2 \) in the above theorem is sharp, since these estimates can fail for \( r \leq 2 \), (see e.g. [4, 8]).

By now, comparatively simple proofs of Lépingle’s theorem can be found in Pisier and Xu [7] and Bourgain [1] (see also [5]). The idea was to leverage known estimates for jump inequalities to recover variational estimates. Let us recall that the number of \( \lambda \)-jumps, denoted by \( N_\lambda(f) \), is equal to the supremum over \( J \in \mathbb{N} \) such that there is an increasing sequence \( n_0 < n_1 < \ldots < n_J \) satisfying

\[
|f_{n_j} - f_{n_{j-1}}| > \lambda
\]

for all \( 1 \leq j \leq J \). The key result concerning \( \lambda \)-jumps is the following theorem.

Theorem 2 ([1, 7]). For each \( p \in [1, \infty) \) there exist \( B_p > 0 \) such that for all \( f \in L^p(X, \nu) \) and \( \lambda > 0 \)

\[
\|\lambda N^{1/2}_\lambda (f)\|_{L^p(\nu)} \leq B_p \|f\|_{L^p(\nu)}, \quad (p > 1),
\]

and

\[
\nu\{\lambda N^{1/2}_\lambda (f) > t\} \leq B_1 t^{-1} \|f\|_{L^1(\nu)}, \quad (p = 1),
\]

for any \( t > 0 \).

The goal of this note is to provide a new and elementary proof of Lépingle’s result. The proofs of this theorem in [1, 7] and [5] used boundedness of \( \lambda N^{1/2}_\lambda (f) \) on a range of \( L^p(X, \nu) \)-spaces for an open interval of \( p \) containing \( p = 2 \). In contrast, our proof uses weaker information. We instead use only the \( L^2(X, \nu) \)-boundedness of \( \lambda N^{1/2}_\lambda (f) \) in addition to the \( L^2(X, \nu) \)-boundedness of the maximal function and of the conditional square function. The significance of our approach is that it sheds new insight into the relationship between the maximal function, conditional square function, and variation operator. Specifically, we prove the following theorem.

Theorem A. There is \( C > 0 \) such that for all \( \delta \in (0, \frac{1}{2}) \), \( r > 2 \) and \( \lambda > 0 \)

\[
(1.1) \quad \nu\{V_r(f) > 3\lambda; M(f) \leq \delta \lambda\} \leq 4 \nu\{s(f) > \delta \lambda\} + \delta^2 \left(1 + \frac{16}{r-2}\right)^2 \cdot \nu\{V_r(f) > \lambda\},
\]

In particular, by taking \( \delta > 0 \) sufficiently small and integrating the distribution functions we obtain that for all \( p \in [1, \infty) \) and \( r > 2 \), \( V_r(f) \in L^p(X, \nu) \) whenever \( M(f) \in L^p(X, \nu) \).

1.1. Acknowledgments. The authors wish to thank Konrad Kolesko, Christoph Thiele, and Jim Wright for helpful discussions, and Michael Lacey and Terence Tao for their careful reading and support. Furthermore, this work was completed during the “Harmonic Analysis and Partial Differential Equations” trimester program at Hausdorff Research Institute for Mathematics; the authors wish to thank HIM for their hospitality.

1.2. Notation. We write \( X \lesssim Y \), or \( Y \gtrsim X \) to denote the estimate \( X \leq CY \) for an absolute constant \( C > 0 \). If the constant \( C \) depends on a parameter, we shall indicate this by subscripts, thus for instance \( X \lesssim_p Y \) denotes the estimate \( X \leq C_p Y \) for some \( C_p > 0 \) depending on \( p \).
2. The Proof

We begin with a preliminary lemma.

**Lemma 1.** There is $C > 0$ such that for any $A \in \mathcal{F}_m$, all $\lambda > 0$, and $\delta \in (0, \frac{1}{2})$

$$\nu \{ A; V_r(f) > \lambda; \mathcal{M}(f) \leq \delta \lambda \} \leq C \lambda^{-2}(r - 2)^{-2} \int_A |f|^2 \, d\nu$$

for each $f \in L^2(X, \nu)$ satisfying

$$\mathbb{E}[f \cdot 1_A |\mathcal{F}_n] = \begin{cases} 0 & \text{if } n \leq m, \\ f_n \cdot 1_A & \text{otherwise.} \end{cases}$$

**Proof.** By homogeneity, it suffices to prove the result with $\lambda = 1$. We can pointwise dominate the variation as in [1, §3]

$$(V_r(f))^r \leq \sum_{l \in \mathbb{Z}} 2^{2l} N_2(f).$$

Let $s = (r + 2)/2$. Since $\mathcal{M}(f) < \delta < 1/2$, the above sum runs over $l \leq 0$, which leads to the containment

$$\{ A; V_r(f) > 1; \mathcal{M}(f) \leq \delta \} \subseteq \left\{ \left\{ f \in A; \sum_{l \leq 0} 2^{2l} N_2(f) > 1 \right\} \right\} \subseteq \bigcup_{l \leq 0} \{ 2^{2l} N_2(g) > c_r \},$$

where $g = f \cdot 1_A$ and

$$c_r^{-1} = \frac{1}{2} \sum_{l \leq 0} 2^{(r-2)/2}.$$ 

Let us observe that $c_r^{-1} \leq 1 + \frac{8}{r}$ for all $r > 2$. In light of Theorem 2, where $B_2$ can be taken to be 1, this immediately leads to the majorization

$$\nu \{ A; V_r(f) > 1; \mathcal{M}(f) \leq \delta \} \leq c_r^{-1} \sum_{l \leq 0} 2^{(s-2)/2} \int_A |f|^2 \, d\nu \leq \left( 1 + \frac{16}{r - 2} \right)^2 \int_A |f|^2 \, d\nu. \quad \Box$$

**Proof of Theorem A.** By homogeneity, it will suffice to prove (1.1) for $\lambda = 1$.

Let $B = \{ s(f) > \delta \}$, $B^* = \{ \mathcal{M}(1_B) > 1/2 \}$ and $G = (B^*)^c$. By Doob’s inequality, we have

$$\nu(B^*) \leq 4 \int |1_B|^2 \, d\nu = 4\nu\{ s(f) > \delta \}.$$

Therefore, it is enough to show that

$$\nu \{ V_r(f) > 3; \mathcal{M}(f) \leq \delta; G \} \leq \delta^2 \cdot \nu \{ V_r(f) > 1 \}.$$ 

For $r \geq 1$ and $m \in \mathbb{N}$, define the pointwise, truncated variation operators

$$V_r(f : n \leq m) = \sup_{n_0 < n_1 < \ldots < n_j \leq m} \left( \sum_{j=1}^{J} |f_{n_j} - f_{n_{j+1}}|^r \right)^{1/r}$$

where the supremum is over all possible increasing subsequences of $\{1, \ldots, m\}$. Let $\sigma$ be a stopping time defined to be equal to the minimal $m \in \mathbb{N}$ such that

$$V_r(f : n \leq m) > 1.$$ 

Then

$$V_r(f) \leq V_r(f - f_\sigma) + 2\mathcal{M}(f) + V_r(f : n < \sigma),$$

thus for $g = f - f_\sigma$ we have

$$\{ V_r(f) > 3; \mathcal{M}(f) \leq \delta; G \} \subseteq \{ V_r(g) > 1; \mathcal{M}(g) \leq 2\delta; G \}. $$
We are going to prove that for each $m \in \mathbb{Z}$
\[
\nu\{V_r(g) > 1; \mathcal{M}(g) \leq 2\delta; G; \sigma = m\} \lesssim \delta^2 \cdot \nu\{\sigma = m\}.
\]
For $n \in \mathbb{Z}$ we define $U_n = \{x : E[\mathbf{1}_B|\mathcal{F}_n](x) \leq 1/2\}$. We notice that, if $x \in G$ then $x \in U_n$ for all $n \in \mathbb{Z}$. Let
\[
\tilde{g}(x) = \sum_{n \in \mathbb{Z}} (g_n(x) - g_{n-1}(x)) \cdot \mathbf{1}_{U_{n-1}}(x).
\]
We observe that $g_n(x) = \tilde{g}_n(x)$ for all $x \in G$ and $n \in \mathbb{Z}$. Indeed, $(g_n - g_{n-1})\mathbf{1}_{U_{n-1}}$ is $\mathcal{F}_n$-measurable and
\[
E[(g_n - g_{n-1})\mathbf{1}_{U_{n-1}}|\mathcal{F}_{n-1}] = 0.
\]
Thus for $x \in G$ we have
\[
\tilde{g}_m(x) = \sum_{n \leq m} (g_n(x) - g_{n-1}(x)) \mathbf{1}_{U_{n-1}}(x) = g_m(x).
\]
Therefore, we obtain
\[
\nu\{V_r(g) > 1; \mathcal{M}(g) \leq 2\delta; G; \sigma = m\} = \nu\{V_r(\tilde{g}) > 1; \mathcal{M}(\tilde{g}) \leq 2\delta; G; \sigma = m\} \leq \nu\{V_r(\tilde{g}) > 1; \mathcal{M}(\tilde{g}) \leq 2\delta; \sigma = m\}.
\]
By Lemma 1 we conclude
\[
\nu\{V_r(\tilde{g}) > 1; \mathcal{M}(\tilde{g}) \leq 2\delta; \sigma = m\} \lesssim (r - 2)^{-2} \int_{\{\sigma = m\}} |\tilde{g}|^2 \, d\nu.
\]
Next, $s$ preserves $L^2$-norm thus
\[
\int_{\{\sigma = m\}} |\tilde{g}|^2 \, d\nu = \int s(\tilde{g} \cdot \mathbf{1}_{\{\sigma = m\}}) \, d\nu = \sum_{n \in \mathbb{Z}} \int_{\{\sigma = m\}} E[|g_n - g_{n-1}|^2|\mathcal{F}_{n-1}] \cdot \mathbf{1}_{U_{n-1}} \, d\nu.
\]
Since $\mathbf{1}_{U_n} \leq 2 \cdot E[\mathbf{1}_B|\mathcal{F}_n]$ for each $n \in \mathbb{N}$, we get
\[
\int_{\{\sigma = m\}} |\tilde{g}|^2 \, d\nu \leq 2 \sum_{n \in \mathbb{Z}} \int_{\{\sigma = m\}} E[|g_n - g_{n-1}|^2|\mathcal{F}_{n-1}] \cdot E[\mathbf{1}_B|\mathcal{F}_{n-1}] \, d\nu = 2 \int_{\{\sigma = m\}} s(\tilde{g})^2 \cdot \mathbf{1}_B \, d\nu
\]
which is bounded by $4\delta^2 \cdot \nu\{\sigma = m\}$. 

3. Applications to dyadic $A_\infty$-weights

We remark that in the case of dyadic filtration on $\mathbb{R}^d$, the proof generalizes to handle measures given by $w$, dyadic $A_\infty$-weights. First, let us recall the following definition.

**Definition 1.** A non-negative locally integrable function $w$ belongs to dyadic $A_\infty$, if for every $\epsilon > 0$ there exists $\gamma > 0$ so that for every dyadic interval $I$ and any measurable set $E \subset I$, if $|E| \leq \gamma \cdot |I|$ then
\[
(3.1) \quad w(E) \leq c \epsilon w(I).
\]
If additionally, there is $C > 0$ such that for all dyadic intervals $I$
\[
C^{-1}w(I_L) \leq w(I) \leq Cw(I_R),
\]
where $I_L$ and $I_R$ are, respectively, left and right children of $I$, then $w$ is called dyadic doubling.

**Corollary 1.** Let $w$ be a dyadic $A_\infty$-weight. There exist $C > 0$ so that for each $\epsilon > 0$ and $r > 2$ there is $\delta > 0$ such that for all $\lambda > 0$
\[
w\{V_r(f) > 3\lambda; \mathcal{M}(f) \leq \delta \lambda\} \leq C \cdot w\{S(f) > \delta \lambda\} + \epsilon \cdot w\{V_r(f) > \lambda\}.
\]
Proof. Using the notation as in the proof of Theorem A we may write
\[ \{ V_r(f) > 3; M(f) \leq \delta; G; \sigma = m \} \subseteq \{ \sigma = m \} \]
and
\[ \left| \{ V_r(f) > 3; M(f) \leq \delta; G; \sigma = m \} \right| \leq C \frac{\delta^2}{(r-2)^2} \cdot \left| \{ \sigma = m \} \right| . \]
Given \( \epsilon > 0 \) we take \( \delta > 0 \) small enough so that \( C \frac{\delta^2}{(r-2)^2} \leq \gamma \). Then, by (3.1) we get
\[ w \{ V_r(f) > 3; M(f) \leq \delta; G; \sigma = m \} \leq \epsilon \cdot w \{ \sigma = m \} . \]
Since for the dyadic filtration \( S(f) = s(f) \) we conclude the proof. \( \square \)

Again, by integrating distribution functions for each \( p \in [1, \infty) \) and \( r > 2 \) we can find \( C_{p,r} > 0 \) such that
\[ C_{p,r} \| V_r(f) \|_{L^p(w)} \leq \| S(f) \|_{L^p(w)} + \| M(f) \|_{L^p(w)} . \]
By [9, §2], for each \( w \), a dyadic \( A_\infty \)-weight there is \( C_p > 0 \)
\[ \| M(f) \|_{L^p(w)} \leq C_p \| S(f) \|_{L^p(w)} , \]
so the square function alone dominates \( V_r \) in \( L^p(w) \). In the case where \( w \) is a dyadic doubling, we have the reverse inequality as well
\[ \| S(f) \|_{L^p(w)} \leq C_p^{-1} \| M(f) \|_{L^p(w)} , \]
and thus the maximal function alone dominates \( V_r \) in \( L^p(w) \).

References

[1] J. Bourgain, Pointwise ergodic theorems for arithmetic sets. With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein, Publ. Math.-Paris 69 (1989), no. 1, 5–45.
[2] D. L. Burkholder, B. J. Davis, and R. F. Gundy, Integral inequalities for convex functions of operators on martingales, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory, University of California Press, 1972, pp. 223–240.
[3] B. J. Davis, On the inequality of the martingale square function, Israel J. Math. 8 (1970), 1494–1504.
[4] R. Jones and G. Wang, Variation inequalities for the Fejér and Poisson kernels, T. Am. Math. Soc. 356 (2004), no. 11, 4493–4518.
[5] R. L. Jones, A. Seeger, and J. Wright, Strong variational and jump inequalities in harmonic analysis, Trans. Amer. Math. Soc. (2008), 6711–6742.
[6] D. Lepingle, La variation d’ordre \( p \) des semi-martingales, Z. Wahrscheinlichkeit. 36 (1976), no. 4, 295–316.
[7] G. Pisier and Q. Xu, The strong \( p \)-variation of martingales and orthogonal series, Probab. Theory Rel. 77 (1988), no. 4, 497–514.
[8] J. Qian, The \( p \)-variation of partial sum processes and the empirical process, Ann. Probab. 26 (1998), no. 3, 1370–1383.
[9] M. Wilson, Weighted Littlewood–Paley theory and exponential-square integrability, Lecture notes in mathematics, vol. 1924, Springer, 2008.