J. Math. Anal. Appl. 366 (2010) 67–80

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Large deviations for martingales and derivatives

S. Butler[a,*], S. Pavlov[b], J. Rosenblatt[a]

[a] Department of Mathematics, University of Illinois at Urbana-Champaign, 273 Altgeld Hall, 1409 West Green Street, Urbana, IL 61801, USA
[b] Renaissance Technologies LLC, 600 Rt. 25A, East Setauket, NY 11733, USA

ARTICLE INFO

Article history:
Received 12 November 2008
Available online 13 January 2010
Submitted by M. Ledoux

Keywords:
Martingales
Lebesgue derivatives
Large deviations

ABSTRACT

Fix a sequence of positive integers (m_n) and a sequence of positive real numbers (w_n). Two closely related sequences of linear operators (T_n) are considered. One sequence has $T_n : L_1(\mathbb{R}) \to L_1(\mathbb{R})$ given by the Lebesgue derivatives $T_n f(x) = D_n f(x) = 2^n \int_0^{1/2^n} f(x+t)\,dt$. The other sequence has $T_n : L_1[0,1) \to L_1[0,1)$ given by the dyadic martingale $T_n f(x) = E(f|\beta_n)(x) = 2^n \int_{(l-1)/2^n}^{l/2^n} f(t)\,dt$ when $(l-1)/2^n \leqslant x < l/2^n$ for $l = 1, \ldots, 2^n$. We prove both positive and negative results concerning the convergence of $\sum_{n=1}^{\infty} m\{|T_{m_n} f(x)| \geqslant w_n\}$.

1. Introduction

We would like to be able to understand completely the large deviation behavior of operators given by martingales or approximate identities. We are considering this type of problem. Take an interval I in $\mathbb{R}$, and a sequence of bounded linear operators (T_n) on $L_p(I) = L_p(I, m)$, where m is the usual Lebesgue measure on $\mathbb{R}$. We want to know for which sequences of positive numbers (w_n) and which $L_p(I)$, $1 \leqslant p < \infty$, there is an inequality of this form: for all $f \in L_p(I)$,

$$\sum_{n=1}^{\infty} m\{x \in I\colon\ |T_n f(x)| \geqslant w_n\} < \infty. \tag{1}$$

Eq. (1) is what we mean here by a large deviation inequality. It should be distinguished from large deviation in probability theory which has a much more concrete (but not totally unrelated) form. A homogeneous large deviation inequality would take this form: for some constant C, and for all $f \in L_p(I)$,

$$\sum_{n=1}^{\infty} m\{x \in I\colon\ |T_n f(x)| \geqslant w_n\} \leqslant C\|f\|_{L_p(I)}^p. \tag{2}$$

One approach to obtaining large deviation inequalities could be to get a strong bound on the maximal function $\sup_{n\geqslant 1} |T_n f|$. However, often there is no such strong bound (although there may be a weak bound). In the absence of a strong bound, an alternative method is needed.

2. Background facts

For Eq. (1) to be non-trivial, we should at least know that the levels (w_n) fail to have $\sum_{n=1}^{\infty} \frac{1}{w_n^p}\|T_n\|_{L_p(I)}^p < \infty$ because we always know that the terms $m\{x \in I\colon\ |T_n f(x)| \geqslant w_n\}$ are no larger than $\frac{1}{w_n^p}\|T_n f\|_{L_p(I)}^p \leqslant \frac{1}{w_n^p}\|T_n\|_{L_p(I)}^p \|f\|_{L_p(I)}^p$. For example,

* Corresponding author.
E-mail addresses: svbutler@math.uiuc.edu (S. Butler), savvap@gmail.com (S. Pavlov), jrsnbltt@math.uiuc.edu (J. Rosenblatt).

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2009.12.054

if (T_n) is the dyadic Lebesgue derivatives, then $\|T_n\|_{L_1(\mathbb{R})} = 1$ for all n. So a non-trivial result for the dyadic derivatives on $L_1(\mathbb{R})$ would require that $\sum_{n=1}^{\infty} \frac{1}{w_n} = \infty$.

There are non-trivial large deviation inequalities for averaging operators in ergodic theory and for some related operators. The basic theorems on this appear in Rosenblatt and Wierdl [3]. See in particular Section 4 in [3] where large deviation inequalities are proved for ergodic averages, superadditive processes, and reversed super-martingales. But it has not been clear if there are non-trivial large deviation inequalities where the operators are instead approximate identities or some related operators. Preliminary results in this context already appeared in Rosenblatt and Wierdl [3]. See in particular Section 5 in [3]. In this article, we clarify these issues more by giving both positive and negative results about large deviation inequalities for dyadic derivatives and the closely related case of the dyadic martingale. But first, in this section, we present some background information and simple facts that we will use in the sequel.

The two sequences of operators that we are focusing on in this article are the dyadic martingale and the dyadic Lebesgue derivatives. The natural domains for these operators are different, so we will have to keep track of this in the results being presented here. We denote the usual Lebesgue measure on $\mathbb{R}$ by m. The dyadic martingale $(E(f|\beta_n)\colon n \geqslant 1)$ on $[0,1)$ is given by the dyadic σ-algebras β_n consisting of atoms $\beta_n(l) = [\frac{l-1}{2^n}, \frac{l}{2^n})$ for l, $1 \leqslant l \leqslant 2^n$. For $f \in L_1[0,1)$, we define $E(f|\beta_n)$ to be the step function whose value on $\beta_n(l)$ is $E_n(l)f = 2^n \int_{\beta_n(l)} f\,dm$ for each l, $1 \leqslant l \leqslant 2^n$. The Lebesgue derivatives $D(\epsilon)f$ are defined for $\epsilon > 0$ and for $f \in L_1(\mathbb{R})$ by $D(\epsilon)f(x) = \frac{1}{\epsilon}\int_0^{\epsilon} f(x+t)\,dm(t)$. From these one gets the dyadic Lebesgue derivatives $D_n f$ as $D(\frac{1}{2^n})f$. That is, $D_n f(x) = 2^n \int_0^{1/2^n} f(x+t)\,dt$. The dyadic Lebesgue derivatives and the dyadic martingale are closely related to one another. For analytic results that show this, see the article by Jones, Kaufman, Rosenblatt, and Wierdl [2], in particular the basic result Theorem 2.1 in [3] on the square function of the differences of these operators.

One expectation about large deviations would naturally be the following. Suppose the operators (T_n) converge in L_p-norm. Fix a function $f_o \in L_p(I)$. One expects then that the faster the time index (m_n) of a subsequence (T_{m_n}) grows, the more likely that there will be a large deviation inequality for $(T_{m_n} f_o)$ with previously fixed levels (w_n). This is true if the levels (w_n) are on the order of $(n^{1/p})$ because of this simple fact.

Lemma 2.1. *Given $1 \leqslant p \leqslant \infty$ and a sequence of bounded linear operators (T_n) on $L_p(I)$, such that $(T_n f)$ converges in $L_p(I)$-norm for all $f \in L_p(I)$, and given a fixed function $f_o \in L_p(I)$, there exists a subsequence (T_{m_n}) such that* $\sup_{n\geqslant 1} |T_{m_n} f_o| \in L_p(I)$.

Proof. Let $Lf_o = \lim_{n\to\infty} T_n f_o$. Take (T_{m_n}) such that $\|T_{m_n} f_o - Lf_o\|_{L_p(I)} \leqslant \frac{1}{2^n}$. Then

$$\sup_{n\geqslant 1} |T_{m_n} f_o| \leqslant \sup_{n\geqslant 1} |T_{m_n} f_o - f_o| + |f_o| \leqslant \sum_{n=1}^{\infty} |T_{m_n} f_o - f_o| + |f_o|.$$

Hence, by the triangle inequality, $\|\sup_{n\geqslant 1} |T_{m_n} f_o|\|_{L_p(I)} \leqslant \sum_{n=1}^{\infty} \frac{1}{2^n} + \|f_o\|_{L_p(I)}$. $\square$

Corollary 2.2. *With the assumptions in Lemma* 2.1*, and (w_n) such that* $\liminf_{n\to\infty} \frac{w_n}{n^{1/p}} > 0$*, for any $f \in L_p(I)$, there exists (m_n) such that*

$$\sum_{n=1}^{\infty} m\{|T_{m_n} f| \geqslant w_n\} < \infty.$$

Proof. Using Lemma 2.1, there exists (m_n) such that

$$\sum_{n=1}^{\infty} m\{|T_{m_n} f| \geqslant n^{1/p}\} < \infty$$

since $\|\sup_{n\geqslant 1} |T_{m_n} f|\|_{L_p(I)}$ is finite. We can replace $(n^{1/p})$ by (w_n) if there exists a constant $c > 0$ such that $w_n \geqslant c n^{1/p}$ for all $n \geqslant 1$. $\square$

Remark 2.3. The results above make it clear why it would be worthwhile to be able to characterize which functions f have $\sup_{n\geqslant 1} |T_{m_n} f| \in L_1(I)$ for a given sequence (T_n) and subsequence (m_n). For example, take the case that (T_n) corresponds to the dyadic derivatives on $L_1(\mathbb{R})$. Hare and Stokolos [1] show that this cannot be characterized as an Orlicz class, so it is not just a matter of how large the function is. Instead it is some aspect of the degree of variability in the function that dictates the speed with which the subsequences $(T_{m_n} f)$ converge in this context.

Part of what is shown in Section 4 is that while speeding up the time index may work to get a large deviation inequality for a fixed function, it may not work to guarantee a large deviation inequality that holds simultaneously for all $f \in L_p(I)$, let alone to guarantee a homogeneous large deviation inequality. On the other hand, it is an interesting aspect of derivatives that having a slower time index can be an advantage. For example, $\sum_{n=1}^{\infty} m\{|D(1/w_n)(f 1_{[0,1)})| \geqslant w_n\} < \infty$ for all $f \in L_1(\mathbb{R})$

as long as (w_n) tends to ∞. See Rosenblatt and Wierdl [3], p. 541. However, it is also observed that there is no homogeneous inequality in this case when $w_n = n$ for all $n \geqslant 1$.

For our purposes, as far as outcomes about large deviations go, it turns out that it does not matter whether we work with the dyadic martingale or dyadic Lebesgue derivatives, although one may be computationally or notationally easier to use than the other in a given situation. The reason for this is the following lemma, which we call the Comparison Lemma. Of course, in the first case we would be working on $L_1[0,1)$ and in the second case we would be working on $L_1(\mathbb{R})$; so we will have to be careful to keep this distinction in mind. A couple of notational conventions are convenient to use here. Given $f \in L_1[0,1)$, we have $E(f|\beta_n) \in L_1[0,1)$, and we take it to be defined on $\mathbb{R}$ by setting it to be 0 off $[0,1)$. Also, given $f \in L_1(\mathbb{R})$, the function $f 1_{[0,1)} \in L_1(\mathbb{R})$ can be taken to be in $L_1[0,1)$, even though it is formally also defined, and equal to zero, outside of $[0,1)$.

Lemma 2.4 *(Comparison Lemma). Given $N \geqslant 1$ and $w > 0$, for a positive function $f \in L_1(\mathbb{R})$, we have*

$$\text{(a)} \qquad m\bigl\{x \in \mathbb{R}\colon\ E(f 1_{[0,1)}|\beta_N)(x) \geqslant w\bigr\} \leqslant 2m\bigl\{x \in \mathbb{R}\colon\ D_N(2f 1_{[0,1)})(x) \geqslant w\bigr\},$$

$$\text{(b)} \qquad m\bigl\{x \in \mathbb{R}\colon\ D_N(f 1_{[0,1)})(x) \geqslant w\bigr\} \leqslant 2m\bigl\{x \in \mathbb{R}\colon\ E(2f 1_{[0,1)}|\beta_N)(x) \geqslant w\bigr\}.$$

Proof. Let $E(f 1_{[0,1)}|\beta_N)(x) \geqslant w$. Then $x \in [0,1)$ and

$$2^N \int_{-1/2^N}^{1/2^N} f 1_{[0,1)}(x+t)\,dt \geqslant w.$$

So, either $2^N \int_0^{1/2^N} f 1_{[0,1)}(x+t)\,dt \geqslant w/2$ or $2^N \int_{-1/2^N}^0 f 1_{[0,1)}(x+t)\,dt \geqslant w/2$. Thus, x must be in

$$\bigl\{u \in \mathbb{R}\colon\ D_N(f 1_{[0,1)})(u) \geqslant w/2\bigr\} \cup \bigl\{u \in \mathbb{R}\colon\ D_N(f 1_{[0,1)})\bigl(u - 1/2^N\bigr) \geqslant w/2\bigr\}.$$

Hence,

$$\begin{aligned} m\bigl\{x \in \mathbb{R}\colon\ E(f 1_{[0,1)}|\beta_N)(x) \geqslant w\bigr\} &\leqslant m\bigl\{x \in \mathbb{R}\colon\ D_N(f 1_{[0,1)})(x) \geqslant w/2\bigr\} \\ &\quad + m\bigl(\bigl\{x + 1/2^N \in \mathbb{R}\colon\ D_N(f 1_{[0,1)})(x) \geqslant w/2\bigr\}\bigr) \\ &\leqslant 2m\bigl\{x \in \mathbb{R}\colon\ D_N(2f 1_{[0,1)})(x) \geqslant w\bigr\}. \end{aligned}$$

This proves (a).

On the other hand, take $x \in \mathbb{R}$ and $D_N(f 1_{[0,1)})(x) \geqslant w$. Then $x \in [-1/2^N, 1)$ and $2^N \int_0^{1/2^N} f 1_{[0,1)}(x+t)\,dt \geqslant w$. If $x \in [-1/2^N, 0)$, then $x + 1/2^N \in \beta_N(1)$ and $2^N \int_{\beta_N(1)} f 1_{[0,1)}(t)\,dt \geqslant w$. If $x \in [1 - 1/2^N, 1)$, then $x \in \beta_N(2^N)$ and $2^N \int_{\beta_N(2^N)} f 1_{[0,1)}(t)\,dt \geqslant w$. Otherwise, $x \in [0, 1 - 1/2^N)$, $x \in \beta_N(l)$ for some $l = 1, \ldots, 2^N - 1$, and

$$2^N \int_{\beta_N(l)} f 1_{[0,1)}(t)\,dt + 2^N \int_{\beta_N(l+1)} f 1_{[0,1)}(t)\,dt \geqslant w.$$

In this case, either $2^N \int_{\beta_N(l)} f 1_{[0,1)}(t)\,dt \geqslant w/2$ or $2^N \int_{\beta_N(l+1)} f 1_{[0,1)}(t)\,dt \geqslant w/2$. So either

$$x \in \beta_N(l) \quad \text{and} \quad 2^N \int_{\beta_N(l)} f 1_{[0,1)}(t)\,dt \geqslant w/2$$

or

$$x + 1/2^N \in \beta_N(l+1) \quad \text{and} \quad 2^N \int_{\beta_N(l+1)} f 1_{[0,1)}(t)\,dt \geqslant w/2.$$

Hence, in all the cases, we would have x in the union of the two sets

$$\bigl\{u \in \mathbb{R}\colon\ E(f 1_{[0,1)}|\beta_N)(u) \geqslant w/2\bigr\}$$

and

$$\bigl\{u \in \mathbb{R}\colon\ E(f 1_{[0,1)}|\beta_N)\bigl(u + 1/2^N\bigr) \geqslant w/2\bigr\}.$$

Therefore,

$$\begin{aligned} m\{x \in \mathbb{R}\colon\ D_N(f 1_{[0,1)})(x) \geqslant w\} &\leqslant m\{x \in \mathbb{R}\colon\ E(2f 1_{[0,1)}|\beta_N)(x) \geqslant w\} \\ &\quad + m(\{x - 1/2^N \in \mathbb{R}\colon\ E(2f 1_{[0,1)}|\beta_N)(x) \geqslant w\}) \\ &\leqslant 2m\{x \in \mathbb{R}\colon\ E(2f 1_{[0,1)}|\beta_N)(x) \geqslant w\}. \end{aligned}$$

This proves (b). □

Remark 2.5. Of course, for $w > 0$, $m\{x \in \mathbb{R}\colon\ E(f 1_{[0,1)}|\beta_N)(x) \geqslant w\} = m\{x \in [0,1)\colon\ E(f 1_{[0,1)}|\beta_N)(x) \geqslant w\}$ because $E(f 1_{[0,1)}|\beta_N)$ is taken to be zero outside $[0,1)$. On the other hand, $D_N(f 1_{[0,1)})$ is only sure to be zero off $[-1/2^n, 1)$ since $D_N(f 1_{[0,1)})(x) > 0$ is possible with $x \in (-1/2^N, 0)$. In any case, this lemma allows us to pass back and forth between dyadic martingales operating on $L_1[0,1)$ and Lebesgue derivatives operating on $L_1(\mathbb{R})$, when proving non-trivial large deviation results or when giving counterexamples to large deviation results. For example, the Comparison Lemma allows us to prove large deviation results for the dyadic martingales operating on $L_1[0,1)$ from large deviation results for the Lebesgue derivatives operating on $L_1(\mathbb{R})$. Indeed, suppose there is a constant C, so that we know that for any $f \in L_1(\mathbb{R})$,

$$\sum_{n=1}^{\infty} m\{x \in \mathbb{R}\colon\ D_n|f|(x) \geqslant w_n\} \leqslant C\|f\|_{L_1(\mathbb{R})}.$$

Now let $g \in L_1[0,1)$. Write $g = f 1_{[0,1)}$ where $f \in L_1(\mathbb{R})$. Then, using Lemma 2.4(a), we obtain:

$$\begin{aligned} \sum_{n=1}^{\infty} m\{x \in [0,1)\colon\ E(|g||\beta_n)(x) \geqslant w_n\} &= \sum_{n=1}^{\infty} m\{x \in \mathbb{R}\colon\ E(|f 1_{[0,1)}||\beta_n)(x) \geqslant w_n\} \\ &\leqslant 2\sum_{n=1}^{\infty} m\{x \in \mathbb{R}\colon\ D_n(2|f 1_{[0,1)}|)(x) \geqslant w_n\} \\ &\leqslant C\|2f 1_{[0,1)}\|_{L_1(\mathbb{R})} = 2C\|g\|_{L_1[0,1)}. \end{aligned}$$

3. Positive results

We can argue that there exist certain positive results about large deviations via transfer from the large deviations for ergodic averages in Rosenblatt and Wierdl [3]. An examination of the large deviation result in [3] as transferred to $\mathbb{Z}$ where it was originally proved, and suitably adapted to a similar statement for $\mathbb{R}$, shows the following large deviation bounds.

Proposition 3.1. *There is a constant C such that for all finite sequences $0 < \rho_1 \leqslant \cdots \leqslant \rho_N$, all $\lambda > 0$, and all $f \in L_1(\mathbb{R})$, we have*

$$\sum_{n=1}^{N} m\{D(\rho_n)|f| \geqslant \lambda n\} \leqslant \frac{C\|f\|_{L_1(\mathbb{R})}}{\lambda}. \tag{3}$$

Proof. This result is implicit in [3], but to make it clear we will use Theorem 5.11 in [3] in the case that the dimension $k = 1$ to prove it. We will use the notation from this article. Fix a sequence (v_n) in $\mathbb{R}$ which will be specified later. Let $I_n = [0, \rho_{N-n+1}] + v_n$ for $n = 1, \dots, N$ and $I_n = I_N$ for $n > N$. Since (ρ_n) is non-decreasing and the sets I_n are intervals, it is easy to see that, with $I_n - I_i = \{x - y\colon\ x \in I_n,\ y \in I_i\}$,

$$Q_n = m\left(\bigcup_{i=n}^{\infty}(I_n - I_i)\right) = m\left(\bigcup_{i=n}^{N}(I_n - I_i)\right) \leqslant \sum_{i=n}^{N} \rho_{N-n+1} + \rho_{N-i+1} \leqslant 2\sum_{i=n}^{N} \rho_{N-n+1} = 2(N-n+1)\rho_{N-n+1}.$$

Now if $I = [0, \rho] + v$, then $N(I, Q)f(x) = \frac{\rho}{Q} D(\rho) f(x+v)$. Hence, for a positive function $f \in L_1(\mathbb{R})$, we have

$$N(I_n, Q_n) f(x) \geqslant \frac{1}{2(N-n+1)} D(\rho_{N-n+1}) f(x + v_n).$$

But if in addition we know that f has bounded support, then for suitably widely spaced (v_n) the functions $N(I_n, Q_n)f$ are disjointly supported. By Theorem 5.11 [3], using sup for the essential supremum, we have

$$m\left\{\sup_{1 \leqslant n \leqslant N} N(I_n, Q_n) f > \lambda\right\} \leqslant \frac{1}{\lambda}\|f\|_{L_1(\mathbb{R})}.$$

This tells us that we have

$$\begin{aligned}\sum_{n=1}^{N} m\{D(\rho_n)f > 2\lambda n\} &= \sum_{n=1}^{N} m\{x\colon D(\rho_{N-n+1})f(x+v_n) > 2\lambda(N-n+1)\}\\ &= m\left\{x\colon \sup_{1\leqslant n\leqslant N}\frac{1}{2(N-n+1)}D(\rho_{N-n+1})f(x+v_n) > \lambda\right\}\\ &\leqslant m\left\{\sup_{1\leqslant n\leqslant N} N(I_n, Q_n)f > \lambda\right\}\\ &\leqslant \frac{1}{\lambda}\|f\|_{L_1(\mathbb{R})}.\end{aligned}$$

Replacing λ by half its value gives then

$$\sum_{n=1}^{N} m\{D(\rho_n)f > \lambda n\} \leqslant \frac{2\|f\|_{L_1(\mathbb{R})}}{\lambda}.$$

Then for all large enough $k \geqslant 1$, we have $\sum_{n=1}^{N} m\{D(\rho_n)f > (\lambda - \frac{1}{k})n\} \leqslant \frac{2\|f\|_{L_1(\mathbb{R})}}{\lambda - \frac{1}{k}}$. Letting k tend to infinity gives Eq. (3) for positive $f \in L_1(\mathbb{R})$ with bounded support. Now the subadditive operator on $L_1(\mathbb{R})$ given by $\sup_{n=1,\ldots,N} D(\rho_n)f$ is continuous in measure, so Eq. (3) holds for all positive functions in $L_1(\mathbb{R})$ by approximating the general function by functions with bounded support. We have Eq. (3) as claimed with the constant $C = 2$. $\square$

As a consequence, there is a constant C such that if we have a finite non-increasing sequence $(m_n\colon a \leqslant n \leqslant b)$, then for any $M \geqslant 1$, we have

$$\sum_{n=a}^{b} m\{D_{m_n}|f| \geqslant M(n-a+1)\} \leqslant \frac{C}{M}\|f\|_{L_1(\mathbb{R})}. \tag{4}$$

This is because $D_{m_n}f = D(\frac{1}{2^{m_n}})f$, so that for $\rho_n = \frac{1}{2^{m_n}}$, $n = 1, \ldots, N$ to be non-decreasing, we need to have $(m_n\colon n = 1, \ldots, N)$ to be non-increasing.

Reversing the summation index gives this: there is a constant C such that for $b \geqslant a \geqslant 1$, and a non-decreasing sequence $(m_n\colon a \leqslant n \leqslant b)$, we have

$$\sum_{n=a}^{b} m\{D_{m_n}|f| \geqslant M(b-n+1)\} \leqslant \frac{C}{M}\|f\|_{L_1(\mathbb{R})}. \tag{5}$$

From these facts, we can obtain positive results for large deviations for both the dyadic martingale and the dyadic Lebesgue derivatives. For example, we have this result.

Proposition 3.2. *There exists a constant C and a sequence (w_n) such that $\lim_{n\to\infty} w_n = \infty$ and $\sum_{n=1}^{\infty}\frac{1}{w_n} = \infty$, so that for $f \in L_1(\mathbb{R})$ we have*

$$\sum_{n=1}^{\infty} m\{x \in \mathbb{R}\colon D_{m_n}|f|(x) \geqslant w_n\} \leqslant C\|f\|_{L_1(\mathbb{R})}$$

and for $g \in L_1[0,1)$ we have

$$\sum_{n=1}^{\infty} m\{x \in [0,1)\colon E(|g|\,|\beta_n)(x) \geqslant w_n\} \leqslant 2C\|g\|_{L_1[0,1)}.$$

Proof. We will construct (w_n) in blocks $B_k = \{M_k + 1, \ldots, M_{k+1}\}$ where (M_k) is strictly increasing and $M_1 \geqslant 1$. We can take $w_n = 2^k(M_{k+1} - n + 1)$ for $n = M_k + 1, \ldots, M_{k+1}$. Then we have $\sum_{n=1}^{\infty}\frac{1}{w_n} = \sum_{k=1}^{\infty}\sum_{n=M_k+1}^{M_{k+1}}\frac{1}{2^k(M_{k+1}-n+1)}$. With an appropriately rapidly increasing (M_k), we have $\sum_{n=1}^{\infty}\frac{1}{w_n} = \infty$. Indeed, $M_k = 2^{2^k}$ will do. But also, for all $f \in L_1(\mathbb{R})$, we have by Eq. (5),

$$\sum_{n=1}^{\infty} m\{x \in \mathbb{R}\colon D_n|f| \geqslant w_n\} = \sum_{k=1}^{\infty}\sum_{n=M_k+1}^{M_{k+1}} m\{D_n|f| \geqslant 2^k(M_{k+1} - n + 1)\} \leqslant \sum_{k=1}^{\infty}\frac{C}{2^k}\|f\|_{L_1(\mathbb{R})} = C\|f\|_{L_1(\mathbb{R})}.$$

By Remark 2.5 we also have for a function $g \in L_1[0,1)$,

$$\sum_{n=1}^{\infty} m\bigl\{x \in [0,1): E\bigl(|g|\big|\beta_n\bigr)(x) \geqslant w_n\bigr\} \leqslant 2C\|g\|_{L_1[0,1)}. \qquad \square$$

We can obtain a related result with the roles of the levels and the subsequence of the operators exchanged.

Proposition 3.3. *There exists a constant C, (m_n) such that $\lim_{n\to\infty} m_n = \infty$ and a non-decreasing sequence (w_n) such that $\sum_{n=1}^{\infty} \frac{1}{w_n} = \infty$, so that for $f \in L_1(\mathbb{R})$ we have*

$$\sum_{n=1}^{\infty} m\bigl\{x \in \mathbb{R}: D_n|f|(x) \geqslant w_n\bigr\} \leqslant C\|f\|_{L_1(\mathbb{R})}$$

and for $g \in L_1[0,1)$ we have

$$\sum_{n=1}^{\infty} m\bigl\{x \in [0,1): E\bigl(|g|\big|\beta_n\bigr)(x) \geqslant w_n\bigr\} \leqslant 2C\|g\|_{L_1[0,1)}.$$

Proof. We let $M_k = 2^{2^k}$ as in the previous result. Again, construct the sequences in blocks $B_k = \{M_k+1,\dots,M_{k+1}\}$. We take $w_n = 2^k n$ for each $n \in B_k$. It is clear that (w_n) is non-decreasing. But also as in the proof of Proposition 3.2, the choice of (M_k) guarantees that we have $\sum_{n=1}^{\infty} \frac{1}{w_n} = \infty$. We could define (m_n) by taking $m_n = M_{k+1} - n + 1$ for $n \in B_k$. But then $m_{M_{k+1}} = 1$. So (m_n) will be unbounded but it will not converge to ∞. Hence, instead we will let $m_n = M_{k+1} - n + 1$ for $M_k + 1 \leqslant n \leqslant M_{k+1} - M_k$ and $m_n = M_k + 1$ for $M_{k+1} - M_k + 1 \leqslant n \leqslant M_{k+1}$. Then (m_n) does converge to ∞. Now we can use the large deviation fact above because the (m_n) are non-increasing on the block B_k. Thus, for all $f \in L_1(\mathbb{R})$ by Eq. (4),

$$\begin{aligned}
\sum_{n=1}^{\infty} m\bigl\{x \in \mathbb{R}: D_{m_n}|f| \geqslant w_n\bigr\} &= \sum_{k=1}^{\infty} \sum_{n=M_k+1}^{M_{k+1}} m\bigl\{x \in \mathbb{R}: D_{m_n}|f| \geqslant 2^k n\bigr\} \\
&\leqslant \sum_{k=1}^{\infty} \sum_{n=M_k+1}^{M_{k+1}} m\bigl\{x \in R: D_{m_n}|f| \geqslant 2^k(n - M_k)\bigr\} \\
&\leqslant \sum_{k=1}^{\infty} \frac{C}{2^k}\|f\|_{L_1(\mathbb{R})} \\
&= C\|f\|_{L_1(\mathbb{R})}.
\end{aligned}$$

By Remark 2.5 we also have for a function $g \in L_1[0,1)$,

$$\sum_{n=1}^{\infty} m\bigl\{x \in [0,1): E\bigl(|g|\big|\beta_n\bigr)(x) \geqslant w_n\bigr\} \leqslant 2C\|g\|_{L_1[0,1)}. \qquad \square$$

Remark 3.4. The operators T_n that we consider in this paper are all positive operators. So if $f \in L_1(I)$, then $|T_n f| \leqslant T_n|f|$. Stating Proposition 3.3 in terms of $D_n|f|$ and $E(|g||\beta_n)$ is formally stronger than stating it in terms of $|D_n f|$ and $|E(g|\beta_n)|$ respectively. But it is not hard to see that the existence of a large deviation result for particular sequences (m_n) and (w_n), and all functions, is independent of whether one uses $D_n|f|$ or $|D_n f|$ (respectively $E(|f||\beta_n)$ or $|E(g|\beta_n)|$), although the value of the constant C may change.

These results can be merged into one more specific result. First, we will be using not just any rearrangement of the whole numbers.

Definition 3.5. A block rearrangement $\pi: \mathbb{N} \to \mathbb{N}$ is a one-to-one, onto map such that there are finite disjoint subsets $B_k \subset \mathbb{N}$, $\mathbb{N} = \bigsqcup_k B_k$, and $\pi: B_k \to B_k$ is one-to-one and onto.

Not all rearrangements are like this, of course. Consider, for example, rearrangement π_0 defined by:

$$\pi_0(n) = \begin{cases} n+2, & \text{if } n \geqslant 2,\ n \text{ even}, \\ 2, & \text{if } n = 1, \\ n-2, & \text{if } n \geqslant 3,\ n \text{ odd}, \end{cases}$$

$$7 \xrightarrow{\pi_0} 5 \xrightarrow{\pi_0} 3 \xrightarrow{\pi_0} 1 \xrightarrow{\pi_0} 2 \xrightarrow{\pi_0} 4 \xrightarrow{\pi_0} 6 \ldots.$$

All orbits $\pi_0^{(j)}(k)$ are infinite and π_0 is not a block rearrangement; indeed, it has no finite invariant subsets.

In the next proposition we are taking $w_n = n \ln n$, so $\sum_{n=2}^{\infty} \frac{1}{w_n} = \infty$.

Proposition 3.6. *There exists a constant C and a block rearrangement $\pi : \mathbb{N} \to \mathbb{N}$ such that for $f \in L_1(\mathbb{R})$ we have*

$$\sum_{n=1}^{\infty} m\bigl\{x \in \mathbb{R}:\ D_{\pi(n)}|f|(x) \geqslant n \ln n\bigr\} \leqslant C\|f\|_{L_1(\mathbb{R})}$$

and for $g \in L_1[0,1)$ we have

$$\sum_{n=1}^{\infty} m\bigl\{x \in [0,1):\ E\bigl(|g|\big|\beta_{\pi(n)}\bigr)(x) \geqslant n \ln n\bigr\} \leqslant 2C\|g\|_{L_1[0,1)}.$$

Proof. Again, according to Remark 2.5 it is enough to consider the case $D_n f$ with $f \in L_1(\mathbb{R})$. We build π in blocks $B_k = [M_k + 1, \ldots, M_{k+1}]$ where $M_k = 2^{2^k}$. Consider the map $B_k \xrightarrow{\pi} B_k$, $\pi(n) = M_{k+1} + M_k - n + 1$. This is a one-to-one, onto map of B_k to B_k. Since $\ln n \geqslant 2^k$ for all $n \in B_k$, and $\pi(n)$ is non-increasing on B_k, using (4) we have:

$$\begin{aligned} \sum_{n \in B_k} m\bigl\{x \in \mathbb{R}:\ D_{\pi(n)}|f| \geqslant n \ln n\bigr\} &\leqslant \sum_{n \in B_k} m\bigl\{x \in \mathbb{R}:\ D_{\pi(n)}|f| \geqslant n2^k\bigr\} \\ &\leqslant \sum_{n \in B_k} m\bigl\{x \in \mathbb{R}:\ D_{\pi(n)}|f| \geqslant (n - M_k)2^k\bigr\} \\ &= \frac{C}{2^k}\|f\|_{L_1(\mathbb{R})}. \end{aligned}$$

Hence,

$$\sum_k \sum_{n \in B_k} m\bigl\{x \in \mathbb{R}:\ D_{\pi(n)}|f| \geqslant n \ln n\bigr\} \leqslant C\|f\|_1. \qquad \square$$

Remark 3.7. In this result it is critical that one allows a rearrangement of the terms of the series, since the result fails if one does not. See Example 4.4(b).

Remark 3.8. We would like to make the trivial observation that we can rearrange the operators or the levels. Indeed, by taking $\pi_o = \pi^{-1}$ with π as above, we get for some block rearrangement π_o of $\mathbb{N}$,

$$\sum_{n=1}^{\infty} m\bigl\{x \in \mathbb{R}:\ D_n|f| \geqslant \pi_o(n) \ln \pi_o(n)\bigr\} \leqslant C\|f\|_{L_1(\mathbb{R})}.$$

4. Negative results

Unlike in the previous section, when we insist that both the levels (w_n) and the time index (m_n) are non-decreasing, we typically get negative results for large deviation inequalities.

Conjecture 4.1. *For any non-decreasing sequence (m_n) such that $m_n \geqslant n$ for all $n \geqslant 1$, and any levels (w_n) which are non-decreasing such that $\sum_{n=1}^{\infty} \frac{1}{w_n} = \infty$, there exists some $f \in L_1[0,1)$ such that*

$$\sum_{n=1}^{\infty} m\bigl\{x \in [0,1):\ \bigl|E(f|\beta_{m_n})(x)\bigr| \geqslant w_n\bigr\} = \infty.$$

Similarly, with the same conditions on (m_n) and (w_n), there exists $f \in L_1(\mathbb{R})$ such that

$$\sum_{n=1}^{\infty} m\bigl\{x \in \mathbb{R}:\ \bigl|D_{m_n} f(x)\bigr| \geqslant w_n\bigr\} = \infty.$$

We cannot prove this conjecture at this time, but we can prove results that give many examples supporting this conjecture. First we handle the levels (w_n) and then we consider time index (m_n).

4.1. The levels (w_n)

The way the following result is used in support of the conjecture above is that we take explicit examples of (w_n) which are growing without bound, but still for which $\sum_{n=1}^{\infty} \frac{1}{w_n} = \infty$, and then we try to find a suitable function g that meets the hypotheses needed below to make the construction work correctly.

Theorem 4.2. *Let $w : \mathbb{R}^+ \to \mathbb{R}^+$, $w_n = w(n)$. Let w be non-decreasing and $\sum_{n=1}^{\infty} \frac{1}{w_n} = \infty$. Suppose there exists a non-negative function g on $\mathbb{R}^+$ such that*

(i) $\sum_{k=1}^{\infty} g(k) = \infty$,
(ii) $[2^{k+1} g(k+1)] - [2^k g(k)] \geqslant 2$, *and*
(iii) $\frac{w(2^k g(k))}{2^k}$ *is decreasing to* 0.

Then there exists positive $f \in L_1(\mathbb{R})$, $f = 0$ off $[0, 1)$ such that

$$\sum_{n=1}^{\infty} m\{x \in \mathbb{R}:\ D_n f(x) \geqslant w_n\} = \infty.$$

Also, there exists positive $h \in L_1[0, 1)$ such that

$$\sum_{n=1}^{\infty} m\{x \in [0, 1):\ E(h|\beta_n)(x) \geqslant w_n\} = \infty.$$

Proof. It is enough to prove the first statement. The second statement follows for $h = f 1_{[0,1)}$ from Comparison Lemma 2.4 and the fact that $m\{x \in \mathbb{R}:\ E(f 1_{[0,1)}|\beta_n)(x) > w\} = m\{x \in [0, 1):\ E(f 1_{[0,1)}|\beta_n)(x) > w\}$ for $w > 0$.

We denote by $|I|$ the length of interval I, by $[x]$ the largest integer not exceeding x, by $\bigsqcup_{i=1}^{k} A_i$ the union of disjoint sets $A_1, \dots, A_k$, and for notational convenience in this proof, by $E(B)f$ the average $\frac{1}{m(B)} \int_B f\,dm$ of the function f over the set B.

For $a \in [0, 1)$ we write $a = \sum_{i=1}^{\infty} a_i 2^{-i}$. Define sets

$$A_1 = \{a:\ a_{[2^2 g(2)]} = 0\},$$

and for $k \geqslant 2$

$$A_k = \{a:\ a_{[2^2 g(2)]} = \cdots = a_{[2^k g(k)]} = 1,\ a_{[2^{k+1} g(k+1)]} = 0\},$$
$$B_k = \{a:\ a_{[2^2 g(2)]} = \cdots = a_{[2^k g(k)]} = 1\}.$$

Notice that sets A_k are disjoint, $m(A_k) = 2^{-k}$, $m(\bigsqcup_{k=1}^{\infty} A_k) = 1$, $m(B_k) = 2^{1-k}$, $B_k = [0, 1) \setminus \bigsqcup_{j=1}^{k-1} A_j$.

Let

$$r(k) = \frac{w(2^{k+1} g(k+1))}{2^{k+1}},$$

so by assumption (iii) $r(k) \searrow 0$. Define $f \geqslant 0$ on $[0, 1)$ by assigning

$$f(a) = 2^{i+2}\big(r(i) - r(i+1)\big), \quad a \in A_i$$

and $f = 0$ off $[0, 1)$. Then

$$\int_{\mathbb{R}} f\,dm = \int_0^1 f\,dm = \sum_{i=1}^{\infty} \int_{A_i} f\,dm = \sum_{i=1}^{\infty} 2^{i+2}\big(r(i) - r(i+1)\big)2^{-i} = 4r(1),$$

so $f \in L_1(\mathbb{R})$. On B_k we have

$$\int_{B_k} f\,dm = \sum_{i=k}^{\infty} \int_{A_i} f\,dm = 4r(k).$$

Let $M_k = w(2^{k+1}g(k+1))$. The average of f on B_k is

$$E(B_k)f = 2^{k-1}\int_{B_k} f\,dm = 2^{k+1}r(k) = M_k.$$

Each B_k consists of intervals B_k^i, $i = 1,\dots,l_k$, where $|B_k^i| = 2^{-[2^k g(k)]}$, $i = 1,\dots,l_k$, with $l_k = 2^{[2^k g(k)]-k+1}$. Similarly, each A_k consists of intervals A_k^j, and each B_{k+1} consists of intervals B_{k+1}^j, each of length $|A_k^j| = |B_{k+1}^j| = 2^{-[2^{k+1}g(k+1)]}$ with $j = 1,\dots,l_{k+1}$. Note that

$$B_k = A_k \sqcup B_{k+1} = A_k \sqcup \left(\bigsqcup_{n=k+1}^{\infty} A_n\right).$$

On each B_k^i and, hence, on B_k the function f consists of repeating pieces. More precisely, each B_k^i is the union of L adjacent disjoint dyadic intervals of the form $A_k^j \sqcup B_{k+1}^j$, where $L = 2^{[2^{k+1}g(k+1)]-[2^k g(k)]-1}$ and A_k^j, B_{k+1}^j are left and right halves of a dyadic interval. From the definition of f it is clear that f looks the same on each of these L sets, which we will call periods of length T. Then $T = |A_k^j| + |B_{k+1}^j| = 2^{-[2^{k+1}g(k+1)]+1}$. Notice that

$$E(B_k)f = E(I)f = M_k$$

where I is any interval inside some B_k^i whose length is a multiple of T.

Fix k. By assumption $[2^{k+1}g(k+1)] - [2^k g(k)] \geqslant 2$, and we may consider integers n that satisfy

$$\left[2^k g(k)\right] + 1 \leqslant n \leqslant \left[2^{k+1}g(k+1)\right] - 1. \tag{6}$$

For such n we see that $M_k = w(2^{k+1}g(k+1)) \geqslant w_n$. For any interval I with $|I| = 2^{-n}$ the second inequality in (6) shows that $|I|$ is a multiple of T. Therefore, if $|I| = 2^{-n}$ and $I \subseteq B_k^i$ then $E(I)f = M_k$. Let α and β denote the left and right ends of the interval B_k^i. For each x between α and $\beta - 2^{-n}$ we may consider interval $I_x = (x, x+2^{-n}) \subseteq B_k^i$. Note that $D_n f(x) = E(I_x)f = M_k$. Then using the first inequality in (6) we see that

$$m\left\{x \in B_k^i\colon D_n f(x) = M_k\right\} \geqslant \left|B_k^i\right| - 2^{-n} \geqslant \frac{1}{2}\left|B_k^i\right|.$$

Therefore

$$m\left\{x \in B_k\colon D_n f(x) = M_k\right\} \geqslant \frac{1}{2}m(B_k) = \frac{1}{2^k}.$$

Then

$$\begin{aligned}
\sum_{n=1}^{\infty} m\left\{x\colon D_n f(x) \geqslant w_n\right\} &\geqslant \sum_{k=2}^{\infty}\sum_{n=[2^k g(k)]+1}^{[2^{k+1}g(k+1)]-1}\left\{x\colon D_n f(x) \geqslant w_n\right\}\\
&\geqslant \sum_{k=2}^{\infty}\sum_{n=[2^k g(k)]+1}^{[2^{k+1}g(k+1)]-1}\left\{x\colon D_n f(x) = M_k\right\}\\
&\geqslant \sum_{k=2}^{\infty}\sum_{n=[2^k g(k)]+1}^{[2^{k+1}g(k+1)]-1}\left\{x \in B_k\colon D_n f(x) = M_k\right\}\\
&\geqslant \sum_{k=2}^{\infty}\sum_{n=[2^k g(k)]+1}^{[2^{k+1}g(k+1)]-1}\frac{1}{2^k}\\
&= \sum_{k=2}^{\infty}\frac{1}{2^k}\left(\left[2^{k+1}g(k+1)\right] - \left[2^k g(k)\right] - 1\right)\\
&\geqslant \sum_{k=2}^{\infty}\frac{1}{2^k}\left(2^{k+1}g(k+1) - 1 - 2^k g(k) - 1\right)\\
&= -1 + \sum_{k=2}^{\infty} 2g(k+1) - g(k)\\
&= \infty.
\end{aligned}$$

The last equality follows easily because

$$\begin{aligned}\sum_{k=2}^{M} 2g(k+1)-g(k) &= 2g(3)-g(2)+\cdots+2g(M+1)-g(M)\\ &= -g(2)+g(3)+\cdots+g(M)+2g(M+1)\\ &\geqslant -g(2)+\sum_{k=3}^{M} g(k),\end{aligned}$$

which tends to ∞ as M tends to ∞. $\square$

Remark 4.3.

(a) It is enough to have the criteria of Theorem 4.2 satisfied eventually.
(b) One instance when (ii) of Theorem 4.2 holds is when $\alpha_k = 2^{k+1}g(k+1)-2^k g(k)$ is non-decreasing. In this case α_k is unbounded and so $\alpha_k \nearrow \infty$. To show that α_k is unbounded, assume that $\alpha_k \leqslant C$ for all k, i.e. $g(k+1) \leqslant 2^{-1}g(k)+2^{-k-1}C$. Then one can show by induction that for all k,

$$g(2k+1) \leqslant \frac{g(1)}{2^{2k}} + \frac{C}{2^{k+1}},$$
$$g(2k+2) \leqslant \frac{g(1)}{2^{2k+1}} + \frac{C}{2^{k+2}} + \frac{C}{2^{2k+2}}.$$

This contradicts the assumption $\sum_{k=1}^{\infty} g(k) = \infty$.

Example 4.4.

(a) It is not hard to check that Theorem 4.2 holds for a given (w_n), once the right candidate for g is determined. For example, if (w_n) is (n), $(n \ln n)$, or $(n \ln n \ln \ln n)$, then we can take $g(k) = \frac{1}{w(k)}$. Unfortunately, other examples do not work with such a simple formula for g.
(b) For example, with $w_n = n \ln n$ for all n, and $g(k) = \frac{1}{w(k)}$ for all k, one gets $\frac{w(2^k g(k))}{2^k} = \frac{\ln 2}{\ln k} + O(\frac{1}{k})$. So $\frac{w(2^k g(k))}{2^k}$ converges to 0 and is decreasing eventually. This is why we can make the assertion in Remark 3.7, that for (T_n) the dyadic martingale (respectively, the dyadic Lebesgue derivatives) there exists a positive function $f \in L_1[0,1)$ (respectively, a positive function $f \in L_1(\mathbb{R})$ with bounded support) such that

$$\sum_{n=1}^{\infty} m\{T_n f(x) \geqslant n \ln n\} = \infty.$$

(c) When $w_n = w(n) = n \ln n \ln \ln n$, and we take $g(k) = \frac{1}{w(k)}$, then it is a little harder to verify the criteria, but they do hold. For example, now $\frac{w(2^k g(k))}{2^k} = \frac{\ln 2}{\ln \ln k} + O(\frac{\ln k}{k \ln \ln k})$, and so $\frac{w(2^k g(k))}{2^k}$ converges to 0 and is also decreasing eventually.

Remark 4.5. Suppose that $w: \mathbb{R}^+ \to \mathbb{R}^+$ is non-decreasing, $w_n = w(n)$ and $\sum_{n=1}^{\infty} \frac{1}{w_n} = \infty$. One might need some additional properties of w in order to construct g so that the conditions of Theorem 4.2 hold. Indeed, suppose (y_k) is a non-decreasing sequence of positive numbers with $\sum_{k=1}^{\infty} \frac{1}{y_k} = \infty$. Because of Example 4.4(a) there is no harm in assuming that $\limsup \frac{y_k}{k} > 0$. Then there is a non-decreasing sequence $w(k)$ equivalent to y_k in the sense that $y_k \leqslant w(k) \leqslant 5y_k$ such that $\sum_{n=1}^{\infty} \frac{1}{w_n} = \infty$, but if g satisfies (i) and (ii) of Theorem 4.2, then (iii) fails, i.e. $(\frac{w(2^k g(k))}{2^k})$ is not decreasing. To see this, we modify an example provided by V. Vlasak. We may define (w_k) as follows: $w(1) = 2y_1$; and for $k \geqslant 1$ let $w(k+1) = w(k)$ if $4w(k) > 5y_{k+1}$ and $w(k+1) = \max\{4w(k), w(k)+y_{k+1}-y_k\}$ if $4w(k) \leqslant 5y_{k+1}$. By induction it is easy to check that $y_k \leqslant w(k) \leqslant 5y_k$. Note that the set $I = \{i\colon\ 4w(i) \leqslant 5y_{i+1}\}$ is infinite because of the assumption $\limsup \frac{y_k}{k} > 0$. Let $i \in I$, i.e. $w(i+1) \geqslant 4w(i)$, and let function g satisfy (i) and (ii) of Theorem 4.2. There exists k such that $2^k g(k) \leqslant i$, $2^{k+1}g(k+1) > i$. Then $w(2^{k+2}g(k+2)) \geqslant w(i+1) \geqslant 4w(i) \geqslant 4w(2^k g(k))$, and $\frac{w(2^{k+2}g(k+2))}{2^{k+2}} \geqslant \frac{w(2^k g(k))}{2^k}$. Since I is an infinite set, we see that (iii) fails.

4.2. The time index (m_n)

The second result in support of Conjecture 4.1 is the following. It turns out that with the dyadic martingale and with the dyadic Lebesgue derivatives, once one has a negative result for a certain level sequence (w_n), one gets such a negative result with all time indices increasing faster than in the basic dyadic case. This is important at least because Lemma 2.1

shows that this subsequence behavior is not true if one first fixes the function. It is also important because the principle does not hold for all sequences of operators. For example, suppose that we have a sequence of operators (T_n) on $L_p(I)$ as in the beginning of Section 2. Suppose that $\|T_n\|_p$ tends to 0 as $n \to \infty$. Then, for any fixed sequence of levels (w_n), there is always a subsequence (T_{m_n}) for which there is the homogeneous large deviation inequality

$$\sum_{n=1}^{\infty} m\big\{x \in I\colon\ T_{m_n}|f| \geqslant w_n\big\} \leqslant C\|f\|_{L_p(I)}^p.$$

To show why we can speed up the index for dyadic martingales and dyadic Lebesgue derivatives, we prove what we call the Multiplicity Lemma, a method of changing facts about the distribution of a martingale into facts about the distribution of a subsequence of that martingale. We believe that this result may have interest independent of the specific application in this article. The result is this.

Lemma 4.6 *(Multiplicity Lemma). Given (m_n) increasing with $m_n \geqslant n$ for all $n \geqslant 1$, there exists a measure-preserving transformation τ of $[0, 1)$ such that for all $k \geqslant 1$ and all $y \in \mathbb{R}$, we have for all $f \in L_1[0, 1)$,*

$$m\big\{E(f|\beta_k) = y\big\} = m\big\{E(f \circ \tau|\beta_{m_k}) = y\big\}. \tag{7}$$

This Multiplicity Lemma was originally proved in steps and the argument was long and not very revealing of what the structure of τ really is. The referee for this article observed that the construction could be greatly simplified; we give here the referee's construction with some details to show that the definition of τ gives the type of mapping that we want. We want to thank the referee for this very important improvement to our paper, as well as for other valuable input that was also given to us.

Proof of the Multiplicity Lemma 4.6. First, up to a measure-theoretic isomorphism, in the usual manner using dyadic expansions, we may identify $[0, 1]$ with Lebesgue measure with the product space $P = \prod_{k=1}^{\infty}\{0, 1\}$ with the product measure $\pi = \prod_{k=1}^{\infty}\mu_k$ with each μ_k being the discrete measure given by $\mu_k(\{0\}) = \mu_k(\{1\}) = 1/2$. With this isomorphism, our dyadic σ-algebras correspond to the σ-algebras in P determined by the first k coordinates, which we again denote by β_k. We define the mapping $\tau: P \to P$ by $\tau(\epsilon_1, \epsilon_2, \epsilon_3, \ldots) = (\epsilon_{m_1}, \epsilon_{m_2}, \epsilon_{m_3}, \ldots)$. Clearly, τ preserves the measure π in the sense that for all π-measurable sets E, $\pi(\tau^{-1}E) = \pi(E)$. This lemma claims that for all $f \in L_1(P, \pi)$, we have $\pi\{E(f|\beta_k) = y\} = \pi\{E(f \circ \tau|\beta_{m_k}) = y\}$ for all k and y.

Let $\beta_k(i)$ be an atom where $E(f|\beta_k) = y$; that is, $y = 2^k \int_{\beta_k(i)} f\,d\pi$. Suppose that specifically $\beta_k(i)$ consists of all sequences $\bar{\epsilon} \in P$ with $\bar{\epsilon} = (\epsilon_1(i), \ldots, \epsilon_k(i), \bar{t})$ for some particular $\epsilon_k(i) \in \{0, 1\}$ and a variable $\bar{t} = (t_1, t_2, t_3, \ldots)$ with each $t_s \in \{0, 1\}$. Let $\beta_{m_k}(j)$ be an atom in β_{m_k} given by the vectors $(e_1(j), \ldots, e_{m_k}(j), \bar{t})$ where the values $e_s(j)$, $s = 1, \ldots, m_k$ are fixed, $e_{m_l}(j) = \epsilon_l(i)$ for all $l = 1, \ldots, k$, and the values t_s are allowed to vary in $\{0, 1\}$. The union of the atoms $\beta_{m_k}(j)$ of this type gives $\tau^{-1}\beta_k(i)$.

Because of the definition of τ,

$$\int\limits_{\beta_{m_k}(j)} f \circ \tau\,d\pi = \int f \circ \tau\big(e_1(j), \ldots, e_{m_k}(j), \bar{t}\big)\,d\pi(\bar{t})$$

is the same value for each $\beta_{m_k}(j)$ as above. But also,

$$\int\limits_{\beta_k(i)} f\,d\pi = \int\limits_{\tau^{-1}\beta_k(i)} f \circ \tau\,d\big(\pi \circ \tau^{-1}\big) = \int\limits_{\tau^{-1}\beta_k(i)} f \circ \tau\,d\pi = \sum_{\beta_{m_k}(j) \subset \tau^{-1}\beta_k(i)} \int\limits_{\beta_{m_k}(j)} f \circ \tau\,d\pi.$$

Since there are $\frac{2^{m_k}}{2^k}$ atoms $\beta_{m_k}(j) \subset \tau^{-1}\beta_k(i)$, we always have $\int_{\beta_k(i)} f\,d\pi = \frac{2^{m_k}}{2^k}\int_{\beta_{m_k}(j)} f \circ \tau\,d\pi$. Hence, for any $\beta_{m_k}(j)$ as above, we have

$$y = 2^k \int\limits_{\beta_k(i)} f\,d\pi = 2^k\left(\frac{2^{m_k}}{2^k}\right) \int\limits_{\beta_{m_k}(j)} f \circ \tau\,d\pi = 2^{m_k} \int\limits_{\beta_{m_k}(j)} f \circ \tau\,d\pi.$$

That is, the atoms $\beta_k(i)$ where $E(f|\beta_k) = y$ give us all of the atoms $\beta_{m_k}(j)$ where $E(f \circ \tau|\beta_{m_k}) = y$ by taking the atoms $\beta_{m_k}(j) \subset \tau^{-1}\beta_k(i)$ for some such value of i. Moreover, distinct atoms $\beta_k(i_1)$ and $\beta_k(i_2)$ on which $E(f|\beta_k) = y$ give pairwise disjoint atoms $\beta_{m_k}(j_1)$ and $\beta_{m_k}(j_2)$ corresponding to $\beta_k(i_1)$ and $\beta_k(i_2)$ respectively as above. Hence, denoting the cardinality of a set A by $\#A$, we have for all k and y,

$$\begin{aligned} m\{E(f|\beta_k)=y\} &= \frac{1}{2^k}\#\left\{i:\ 2^k\int\limits_{\beta_k(i)} f\,d\pi = y\right\} \\ &= \frac{1}{2^k}\left(\frac{2^k}{2^{m_k}}\right)\#\left\{j:\ \beta_{m_k}(j)\subset\tau^{-1}\beta_k(i)\text{ where } 2^k\int\limits_{\beta_k(i)} f\,d\pi = y\right\} \\ &= \frac{1}{2^{m_k}}\#\left\{j:\ 2^{m_k}\int\limits_{\beta_{m_k}(j)} f\circ\tau\,d\pi = y\right\} \\ &= m\{E(f\circ\tau|\beta_{m_k})=y\}. \qquad \square \end{aligned}$$

The Multiplicity Lemma 4.6 and the Comparison Lemma 2.4 allow us to improve the results that come out of Theorem 4.2. The following is the best we can do at this time in support of Conjecture 4.1.

Corollary 4.7. *If (w_n) satisfies the conditions of Theorem 4.2, then for any sequence (m_n) that is increasing and has $m_n \geqslant n$ for all $n \geqslant 1$, with (T_n) being the dyadic martingale (respectively, the dyadic Lebesgue derivatives) there exists $f \in L_1[0,1)$ (respectively, $f \in L_1(\mathbb{R})$ with bounded support) such that*

$$\sum_{n=1}^{\infty} m\{|T_{m_n} f| \geqslant w_n\} = \infty.$$

Remark 4.8. For example, with $w_n = n$ and $m_n = n$ for all $n \geqslant 1$, we do not have a large deviation inequality for the dyadic martingale. Although, according to Corollary 2.2, speeding up the time index can lead to a large deviation inequality function by function, the Multiplicity Lemma shows that it cannot ever lead to one that holds simultaneously for all the functions in $L_1[0,1)$.

Remark 4.9. We do not know yet what to say about

$$\sum_{n=1}^{\infty} m\{x\in\mathbb{R}:\ |D(\epsilon_n)f(x)| \geqslant w_n\}$$

when (w_n) is increasing, $\sum_{n=1}^{\infty}\frac{1}{w_n}=\infty$ and $\epsilon_n = O(\frac{1}{n^k})$, $k=2,3,\ldots$. For example, take the case $\epsilon_n = \frac{1}{n^2}$ and $w_n = n$ for all n. We cannot use the Comparison Lemma 2.4 and Theorem 4.2, or the Multiplicity Lemma 4.6, until we take in place of (ϵ_n) a subsequence (ϵ_{m_n}) with $m_n \geqslant \sqrt{2^n}$. So we do not know at this time whether or not there can be a large deviation inequality in the case that $\epsilon_n = \frac{1}{n^2}$ and $w_n = n$ for all n.

5. Integrability of modulated maximal functions

The Multiplicity Lemma allows us to speed up the index on the martingale in any negative result. Hence, the Multiplicity Lemma and Theorem 4.2 allow us to answer the following type of question in particular instances. Take the Lebesgue derivatives $(D_{m_n} f)$ for some increasing sequence (m_n). Take some increasing function $\Phi:\mathbb{R}^+\to\mathbb{R}^+$. Let $\Phi_1(x)=\Phi(x)1_{[1,\infty)}(x)$ for all $x\in\mathbb{R}^+$. Now, under what conditions will we have $\Phi_1(\sup_{n\geqslant 1}|D_{m_n}f|)$ integrable for all $f\in L_1(\mathbb{R})$ with bounded support? We consider Φ_1 here, instead of Φ, and functions with bounded support, because we want to focus on the large values of the functions in $(D_{m_n} f)$.

First, it is not hard to see the following fact which answers this question negatively if $\Phi(x)=x$ for all $x>0$. This result may be known, but we have not been able to find it in the literature and so provide a proof of it here.

Proposition 5.1. *If (ϵ_n) is non-increasing and converges to 0, then there is a function $f\in L_1(\mathbb{R})$ supported in $[0,1)$ such that $f^* = \sup_{n\geqslant 1}|D(\epsilon_n)f|$ is not integrable.*

Proof. Let (f_s) be a sequence of positive functions in $L_1(\mathbb{R})$, each supported in $[0,1)$, with $\|f_s\|_{L_1(\mathbb{R})}=1$ for all s, which is an approximate identity for $L_1(\mathbb{R})$. That is, assume that the sequence of convolutions $(f_s\star f)$ converges in $L_1(\mathbb{R})$-norm to f for all $f\in L_1(\mathbb{R})$. For example, if for any $\delta>0$, the supports of the functions f_s are eventually in $[0,\delta)$, then the sequence would be an approximate identity.

It is easy to see that $D(\epsilon_n)f_s = f_s\star\frac{1}{\epsilon_n}1_{[-\epsilon_n,0]}$. Because (f_s) is an approximate identity in $L_1(\mathbb{R})$, a standard argument shows that $\sup_{1\leqslant n\leqslant N} D(\epsilon_n)f_s$ converges in $L_1(\mathbb{R})$-norm to $\sup_{1\leqslant n\leqslant N}\frac{1}{\epsilon_n}1_{[-\epsilon_n,0]}$. But $\int_{-\infty}^{\infty}\sup_{1\leqslant n\leqslant N}\frac{1}{\epsilon_n}1_{[-\epsilon_n,0]}(x)\,dx = 1+\sum_{n=1}^{N-1}\frac{1}{\epsilon_n}(\epsilon_n-\epsilon_{n+1})$.

But $\sum_{n=1}^{\infty} \frac{1}{\epsilon_n}(\epsilon_n - \epsilon_{n+1}) = \infty$. Indeed, for any non-decreasing (m_n),

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{1}{\epsilon_n}(\epsilon_n - \epsilon_{n+1}) &= \sum_{n=1}^{\infty} \sum_{k=m_n}^{m_{n+1}-1} \frac{1}{\epsilon_k}(\epsilon_k - \epsilon_{k+1}) \\ &\geqslant \sum_{n=1}^{\infty} \sum_{k=m_n}^{m_{n+1}-1} \frac{1}{\epsilon_{m_n}}(\epsilon_k - \epsilon_{k+1}) \\ &= \sum_{n=1}^{\infty} \frac{1}{\epsilon_{m_n}} \sum_{k=m_n}^{m_{n+1}-1} (\epsilon_k - \epsilon_{k+1}) \\ &= \sum_{n=1}^{\infty} \frac{1}{\epsilon_{m_n}}(\epsilon_{m_n} - \epsilon_{m_{n+1}}).\end{aligned}$$

But then if (m_n) is increasing fast enough, one could have $\frac{\epsilon_{m_{n+1}}}{\epsilon_{m_n}} \leqslant \frac{1}{2}$ for all n, and so $\sum_{n=1}^{\infty} \frac{1}{\epsilon_{m_n}}(\epsilon_{m_n} - \epsilon_{m_{n+1}}) \geqslant \sum_{n=1}^{\infty} \frac{1}{2} = \infty$.

Thus, for all $m, M \geqslant 1$, for sufficiently large s, with $h_m(s) = \frac{1}{2^m} f_s$, we would have

$$\|h_m(s)\|_{L_1(\mathbb{R})} = \frac{1}{2^m} \quad \text{and} \quad \Big\|\sup_{n\geqslant 1} D(\epsilon_n) h_m(s)\Big\|_{L_1(\mathbb{R})} \geqslant M.$$

Therefore, for some rapidly increasing sequence (s_m), the series $h = \sum_{m=1}^{\infty} h_m(s_m)$ will give a positive function $h \in L_1(\mathbb{R})$ that is supported in $[0, 1)$, such that

$$\Big\|\sup_{n\geqslant 1} D(\epsilon_n) h\Big\|_{L_1(\mathbb{R})} = \infty. \qquad \square$$

However, if we modulate $f^* = \sup_{n\geqslant 1} |D(\epsilon_n) f|$, that is consider $\Phi_1(f^*)$ for a suitable increasing function Φ, then the resulting function could be integrable. For example if $\Phi(x) = \sqrt{x}$ for $x \geqslant 0$, then the weak $(1,1)$ inequality for Lebesgue derivatives easily shows that $\Phi_1(f^*)$ is integrable on $\mathbb{R}$ for all $f \in L_1(\mathbb{R})$ with bounded support, and for all (ϵ_n). Indeed, $\Phi_1(f^*)$ would also have bounded support, say on some bounded interval I. Then we have

$$\begin{aligned}\|\Phi_1(f^*)\|_{L_1(\mathbb{R})} &\leqslant m\{0 < \Phi_1(f^*) < 1\} + \sum_{n=1}^{\infty} (n+1)\, m\{n \leqslant \Phi_1(f^*) < n+1\} \\ &\leqslant 2m(I) + \sum_{n=1}^{\infty} n\, m\{n \leqslant \Phi_1(f^*) < n+1\} \\ &= 2m(I) + \sum_{n=1}^{\infty} m\{\Phi_1(f^*) \geqslant n\} \\ &\leqslant 2m(I) + \sum_{n=1}^{\infty} m\{\Phi(f^*) \geqslant n\} \\ &= 2m(I) + \sum_{n=1}^{\infty} m\{f^* \geqslant n^2\} \\ &\leqslant 2m(I) + \sum_{n=1}^{\infty} \frac{1}{n^2} \|f\|_{L_1(\mathbb{R})}.\end{aligned}$$

The above estimate works because $\sum_{n=1}^{\infty} \frac{1}{\Phi^{-1}(n)}$ converges. But what can be said if we were in the situation where $\sum_{n=1}^{\infty} \frac{1}{\Phi^{-1}(n)} = \infty$? For example, take $\Phi(x) = x/(\ln^+ x + 1)$. The results in the previous section on large deviations allow us to answer this again in the negative. We will need to assume for part of this that Φ is scaling, i.e. any constant C, there is another constant K such that $\Phi(Cx) \leqslant K\Phi(x)$ for all sufficiently large x. We have the following result.

Corollary 5.2. *If Φ is increasing, and $w(x) = \Phi^{-1}(x)$ satisfies the conditions of Theorem 4.2, then there is never an increasing sequence (m_n), $m_n \geqslant n$ such that we would have $\Phi_1(\sup_{n\geqslant 1} |D_{m_n} f|)$ integrable for all $f \in L_1(\mathbb{R})$ with bounded support. If Φ is also scaling, then there is no sequence (ϵ_n) decreasing to 0 such that $\Phi_1(\sup_{n\geqslant 1} |D(\epsilon_n) f|)$ is integrable for all $f \in L_1(\mathbb{R})$ with bounded support.*

Proof. We reduce the second part of this result to the first part and then give that proof. Assume there exists a decreasing sequence (ϵ_n) such that $\|\Phi_1(\sup_{n\geqslant 1}|D(\epsilon_n)f|)\|_{L_1(\mathbb{R})}$ is finite for all $f \in L_1(\mathbb{R})$ with bounded support. It follows by the scaling property of Φ that there is an increasing sequence (m_n) with $m_n \geqslant n$ such that $\|\Phi_1(\sup_{n\geqslant 1}|D(\frac{1}{2^{m_n}})f|)\|_{L_1(\mathbb{R})}$ is finite for all $f \in L_1(R)$ with bounded support.

Assume now that Φ is increasing, $w(x) = \Phi^{-1}(x)$ satisfies the conditions of Theorem 4.2, and $(m_n), m_n \geqslant n$ is increasing. Corollary 4.7 shows that there exists a positive function $f \in L_1[0,1)$ such that $\sum_{n=1}^{\infty} m\{D(\frac{1}{2^{m_n}})f \geqslant w_n\} = \infty$. Let $f^* = \sup_{n\geqslant 1}|D(\frac{1}{2^{m_n}})f|$. Then we have f^* supported on $[-1,1]$. We can estimate that

$$\begin{aligned}
\sum_{n=1}^{\infty} m\left\{\left|D\left(\frac{1}{2^{m_n}}\right)f\right| \geqslant w_n\right\} &\leqslant \sum_{n=1}^{\infty} m\left\{\sup_{k\geqslant 1}\left|D\left(\frac{1}{2^{m_k}}\right)f\right| \geqslant w_n\right\} \\
&= \sum_{n=1}^{\infty} m\left\{\sup_{k\geqslant 1}\left|D\left(\frac{1}{2^{m_k}}\right)f\right| \geqslant \Phi^{-1}(n)\right\} \\
&= \sum_{n=1}^{\infty} m\{\Phi(f^*) \geqslant n\} \\
&= \sum_{n=1}^{\infty} m\big(\{\Phi(f^*) \geqslant n\} \cap \{0 < f^* < 1\}\big) + \sum_{n=1}^{\infty} m\big(\{\Phi(f^*) \geqslant n\} \cap \{f^* \geqslant 1\}\big) \\
&\leqslant 2\sharp\{n\colon \Phi^{-1}(n) < 1\} + \sum_{n=1}^{\infty} m\big(\{\Phi(f^*) \geqslant n\} \cap \{f^* \geqslant 1\}\big) \\
&\leqslant 2\sharp\{n\colon n \leqslant \Phi(1)\} + \sum_{n=1}^{\infty} m\{\Phi_1(f^*) \geqslant n\} \\
&\leqslant K_\Phi + \|\Phi_1(f^*)\|_{L_1(\mathbb{R})}
\end{aligned}$$

for some constant K_Φ that depends only on Φ. Then $\|\Phi_1(\sup_{n\geqslant 1}|D(\frac{1}{2^{m_n}})f|)\|_{L_1(\mathbb{R})}$ is infinite for this function $f \in L_1[0,1)$. $\square$

Remark 5.3. In particular, assume Conjecture 4.1 is true. Then there are no non-trivial large deviation inequalities for (m_n) non-decreasing such that $m_n \geqslant n$, and (w_n) non-decreasing. It follows as above that no modulation by an increasing, scaling Φ would give, for (ϵ_n) decreasing to 0, $\Phi_1(\sup_{n\geqslant 1}|D(\epsilon_n)f|)$ always integrable for functions $f \in L_1(\mathbb{R})$ with bounded support, unless $\sum_{n=1}^{\infty}\frac{1}{\Phi^{-1}(n)} < \infty$.

Acknowledgments

The authors would like to thank the referee of this paper. The referee not only read this article very carefully several times, but also contributed the idea of proving the Multiplicity Lemma the way it is now proved, greatly clarifying it and shortening the proof from the one that the authors originally had.

References

[1] K. Hare, A. Stokolos, On weak type inequalities for rare maximal functions, Colloq. Math. 83 (2) (2000) 173–182.
[2] R. Jones, R. Kaufman, J. Rosenblatt, M. Wierdl, Oscillation in ergodic theory, Ergodic Theory Dynam. Systems 18 (1998) 889–935.
[3] J. Rosenblatt, M. Wierdl, A new maximal inequality and its applications, Ergodic Theory Dynam. Systems 12 (1992) 509–558.