

On an inequality of Marcinkiewicz and Zygmund

By J. MOGYORÓDI (Budapest)

1) The following classical inequality of Marcinkiewicz and Zygmund [1] is well-known:

let $X_1, X_2, \dots$ be a sequence of independent random variables with $E(X_i)=0$, $i=1, 2, \dots$ and let $p \geq 1$. Then there exists a constant A such that

$$E\left(\max_{1 \leq k \leq n} \left|\sum_{i=1}^k X_i\right|^p\right) \leq AE\left(\left|\sum_{i=1}^n X_i\right|^p\right).$$

The constant A is absolute in the sense that it does not depend on p nor on n ($n=1, 2, \dots$).

The idea of the proof is the following: one gives the inequality

$$E\left(\max_{1 \leq k \leq n} \left|\sum_{i=1}^k X_i\right|^p\right) \leq C_p E\left(\left|\sum_{i=1}^n X_i\right|^p\right),$$

(where C_p is a constant depending only on p) in two ways: in the first one we obtain

$$C_p = \left(\frac{p}{p-1}\right)^p \quad (p > 1),$$

while in the second

$$C_p = 2^{p+2}, \quad (p \geq 1).$$

The first constant is non-bounded in the neighbourhood of $p=1$, while the second in the neighborhood of $p=+\infty$. Thus the solution $A>1$ of the equation

$$\left(\frac{p}{p-1}\right)^p = 2^{p+2}$$

is convenient for every $p \geq 1$.

Let $\Phi(x)=x^p$, ($x \geq 0$) which, for $p \geq 1$, is a convex function and put $S_k=X_1+\dots+X_k$; $k=1, 2, \dots$. Then the preceding inequality can be written in the following form:

$$E\left(\max_{1 \leq k \leq n} \Phi(|S_k|)\right) \leq AE(\Phi(|S_n|)).$$

The aim of the present note is to generalize this inequality to a larger class of convex functions $\Phi(x)$ with an absolute constant A which is convenient for each member of this class.

2) Let $\Phi(x)$ be a Young function, i.e. let

$$(1) \quad \Phi(x) = \int_0^x \varphi(t) dt, \quad x \geq 0,$$

where $\varphi(0)=0$, $\lim_{t \rightarrow +\infty} \varphi(t) = +\infty$, $\varphi(t)$ is non-decreasing and continuous from the right. A Young function $\Phi(x)$ is convex. Let us suppose further that the following growth condition is satisfied: for $x > 0$

$$(2) \quad \Phi(2x) \leq c\Phi(x),$$

where c is a positive constant. Putting

$$p = \sup_{x>0} \frac{x\varphi(x)}{\Phi(x)}$$

one easily finds that $1 < p \leq c-1$, further that for every $q \geq 1$ and for $x \geq 0$ we have

$$(3) \quad \Phi(qx) \leq q^p \Phi(x),$$

particularly

$$\Phi(2x) \leq 2^p \Phi(x).$$

Further let us consider the Young function $\Psi(x)$, which is the conjugate of $\Phi(x)$ in the sense of Young, i.e. let $\psi(0)=0$

$$\psi(x) = \sup \{t: \varphi(t) \leq x\}, \quad x > 0,$$

and

$$\Psi(x) = \int_0^x \psi(t) dt, \quad x \geq 0.$$

$\psi(x)$ is the inverse of $\varphi(x)$ and we have $\varphi(\psi(x)) \leq x$, $\psi(\varphi(x)) \leq x$.

It is easily seen that the following inequality holds: for every $u \geq 0$, $v \geq 0$

$$uv \leq \Phi(u) + \Psi(v).$$

This inequality is due to Young. In the particular cases $v = \varphi(u)$ and $u = \psi(v)$ we have the equalities

$$u\varphi(u) = \Phi(u) + \Psi(\varphi(u)),$$

$$v\psi(v) = \Phi(\psi(v)) + \Psi(v).$$

In some cases in this paper we also suppose that for $x > 0$ the growth condition

$$(4) \quad \Psi(2x) \leq c'\Psi(x)$$

holds, where c' is a positive constant. Putting

$$q = \sup_{x>0} \frac{x\psi(x)}{\Psi(x)}$$

we have as above $1 < q \leq c'-1$ and also the analogue of inequality (3).

The growth condition (2) (resp. (4)) implies that $\Phi(x)=0$ identically, or $\Phi(x)>0$, if $x>0$ (resp. $\Psi(x)=0$ identically, or $\Psi(x)>0$, if $x>0$). It follows that $\varphi(t)>0$ for $t>0$ (resp. $\psi(t)>0$, if $t>0$).

For the proof of these we refer to [2], [3] and [5].

We have

$$q = \sup_{x>0} \frac{x\psi(x)}{\Psi(x)} \cong \sup_{x>0} \frac{\varphi(x)x}{\Psi(\varphi(x))},$$

since $\psi(\varphi(x)) \cong x$ and the set of the values of $\varphi(x)$, when x varies over $(0, +\infty)$, is contained in the set $\{x: x>0\}$. Further, from the above inequality of Young it follows that

$$q \cong \sup_{x>0} \frac{x\varphi(x)}{x\varphi(x) - \Phi(x)} \cong \sup_{x>0} \frac{x\varphi(x)/\Phi(x)}{x\varphi(x)/\Phi(x) - 1} \cong \sup_{x>0} \frac{x\varphi(x)/\Phi(x)}{p-1} = \frac{p}{p-1}.$$

Consequently, we have

$$(5) \quad \frac{1}{p} + \frac{1}{q} \cong 1.$$

It follows that $q \uparrow +\infty$ as $p \downarrow 1$ (resp. $p \uparrow +\infty$ as $q \downarrow 1$).

3) GARSIA [3] proved the following inequality: let $\Phi(x)$ be a Young function and suppose that the growth condition (2) is satisfied. Let further (f_n, F_n) be a non-negative submartingale and put $f_n^* = \max_{1 \leq k \leq n} f_k$.

Then

$$E(\Psi(f_n^*)) \cong pE(\Psi(pf_n)),$$

where Ψ is the conjugate of Φ in the sense of Young.

We will derive a consequence of this inequality. First we reformulate Garsia's inequality in a more precise form. Namely, we can drop the factor p staying before the expectation on the right hand side of the preceding inequality. We have thus

Lemma 1. *If $\Phi(x)$ and $\Psi(x)$ are conjugate Young functions and Φ satisfies the growth condition (2) with*

$$1 < p = \sup_{x>0} \frac{x\varphi(x)}{\Phi(x)} < +\infty,$$

further it (f_n, F_n) is a non-negative submartingale, then

$$E(\Psi(f_n^*)) \cong E(\Psi(pf_n)).$$

Moreover, if $\Phi(x)$ and $\Psi(x)$ satisfy the growth conditions (2) and (4) respectively, with

$$1 < q = \sup_{x>0} \frac{x\psi(x)}{\Psi(x)} < +\infty,$$

then

$$E(\Phi(f_n^*)) \cong q^p E(\Phi(f_n)).$$

PROOF. Let $a > 0$ be arbitrary and take

$$g_n = \min(f_n, a), \quad n = 1, 2, \dots$$

It is easy to see that Doob's classical inequality

$$\lambda E(\chi(f_n^* > \lambda)) \leq E(f_n \chi(f_n^* > \lambda)),$$

where $\chi(A)$ denotes the indicator of the event A and $\lambda > 0$ is arbitrary, remains valid if we substitute f_n^* by $g_n^* = \max_{1 \leq i \leq n} g_i$, i.e. we have

$$\lambda E(\chi(g_n^* > \lambda)) \leq E(f_n \chi(g_n^* > \lambda)).$$

Let us integrate this inequality on $(0, +\infty)$ with respect to the positive and σ -finite measure generated by $\psi(\lambda)$ and use the Fubini theorem. We obtain

$$E\left(\int_0^{g_n^*} \lambda d\psi(\lambda)\right) \leq E\left(f_n \int_0^{g_n^*} d\psi(\lambda)\right).$$

Recalling that

$$\int_0^x \lambda d\psi(\lambda) = x\psi(x) - \Psi(x) = \Phi(\psi(x)),$$

we have

$$E(\Phi(\psi(g_n^*))) \leq E(f_n \psi(g_n^*)) = \frac{1}{p} E(pf_n \psi(g_n^*)).$$

The right hand side will be majorated by the aid of the inequality of Young:

$$v\psi(t) \leq \Phi(\psi(t)) + \Phi(v).$$

This gives

$$E(\Phi(\psi(g_n^*))) \leq \frac{1}{p} E(\Phi(\psi(g_n^*))) + \frac{1}{p} E(\Psi(pf_n)).$$

Since $E(\Phi(\psi(g_n^*))) \leq \Phi(\psi(a))$, we obtain

$$\left(1 - \frac{1}{p}\right) E(\Phi(\psi(g_n^*))) \leq \frac{1}{p} E(\Psi(pf_n)).$$

Notice also that by the growth condition (2)

$$p = \sup_{x>0} \frac{x\varphi(x)}{\Phi(x)} \geq \sup_{x\geq 0} \frac{\psi(x)x}{\Phi(\psi(x))},$$

and thus from

$$\Phi(\psi(x)) = x\psi(x) - \Psi(x)$$

we have

$$1 = \frac{x\psi(x)}{\Phi(\psi(x))} - \frac{\Psi(x)}{\Phi(\psi(x))} \leq p - \frac{\Psi(x)}{\Phi(\psi(x))}.$$

This implies that

$$(p-1)\Phi(\psi(x)) \geq \Psi(x).$$

Comparing this with the preceding inequality one has

$$\frac{p-1}{p} \frac{1}{p-1} E(\Psi(g_n^*)) \leq \frac{1}{p} E(\Psi(pf_n)),$$

or, in other words

$$E(\Psi(g_n^*)) \leq E(\Psi(pf_n)).$$

If $\alpha \uparrow +\infty$ then $g_n^* \uparrow f_n^*$ and thus by the Beppo Levi theorem

$$E(\Psi(f_n^*)) \leq E(\Psi(pf_n)).$$

The first part of our assertion is thus proved.

If both $\Phi(x)$ and $\Psi(x)$ satisfy the growth condition (2) and (4) respectively, then the role of these is symmetric. Thus on the basis of the preceding inequality we can write

$$E(\Phi(f_n^*)) \leq E(\Phi(qf_n)).$$

Since $q > 1$ we have $\Phi(qf_n) \leq q^p \Phi(f_n)$. This and the preceding inequality give the second part of the assertion.

REMARK. Our inequality is more precise than Garsia's one. The second part of the assertion is a direct generalization of the classical inequality of DOOB; if $p > 1$ and $p^{-1} + q^{-1} = 1$ then

$$E(f_n^{*p}) \leq q^p E(f_n^p).$$

4) BICKEL proved the following inequality ([6]): let $X_1, X_2, \dots, X_n$ be independent and symmetrically distributed random variables. Let further $g(x)$ be a convex function. Then

$$(6) \quad P\left(\max_{1 \leq k \leq n} g(S_k) \geq \varepsilon\right) \leq 2P(g(S_n) \geq \varepsilon), \quad \varepsilon > 0,$$

where $S_k = X_1 + \dots + X_k$, $k = 1, 2, \dots, n$.

This inequality is a generalization of the Lévy inequality.

We now prove the following

Lemma 2. Let $X_1, X_2, \dots, X_n$ be independent random variables with $E(X_i) = 0$, $i = 1, 2, \dots, n$. Let further $\Phi(x)$ be a Young function satisfying the growth condition

$$\Phi(2x) \leq c\Phi(x)$$

for every $x \geq 0$ with a constant c . Put

$$p = \sup_{x > 0} \frac{x\phi(x)}{\Phi(x)}.$$

Then

$$E\left(\max_{1 \leq k \leq n} \Phi(|S_k|)\right) \leq 2^{p+2} E(\Phi(|S_n|)).$$

REMARK. BICKEL [6] proved the same inequality but his constant is $4c$. For our purposes we need the constant 2^{p+2} , which facilitates the comparison of this to the constant of the inequality of Lemma 1. Notice that nothing is supposed about the behaviour of the conjugate function $\Psi(x)$.

PROOF. Suppose first that $X_1, X_2, \dots, X_n$ are symmetrically distributed. Let us consider the function $\Phi_1(x)$ defined for $x \in R$ such that

$$\Phi_1(x) = \Phi_1(-x)$$

and for $x \geq 0$ we put $\Phi_1(x) = \Phi(x)$. Then $\Phi_1(x)$ is also convex and by (6) we have

$$E\left(\max_{1 \leq k \leq n} \Phi_1(S_k)\right) \leq 2E(\Phi_1(S_n)).$$

Since $\Phi_1(S_k) = \Phi(|S_k|)$ we obtain that

$$E\left(\max_{1 \leq k \leq n} \Phi(|S_k|)\right) \leq 2E(\Phi(|S_n|)).$$

In the general case we proceed as follows: let $X'_1, X'_2, \dots, X'_n$ be random variables such that $X_1, \dots, X_n, X'_1, \dots, X'_n$ are independent, further for $i=1, 2, \dots, n$ the variables X_i and X'_i have the same distribution. Put $Z_i = X_i - X'_i$, $S'_i = X'_1 + \dots + X'_i$, $i=1, \dots, n$. Then the random variables Z_i are independent and symmetrically distributed. Thus by the preceding remark

$$E\left(\max_{1 \leq k \leq n} \Phi(|S_k - S'_k|)\right) \leq 2E(\Phi(|S_n - S'_n|)).$$

But by (3)

$$\begin{aligned} \Phi(|S_n - S'_n|) &\leq \Phi(|S_n| + |S'_n|) \leq \Phi(2 \max(|S_n|, |S'_n|)) \leq \\ &\leq 2^p \Phi(\max(|S_n|, |S'_n|)) \leq 2^p (\Phi(|S_n|) + \Phi(|S'_n|)). \end{aligned}$$

Consequently,

$$(7) \quad E\left(\max_{1 \leq k \leq n} (|S_k - S'_k|)\right) \leq 2^{p+2} E(\Phi(|S_n|)).$$

On the other hand, if F_n denotes the smallest σ -field generated by $X_1, X_2, \dots, X_n$, we obtain by the submartingale property of the partial sums $|S_k - S'_k|$ that

$$E(\Phi(|S_k - S'_k|) | F_n) \geq \Phi(|S_k|),$$

from which

$$\max_{1 \leq k < n} E(\Phi(|S_k - S'_k|) | F_n) \geq \max_{1 \leq k \leq n} \Phi(|S_k|),$$

or, in other words

$$E\left(\max_{1 \leq k \leq n} \Phi(|S_k - S'_k|) | F_n\right) \geq \max_{1 \leq k \leq n} \Phi(|S_k|)$$

with probability 1. It follows that

$$(8) \quad E\left(\max_{1 \leq k \leq n} \Phi(|S_k - S'_k|)\right) \geq E\left(\max_{1 \leq k \leq n} \Phi(|S_k|)\right).$$

(7) and (8) together give the desired result.

5) We are now in the position to formulate the generalization of the Marcinkiewicz—Zygmund inequality.

We say that the Young function Φ belongs to the class C if $\Phi(x)$ and its conjugate $\Psi(x)$ satisfy the growth condition (2) and (4) respectively and the corresponding quantities p and q are such that if $p \geq p_0 > 1$ then $q^p \leq k$, where k is a positive constant.

Theorem. Let $X_1, X_2, \dots$ be independent random variables with $E(X_i) = 0$, $i=1, 2, \dots$. Then for every $\Phi \in C$ there exists an absolute constant A such that

$$E\left(\max_{1 \leq k \leq n} \Phi(|S_k|)\right) \leq AE(\Phi(|S_n|)).$$

PROOF. By Lemma 1. we have

$$E\left(\max_{1 \leq k \leq n} \Phi(|S_k|)\right) \leq q^p E(\Phi(|S_n|)).$$

By inequality (5) we see that $q^p \uparrow +\infty$ as $p \downarrow 1$.

By Lemma 2 we obtain

$$* \quad E\left(\max_{1 \leq k \leq n} \Phi(|S_k|)\right) \leq 2^{p+2} E(\Phi(|S_n|)).$$

Here $2^{p+2} \downarrow 8$ as $p \downarrow 1$ and $2^{p+2} \uparrow +\infty$ as $p \uparrow +\infty$.

Since by our supposition we have for every $p \geq p_0 > 1$ that $q^p \leq k$ and since for $1 < p \leq p_0$ the function 2^{p+2} is bounded and increasing we see that $A = \max(2^{p_0+2}, k)$ is convenient for every $p > 1$. This proves the assertion.

6) The class C considered in the Theorem obviously contains the functions $\Phi(x) = x^\alpha/\alpha$, where $\alpha > 1$. In fact, in this case $\Psi(x) = x^\beta/\beta$, where $\alpha^{-1} + \beta^{-1} = 1$. Now $p = \alpha$ and $q = \beta$; it follows that $q^p = \beta^\alpha \rightarrow e$ as $\alpha \rightarrow +\infty$.

To show that C is larger than this special class, let us consider the following function investigated by KRASNOSELSKII and RUTICKII. Let $R(x)$ be a non-negative and non-decreasing function, which is differentiable and has continuous derivative $r(x)$. Suppose that

$$\inf_{u>0} \frac{ur(u)}{R(u)} = 0, \quad \sup_{u>0} \frac{ur(u)}{R(u)} = \gamma < +\infty.$$

Assume further that for $\alpha > 1$ the function

$$\Phi(u) = \frac{u^\alpha}{\alpha} R(u)$$

is a Young function, i.e.

$$\varphi(u) = \frac{d}{du} \left(\frac{u^\alpha}{\alpha} R(u) \right) = u^{\alpha-1} R(u) \left[1 + \frac{ur(u)}{\alpha R(u)} \right]$$

is strictly increasing and $\varphi(u) \rightarrow +\infty$ as $u \rightarrow +\infty$.

It follows easily that

$$p = \sup_{u>0} \frac{u\varphi(u)}{\Phi(u)} = \alpha + \gamma$$

is finite. Further we have

$$\Psi(\varphi(u)) = u\varphi(u) - \Phi(u) = u^\alpha R(u) \left[1 + \frac{ur(u)}{\alpha R(u)} - \frac{1}{\alpha} \right],$$

thus

$$\begin{aligned} q &= \sup_{u>0} \frac{u\psi(u)}{\Psi(u)} = \sup_{\varphi(u)>0} \frac{u\varphi(u)}{\Psi(\varphi(u))} = \sup_{u>0} \frac{u\varphi(u)}{\Psi(\varphi(u))} = \\ &= \sup_{u>0} \frac{1 + \frac{ur(u)}{\alpha R(u)}}{1 + \frac{ur(u)}{\alpha R(u)} - \frac{1}{\alpha}} = 1 + \frac{1}{\alpha - 1 + \inf_{u>0} \frac{ur(u)}{R(u)}} = 1 + \frac{1}{\alpha - 1} = \frac{\alpha}{\alpha - 1}. \end{aligned}$$

From this it follows that

$$q^p = \left(\frac{\alpha}{\alpha-1} \right)^{\alpha+\gamma} \rightarrow e,$$

if $\alpha \rightarrow +\infty$. Thus the functions $\frac{u^\alpha}{\alpha} R(u)$ belong to the class C .

As an example, let $R(u) = (\log(1+u))^\gamma$, where $\gamma \geq 1$. In this case

$$r(u) = \frac{\gamma(\log(1+u))^{\gamma-1}}{1+u}$$

and

$$\sup_{u>0} \frac{ur(u)}{R(u)} = \gamma \sup_{u>0} \frac{u}{(1+u)\log(1+u)} = \gamma; \quad \inf_{u>0} \frac{ur(u)}{R(u)} = 0.$$

Further,

$$\varphi(u) = u^{\alpha-1}(\log(1+u))^\gamma + \frac{u^\alpha}{\alpha} \frac{(\log(1+u))^{\gamma-1}}{1+u}.$$

For $\gamma \geq 1$ the functions $(\log(1+u))^{\gamma-1}$ and $u^\alpha/(u+1)$ are increasing and continuous. Thus $\varphi(u)$ strictly increases, it is continuous and $\lim_{u \rightarrow +\infty} \varphi(u) = +\infty$. Consequently,

$$\Phi(u) = \frac{u^\alpha}{\alpha} (\log(1+u))^\gamma; \quad \gamma \geq 1, \alpha > 1$$

is a Young function.

Let $X_1, X_2, \dots$ be a sequence of independent random variables with zero mean value and put $S_k = X_1 + \dots + X_k$. Then on the basis of the preceding example and the theorem we have for $\gamma \geq 1$ and $\alpha > 1$

$$E\left(\max_{1 \leq k \leq n} |S_k|^\alpha (\log(1+|S_k|))^\gamma\right) \leq AE(|S_n|^\alpha (\log(1+|S_n|))^\gamma),$$

where A is an absolute constant.

References

- [1] J. MARCINKIEWICZ and A. ZYGMUND, Quelques théorèmes sur les fonctions indépendantes. *Studia Mathematica*. **7** (1938), 104—120.
- [2] M. A. KRASNOSELSKII and YA. B. RUTICKII, Convex functions and Orlicz spaces. (Translated by L. F. BORON) Noordhoff, Groningen, 1964.
- [3] A. M. GARSIA, On a convex function inequality for martingales. *Annals of Probability*. **1** (1973) 171—174.
- [4] A. M. GARSIA, Martingale inequalities. *Seminar Notes on Recent Progress*. Benjamin, Reading, 1973.
- [5] J. NEVEU, Martingales à temps discret. Masson, Paris, 1972.
- [6] P. J. BICKEL, A Hájek—Rényi extension of Lévy's inequality and some applications. *Acta Math Acad. Sci. Hungaricae*. **21** (1970) 199—206.

(Received October 6, 1976.)