On a Sobolev type inequality and its applications

by

WITOLD BEDNORZ (Warszawa)

Abstract. Assume $\|\cdot\|$ is a norm on $\mathbb{R}^n$ and $\|\cdot\|_*$ its dual. Consider the closed ball $T := B_{\|\cdot\|}(0, r)$, $r > 0$. Suppose φ is an Orlicz function and ψ its conjugate. We prove that for arbitrary $A, B > 0$ and for each Lipschitz function f on T,

$$
\sup_{s,t \in T} |f(s) - f(t)| \le 6AB\left(\int_0^r \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon \right.
$$
$$
\left. + \frac{1}{n|B_{\|\cdot\|}(0,1)|} \int_T \varphi\left(\frac{1}{B}\,\|\nabla f(u)\|_*\right) du\right),
$$

where $|\cdot|$ is the Lebesgue measure on $\mathbb{R}^n$. This is a strengthening of the Sobolev inequality obtained by M. Talagrand. We use this inequality to state, for a given concave, strictly increasing function $\eta : \mathbb{R}_+ \to \mathbb{R}$ with $\eta(0) = 0$, a necessary and sufficient condition on φ so that each separable process $X(t)$, $t \in T$, which satisfies

$$
\|X(s) - X(t)\|_\varphi \le \eta(\|s - t\|) \quad \text{for } s, t \in T
$$

is a.s. sample bounded.

1. Introduction. Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ and $B_{\|\cdot\|}(x, r) := \{y \in \mathbb{R}^n : \|x - y\| \le r\}$. In particular, we will consider the closed ball $T := B_{\|\cdot\|}(0, r)$, $r > 0$.

Let $\langle \cdot, \cdot \rangle$ be the canonical scalar product (that is, $\langle u, v \rangle := \sum_{i=1}^n u_i v_i$ for $u, v \in \mathbb{R}^n$) and $\|\cdot\|_*$ the dual norm, i.e.

$$
\|v\|_* := \sup_{u \in B_{\|\cdot\|}(0,1)} |\langle u, v \rangle| \quad \text{for } v \in \mathbb{R}^n.
$$

We say that $\varphi : \mathbb{R}_+ \to \mathbb{R}$ is an *Orlicz function* if it is convex, strictly increasing, $\varphi(0) = 0$ and $\lim_{x \to 0} \varphi(x)/x = 0$, $\lim_{x \to \infty} \varphi(x)/x = \infty$. For each Orlicz function φ we define its *conjugate*

$$
\psi(x) := \sup_{y \ge 0}(xy - \varphi(y)) \quad \text{for } x \ge 0.
$$

It is also an Orlicz function. Moreover, it is well known that φ is the conjugate function for ψ, that is, $\varphi(x) = \sup_{y \geq 0}(xy - \psi(y))$. The definition implies the Young inequality

$$(1) \qquad\qquad xy \leq \varphi(x) + \psi(y) \quad \text{ for } x, y \geq 0.$$

From now on we assume that φ, ψ are conjugate Orlicz functions.

We will prove the following Sobolev type inequality and give its applications to the theory of stochastic processes.

THEOREM 1. *For any $A, B > 0$ and for each Lipschitz function f on T,*

$$\sup_{s,t \in T} |f(s) - f(t)| \leq 6AB \bigg(\int_0^r \psi \bigg(\frac{1}{A\varepsilon^{n-1}} \bigg) \varepsilon^{n-1} \, d\varepsilon$$
$$+ \frac{1}{n|B_{\|\cdot\|}(0,1)|} \int_T \varphi \bigg(\frac{1}{B} \, \|\nabla f(u)\|_* \bigg) \, du \bigg),$$

where $|\cdot|$ is the Lebesgue measure on $\mathbb{R}^n$.

Talagrand obtained this inequality in the proof of Theorem 5.1 of [9] for the norm $\|s\| = \sup_{i=1}^n |s_i|$.

Since for each $s, t \in \mathbb{R}^n$ we have $s \in B_{\|\cdot\|}(t, \|s - t\|)$, the above theorem implies some regularity of f. Namely, for $\varphi(x) \equiv x^p/p$, $p > n$, we obtain the following classical result (which can be deduced from Lemma 7.16 in [7] by using the Hölder inequality).

COROLLARY 1. *Suppose $p > n$. Then for each Lipschitz function f on $\mathbb{R}^n$,*

$$\sup_{s,t \in \mathbb{R}^n} \frac{|f(s) - f(t)|}{\|s - t\|^{1-n/p}} \leq \frac{6 \big(\frac{p-1}{p-n} \big)^{1-1/p}}{(n|B_{\|\cdot\|}(0,1)|)^{1/p}} \bigg(\int_{\mathbb{R}^n} \|\nabla f(u)\|_*^p \, du \bigg)^{1/p}.$$

Proof. The conjugate function for $\varphi(x) \equiv x^p/p$ is $\psi(x) \equiv x^q/q$, where $1/p + 1/q = 1$. By Theorem 1 for any $A, B > 0$ we have

$$|f(s) - f(t)| \leq 6AB \bigg(\int_0^{\|s-t\|} q^{-1} \bigg(\frac{1}{A\varepsilon^{n-1}} \bigg)^q \varepsilon^{n-1} \, d\varepsilon$$
$$+ \frac{1}{n|B_{\|\cdot\|}(0,1)|} \int_{B_{\|\cdot\|}(t,\|s-t\|)} p^{-1} \bigg(\frac{1}{B} \, \|\nabla f(u)\|_* \bigg)^p \, du \bigg).$$

We take

$$A = \bigg(\int_0^{\|s-t\|} \varepsilon^{(1-q)(n-1)} \, d\varepsilon \bigg)^{1/q}, \quad B = \bigg(\frac{1}{n|B_{\|\cdot\|}(0,1)|} \int_{\mathbb{R}^n} (\|\nabla f(u)\|_*)^p \, du \bigg)^{1/p}.$$

Consequently $(T = B_{\|\cdot\|}(0, r))$,

$$|f(s) - f(t)| \leq 6AB = \frac{6\left(\frac{p-1}{p-n}\right)^{1/q}\|s - t\|^{1-n/p}}{(n|B_{\|\cdot\|}(0,1)|)^{1/p}}\left(\int\limits_T \|\nabla f(u)\|_*^p \, du\right)^{1/p}. \ \blacksquare$$

Another immediate consequence of Theorem 1 is a sufficient condition for embedding of the Sobolev space $W_0^{1,\varphi}(T)$ into $L_\infty(T)$.

COROLLARY 2. *If*

$$\int\limits_0^r \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1} \, d\varepsilon < \infty \quad \text{for some } A > 0$$

then $W_0^{1,\varphi}(T)$ *embeds into* $L_\infty(T)$.

This can also be deduced from part II of Theorem 1.1 in the paper by A. Cianchi [3].

To give applications to stochastic processes we need some definitions. Let (K, d) be a compact metric space. Denote by $\mathfrak{B}(K)$ the space of bounded Borel functions on K, by $C(K)$ the Banach space of continuous functions on K with sup-norm, and by $\mathrm{Lip}(K)$ the space of Lipschitz functions on T with the norm

$$\|f\|_{\mathrm{Lip}} := \sup_{s \neq t} \frac{|f(s) - f(t)|}{d(s, t)} + D(K)^{-1}\|f\|_\infty,$$

where $D(K) := \sup\{d(s, t) : s, t \in K\}$ (the diameter of K). Let $\mathcal{P}(K)$ be the set of all Borel probability measures on K. For each $\nu \in \mathcal{P}(K)$, $f \in \mathfrak{B}(K)$ and $A \in \mathcal{B}(K)$ (with $\nu(A) > 0$) we write

$$\oint\limits_A f(u) \, \nu(du) := \frac{1}{\nu(A)} \int\limits_A f(u) \, \nu(du).$$

Suppose X is a random variable. We define the Luxemburg norm $\|X\|_\varphi := \inf\{c > 0 : \mathbf{E}\varphi(|X|/c) \leq 1\}$. For a fixed probability space the Banach space L_φ consists of all random variables for which $\|X\|_\varphi < \infty$.

In this paper we consider only separable processes (for the definition see the introduction in the book by Ledoux–Talagrand [8]). For each separable $X(t)$, $t \in K$, we have

$$(2) \qquad \mathbf{E}\sup_{s,t \in K}|X(s) - X(t)| = \sup_{F \subset K}\mathbf{E}\sup_{s,t \in F}|X(s) - X(t)|,$$

where the first supremum on the right-hand side is taken over all finite subsets of K. Let us impose the Lipschitz condition on increments of our processes, that is, for each $X(t)$, $t \in K$, we assume that

$$(3) \qquad \sup_{s,t \in K}\mathbf{E}\varphi\left(\frac{|X(s) - X(t)|}{d(s, t)}\right) \leq 1.$$

This condition can be rewritten in terms of Luxemburg norms as

$$\|X(s) - X(t)\|_\varphi \leq d(s,t) \quad \text{for } s,t \in K.$$

In the theory of stochastic processes a lot of effort has been put into finding criteria for boundedness or continuity of stochastic processes. In most cases they are of the following Kolmogorov type: some assumptions on the Orlicz function φ and the metric space (K,d) ensure that for each separable process X on K the condition (3) implies that X is bounded a.s.

It is not difficult to prove that under the above assumptions on φ and K the following two conditions are equivalent:

1. Each separable process $X(t)$, $t \in K$, which satisfies (3) is a.s. bounded.
2. There exists a universal constant $S < \infty$ such that for each process X the condition (3) implies

$$(4) \qquad \mathbf{E} \sup_{s,t \in K} |X(s) - X(t)| \leq S.$$

The minimal constant S is denoted by $S(K,d,\varphi)$. For a proof of the above statement we refer to M. Talagrand [9, Theorem 2.3].

REMARK 1. In terms of absolutely summing operators each of the above conditions is equivalent to the fact that the injection operator $J : \mathrm{Lip}(K) \to C(K)$ is $(\varphi, 1)$ absolutely summing, in the sense of P. Assouad [1].

By far the strongest criteria for finiteness of $S(K,d,\varphi)$ were obtained using the concept of majorizing measures which was introduced by X. Fernique in the early 70s. It served him and M. Talagrand to characterize bounded Gaussian processes. To explain the concept briefly, we introduce the following definitions.

For $t \in K$ and $\varepsilon \geq 0$, we set

$$B(t,\varepsilon) := \{s \in K : d(s,t) \leq \varepsilon\}, \quad S(t,\varepsilon) := \{s \in K : d(s,t) = \varepsilon\}.$$

We say that $m \in \mathcal{P}(K)$ is a *majorizing measure* (with respect to φ and d) if

$$\mathcal{M}(m,\varphi) := \sup_{t \in K} \int_0^{D(K)} \varphi^{-1}\left(\frac{1}{m(B(t,\varepsilon))}\right) d\varepsilon < \infty.$$

X. Fernique [5], [6] proved that if φ has exponential growth then the existence of a majorizing measure is necessary and sufficient for $S(K,d,\varphi)$ to be finite. Generalizing results of Fernique, Talagrand and others, the author [2] succeeded in proving that for each Orlicz function φ the existence of a majorizing measure always implies $S(K,d,\varphi) < \infty$. However, as will be seen in the next sections, the existence of such a measure is not always necessary for finiteness of $S(K,d,\varphi)$. So it is still an open problem to characterize (K,d) and φ for which all processes satisfying (3) are a.s. sample bounded.

This problem was studied in depth by M. Talagrand [9]. He managed to find such a characterization of φ, (K, d) in two particular cases, important for applications: when d is the Euclidean distance on $\mathbb{R}^n$ and K is a ball in $\mathbb{R}^n$, and when $K = [-1, 1]$ and $d(x, y) = \eta(|x - y|)$ where η is a concave, strictly increasing function with $\eta(0) = 0$. Generalizing his ideas and using Theorem 1 we find a characterization in the case when $K = T = B_{\|\cdot\|}(0, r)$ and $d(x, y) = \eta(\|x - y\| \|)$ (η is concave, strictly increasing, with $\eta(0) = 0$).

By the definition $B(t, \varepsilon) = B_{\|\cdot\|}(t, \eta^{-1}(\varepsilon)) \cap T$. Let λ be the normalized Lebesgue measure on T, that is, $\lambda(A) = |A|/|T|$ for each $A \in \mathcal{B}(T)$, where $|\cdot|$ is the standard Lebesgue measure on $\mathbb{R}^n$. Note that

$$\lambda(B(t, \varepsilon)) \leq \eta^{-1}(\varepsilon)^n/r^n$$

with equality if $B_{\|\cdot\|}(t, \eta^{-1}(\varepsilon)) \subset T$.

The function $\eta(y)/y$ is positive and decreasing. Assume that $\eta'(0) = \infty$ (the case of finite derivative will be considered later). Following M. Talagrand [9, Theorem 5.2] we introduce a sequence $(r_k)_{k \geq 0}$ by setting $r_0 = \eta(r)$, and for $k \geq 0$,

$$r_{k+1} := \inf\left\{\varepsilon \geq 0 : r_k \leq 2\varepsilon \text{ or } \frac{\varepsilon}{\eta^{-1}(\varepsilon)} \leq 2\frac{r_k}{\eta^{-1}(r_k)}\right\}.$$

The sequence $(r_k)_{k \geq 0}$ decreases to 0, since $r_{k+1} \leq r_k/2$. The assumption $\eta'(0) = \infty$ guarantees that $r_k > 0$. There are two possibilities:

$$r_k = 2r_{k+1} \quad \text{or} \quad \frac{r_{k+1}}{\eta^{-1}(r_{k+1})} = 2\frac{r_k}{\eta^{-1}(r_k)}.$$

Denote by I the set of $k \geq 0$ for which the first possibility holds, and the rest by J. Notice that necessarily

$$(5) \qquad\qquad 2r_{k+1} \leq r_k, \quad 2\frac{r_k}{\eta^{-1}(r_k)} \leq \frac{r_{k+1}}{\eta^{-1}(r_{k+1})}.$$

For $k \geq 0$ we define $\mathcal{S}_k$ to be the infimum of $c > 0$ such that

$$\int_{r_{k+1}}^{r_k} \frac{\lambda(B(0, \varepsilon))}{\varepsilon} \psi\left(\frac{\varepsilon}{c\lambda(B(0, \varepsilon))}\right) d\varepsilon$$

$$= \int_{r_{k+1}}^{r_k} \frac{\eta^{-1}(\varepsilon)^n}{r^n \varepsilon} \psi\left(\frac{r^n \varepsilon}{c\eta^{-1}(\varepsilon)^n}\right) d\varepsilon \leq 1.$$

Observe that if $\eta'(0) < \infty$, then there exists $m \geq 0$ such that $r_m > 0$ and $r_{m+1} = 0$. Consequently, in this case $\mathcal{S}_k = 0$ for $k > m$ and $\mathcal{S}_m$ can be either finite or infinite.

Let us state the main result of the paper.

THEOREM 2. *The following inequality holds*:

$$(6) \qquad K^{-1}\sum_{k\geq 0}\mathcal{S}_k \leq S(T,d,\varphi) \leq K\sum_{k\geq 0}\mathcal{S}_k,$$

where the constant K depends only on n.

In fact, we show that $\sum_{k\geq 0}\mathcal{S}_k \leq 3(n+2)S(T,d,\varphi)$ and $S(T,d,\varphi) \leq (a+bn^2)\sum_{k\geq 0}\mathcal{S}_k$ for $k\geq 0$, where a,b are universal constants.

COROLLARY 3. *Let $\varphi(x) \equiv x^p/p$, $p > 1$ and $\eta(x) = x^\alpha$, $0 < \alpha \leq 1$. Then $S(T,d,\varphi) < \infty$ if and only if $n < p\alpha$.*

In the case of $T = B_{\|\cdot\|}(0,1)$ and $d(s,t) = \|s-t\|$ (that is, $\eta(x) \equiv x$) we have $m = 0$ in the above construction. Thus $S(T,d,\varphi)$ is comparable with $\mathcal{S}_0$, where $\mathcal{S}_0$ is such that

$$\int\limits_0^1 \varepsilon^{n-1}\psi\left(\frac{1}{\mathcal{S}_0\varepsilon^{n-1}}\right) d\varepsilon = 1,$$

up to a constant which depends only on n. When d is the Euclidean distance this corollary was proved by M. Talagrand ([9, Theorem 5.1]).

In the case of $T = [-1,1]$, $\eta'(0) = \infty$ it can be observed that for some universal $C > 0$,

$$C^{-1}r_k\varphi^{-1}\left(\frac{1}{\eta^{-1}(r_k)}\right) \leq \mathcal{S}_k \leq Cr_k\varphi^{-1}\left(\frac{1}{\eta^{-1}(r_k)}\right) \qquad \text{for } k \geq 0.$$

Consequently, $S(T,d,\varphi)$ is comparable with $\sum_{k\geq 0} r_k\varphi^{-1}(1/\eta^{-1}(r_k))$. This result was also obtained by M. Talagrand ([9, Theorem 5.2]).

2. Preliminary results. We recall that φ, ψ are conjugate Orlicz functions.

LEMMA 1. *The following inequalities hold*:

$$\varphi\left(\frac{\psi(x)}{x}\right) \leq \psi(x) \leq \varphi\left(\frac{2\psi(x)}{x}\right) \qquad \text{for } x \geq 0,$$

$$(7) \qquad \psi\left(\frac{\varphi(x)}{x}\right) \leq \varphi(x) \leq \psi\left(\frac{2\varphi(x)}{x}\right) \qquad \text{for } x \geq 0.$$

In the symmetric version we have

$$x \leq \varphi^{-1}(x)\psi^{-1}(x) \leq 2x \qquad \text{for } x \geq 0.$$

Proof. Fix $x \geq 0$. By the Young inequality we obtain

$$2\psi(x) = \frac{2\psi(x)}{x}x \leq \psi(x) + \varphi\left(\frac{2\psi(x)}{x}\right).$$

Hence $\psi(x) \leq \varphi(2\psi(x)/x)$. To prove the right-hand inequality of (7) notice that since $\varphi(x) = \sup_{y \geq 0}(xy - \psi(y))$, we have, for some $y \geq 0$,

$$\varphi\left(\frac{\psi(x)}{x}\right) = y\,\frac{\psi(x)}{x} - \psi(y).$$

It remains to prove that

$$(8) \qquad\qquad y\,\frac{\psi(x)}{x} - \psi(y) \leq \psi(x).$$

If $x \leq y$, then the convexity of ψ gives $y\psi(x)/x \leq \psi(y)$. If $x > y$, then $y\psi(x)/x \leq \psi(x)$. This yields (8), and consequently

$$\varphi\left(\frac{\psi(x)}{x}\right) \leq \psi(x) \leq \varphi\left(\frac{2\psi(x)}{x}\right). \quad \blacksquare$$

LEMMA 2. *The functions* φ, ψ *have the following properties*:

- $x\varphi(1/x)$, $x\psi(1/x)$ *are convex and decreasing*;
- $x\varphi^{-1}(1/x)$, $x\psi^{-1}(1/x)$ *are concave and increasing*.

Proof. It is enough to prove the result for φ. By the definition $\varphi(x) = \sup_{y \geq 0}(yx - \psi(y))$, so $x\varphi(1/x) = \sup_{y \geq 0}(y - x\psi(y))$. The supremum of convex functions is a convex function, the supremum of decreasing functions is a decreasing function.

Similarly we observe that $\varphi^{-1}(x) = \inf_{y \geq 0}(x + \psi(y))/y$. Hence $x\varphi^{-1}(1/x) = \inf_{y \geq 0}(1 + x\psi(y))/y$. The infimum of concave functions is a concave function, the infimum of increasing functions is an increasing function. $\blacksquare$

3. Proof of Theorem 1. Fix points $t = (t_i)_{i=1}^n, s = (s_i)_{i=1}^n \in T$. Let g be a smooth function on $\mathbb{R}^n$. We define $F_t : T \times [0, r] \to T$ by the formula

$$F_t(u, \varepsilon) = \left(1 - \frac{\varepsilon}{r}\right)t + \frac{\varepsilon}{r}\,u.$$

We have

$$(9) \qquad\qquad r^n \int_T g(u)\,du = \int_0^r \frac{\partial}{\partial \varepsilon}\left(\int_T g(F_t(u, \varepsilon))\varepsilon^n\,du\right)d\varepsilon.$$

It can be easily verified that

$$\frac{\partial}{\partial \varepsilon}g(F_t(u, \varepsilon)) = r^{-1}\sum_{i=1}^n (u_i - t_i)\frac{\partial}{\partial x_i}g(F_t(u, \varepsilon))$$

$$= \varepsilon^{-1}\sum_{i=1}^n (u_i - t_i)\frac{\partial}{\partial u_i}g(F_t(u, \varepsilon)).$$

Hence

$$\frac{\partial}{\partial \varepsilon}(g(F_t(u,\varepsilon))\varepsilon^n) = n\varepsilon^{n-1}g(F_t(u,\varepsilon)) + \varepsilon^{n-1}\sum_{i=1}^{n}(u_i - t_i)\frac{\partial}{\partial u_i}g(F_t(u,\varepsilon)),$$

which yields

$$\frac{\partial}{\partial \varepsilon}(g(F_t(u,\varepsilon))\varepsilon^n) = \varepsilon^{n-1}\sum_{i=1}^{n}\frac{\partial}{\partial u_i}(g(F_t(u,\varepsilon))(u_i - t_i)).$$

Applying the generalized Green–Gauss theorem (see Theorem 4.5.6 in [4]) which holds for Lipschitz boundaries, we get

$$\int_T \sum_{i=1}^{n}\frac{\partial}{\partial u_i}(g(F_t(u,\varepsilon))(u_i - t_i))\,du = \int_{\partial T} g(F_t(u,\varepsilon))\langle u - t, n(u)\rangle\,\sigma_{\partial T}(du),$$

where $\sigma_{\partial T}$ is the Lebesgue measure on the manifold ∂T, and $n(u)$ the normal vector to the boundary at $u \in \partial T$ such that $\langle n(u), n(u)\rangle = 1$ ($n(u)$ is well defined $\sigma_{\partial T}$-a.s.). Notice that the convexity of T yields $\langle u - t, n(u)\rangle \geq 0$. Setting $\sigma_t(u) := \langle u - t, n(u)\rangle\sigma_{\partial T}(du)$, we obtain, due to (9),

$$r^n \int_T g(u)\,du = \int_0^r \int_{\partial T} g(F_t(u,\varepsilon))\varepsilon^{n-1}\,\sigma_t(du)\,d\varepsilon.$$

By the standard approximation this equality can be easily generalized to any bounded Borel function g on T. We also verify that $n|T| = \sigma_t(\partial T)$ (consider $g \equiv 1$), and consequently for each $g \in \mathfrak{B}(T)$,

$$(10) \qquad n^{-1}r^n \int_T g(u)\,du = \int_0^r \int_{\partial T} g(F_t(u,\varepsilon))\varepsilon^{n-1}\,\sigma_t(du)\,d\varepsilon.$$

We define $\partial F_t : \partial T \times [0,1] \to T$ by $\partial F_t = F_t|\partial T \times [0,1]$. Then (10) implies

$$(11) \qquad |A|/|T| = \sigma_t \otimes d(\varepsilon/r)^n((\partial F_t)^{-1}(A)) \quad \text{for } A \in \mathcal{B}(T).$$

Define $a_t : [0,r] \to \mathbb{R}$ by

$$a_t(\varepsilon) := \int_{\partial T} f(F_t(u,\varepsilon))\,\sigma_t(du) = \int_{\partial T} f\left(\left(1 - \frac{\varepsilon}{r}\right)t + \frac{\varepsilon}{r}u\right)\sigma_t(du).$$

Clearly a_t satisfies the Lipschitz condition (because f does) and

$$a_t(0) = f(t), \qquad a_t(r) = \int_{\partial T} f(u)\,\sigma_t(du).$$

Since f is Lipschitz, it has bounded ∇f, $|\cdot|$-a.s. on T. We can check that if f is differentiable at $F_t(u,\varepsilon)$ then

$$\frac{\partial}{\partial \varepsilon}f(F_t(u,\varepsilon)) = r^{-1}\langle \nabla f(F_t(u,\varepsilon)), u - t\rangle.$$

By the Fubini theorem and (11) we find that $f(F_t)$ is differentiable $d\varepsilon$-a.s. for $\sigma_t(du)$-almost all $u \in \partial T$. Consequently, $d\varepsilon$-a.s. we have

$$a'_t(\varepsilon) = r^{-1} \int_{\partial T} \langle \nabla f(F_t(u, \varepsilon)), u - t \rangle \, \sigma_t(du).$$

Then

$$|a'_t(\varepsilon)| \le r^{-1} \int_{\partial T} \|u - t\| \, \|\nabla f(F_t(u, \varepsilon))\|_* \, \sigma_t(du).$$

Clearly $\|u - t\| \le 2r$. Observe that $b_t : [0, r] \to \mathbb{R}$ is $d\varepsilon$-a.s. well defined by the formula

$$b_t(\varepsilon) := \int_{\partial T} \|\nabla f(F_t(u, \varepsilon))\|_* \, \sigma_t(du).$$

Hence $|a'_t(\varepsilon)| \le 2b_t(\varepsilon)$ $d\varepsilon$-a.s. By the Jensen inequality and (10) for each $B > 0$ we obtain

$$\int_0^r \varphi\left(\frac{1}{B} b_t(\varepsilon)\right) \varepsilon^{n-1} \, d\varepsilon \le \int_{\partial T} \int_0^r \varphi\left(\frac{1}{B} \|\nabla f(F_t(u, \varepsilon))\|_*\right) \varepsilon^{n-1} \, d\varepsilon \, \sigma_t(du)$$

$$= n^{-1} r^n \int_T \varphi\left(\frac{1}{B} \|\nabla f(u)\|_*\right) du.$$

The Young inequality (1) gives

$$\frac{b_t(\varepsilon)}{AB\varepsilon^{n-1}} \le \psi\left(\frac{1}{A\varepsilon^{n-1}}\right) + \varphi\left(\frac{1}{B} b_t(\varepsilon)\right).$$

Since a_t is Lipschitz we get

$$\left| f(t) - \int_{\partial T} f(u) \, \sigma_t(du) \right| = |a_t(0) - a_t(r)| = \left| \int_0^r a'_t(\varepsilon) \, d\varepsilon \right| \le 2 \int_0^r b_t(\varepsilon) \, d\varepsilon.$$

Thus

$$(12) \quad \left| f(t) - \int_{\partial T} f(u) \sigma_t(u) \right|$$

$$\le 2AB \left(\int_0^r \psi\left(\frac{1}{A\varepsilon^{n-1}}\right) \varepsilon^{n-1} \, d\varepsilon + n^{-1} r^n \int_T \varphi\left(\frac{1}{B} \|\nabla f(u)\|_*\right) du \right).$$

Again due to the generalized Green–Gauss theorem (this version holds for Lipschitz functions and Lipschitz boundaries) we obtain

$$\left| \int_{\partial T} f(u) \, \sigma_t(du) - \int_{\partial T} f(u) \, \sigma_s(du) \right|$$

$$= \frac{1}{n|T|} \left| \int_{\partial T} f(u) \langle s - t, n(u) \rangle \, \sigma_{\partial T}(du) \right|$$

$$= n^{-1} \left| \int_T \langle \nabla f(u), s - t \rangle \, du \right| \le 2r n^{-1} \int_T \|\nabla f(u)\|_* \, du.$$

By the Young inequality and since $y\psi(1/y)$ is decreasing we obtain

$$\frac{\|\nabla f(u)\|_*}{ABr^{n-1}} \leq \psi\left(\frac{1}{Ar^{n-1}}\right) + \varphi\left(\frac{1}{B}\|\nabla f(u)\|_*\right)$$

$$\leq n\int_0^r \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\frac{\varepsilon^{n-1}}{r^n}\,d\varepsilon + \varphi\left(\frac{1}{B}\|\nabla f(u)\|_*\right).$$

It follows that

$$(13) \quad \left|\oint_{\partial T} f(u)\,\sigma_t(du) - \oint_{\partial T} f(u)\,\sigma_s(du)\right|$$

$$\leq 2AB\left(\int_0^r \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon + n^{-1}r^n\oint_T \varphi\left(\frac{1}{B}\|\nabla f(u)\|_*\right)du\right).$$

By the definition $|T| = |B(0,r)| = r^n|B_{\|\cdot\|}(0,1)|$. Inequalities (12), (13) yield

$$|f(s) - f(t)| \leq \left|f(t) - \oint_{\partial T} f(u)\,\sigma_t(du)\right|$$

$$+ \left|\oint_{\partial T} f(u)\,\sigma_t(du) - \oint_{\partial T} f(u)\,\sigma_s(du)\right| + \left|f(s) - \oint_{\partial T} f(u)\,\sigma_s(du)\right|$$

$$\leq 6AB\left(\int_0^r \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon + \frac{1}{n|B_{\|\cdot\|}(0,1)|}\int_T \varphi\left(\frac{1}{B}\|\nabla f(u)\|_*\right)du\right). \quad\blacksquare$$

REMARK 2. In the above notation, for each $t \in T$ we have $\sigma_t(\partial T) = n|T| = nr^n|B_{\|\cdot\|}(0,1)|$.

COROLLARY 4. *For each* $t \in T$,

$$\left|f(t) - \oint_T f(u)\,du\right| \leq 6AB\left(\int_0^r \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon\right.$$

$$\left. + \frac{1}{n|B_{\|\cdot\|}(0,1)|}\int_T \varphi\left(\frac{1}{B}\|\nabla f(u)\|_*\right)du\right).$$

Proof. This observation is obvious. Theorem 1 yields

$$|f(t) - f(u)| \leq 6AB\left(\int_0^r \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon\right.$$

$$\left. + \frac{1}{n|B_{\|\cdot\|}(0,1)|}\int_T \varphi\left(\frac{1}{B}\|\nabla f(u)\|_*\right)du\right).$$

Integrating both sides and using $|f(t) - \oint_T f(u)\,du| \leq \oint_T |f(t) - f(u)|\,du$ we obtain the assertion. $\blacksquare$

4. Construction of the optimal process. We assume that $\eta'(0) = \infty$. In this section we prove the left-hand inequality of (6) in Theorem 2.

We define a stochastic process on the probability space $(T, \mathcal{B}(T), \lambda)$ by the formula

$$X(t, \omega) := \int_{d(\omega,0)}^{d(\omega,t)} g(\varepsilon) \, d\varepsilon, \quad \text{for } t, \omega \in T,$$

where $g(\varepsilon)$ is a positive function, integrable on each interval $[\delta, \eta(r)]$, $\delta > 0$, and such that $g(\varepsilon) = 0$ for $\varepsilon > \eta(r)$. Notice that the process X is separable. Suppose we have shown that

$$\mathbf{E}\varphi\left(\frac{|X(s) - X(t)|}{d(s,t)}\right) = \int_T \varphi\left(\frac{|X(s,\omega) - X(t,\omega)|}{d(s,t)}\right) \lambda(d\omega) \le 1.$$

Then the process $X(t)$, $t \in T$, satisfies condition (3). Since

$$X(\omega, \omega) = -\int_0^{d(\omega,0)} g(\varepsilon) \, d\varepsilon, \quad X\left(\frac{\|\omega\| - r}{\|\omega\|}\,\omega, \omega\right) = \int_{d(\omega,0)}^{\eta(r)} g(\varepsilon) \, d\varepsilon,$$

we have $\sup_{s,t \in T} |X(s,\omega) - X(t,\omega)| = \int_0^{\eta(r)} g(\varepsilon) \, d\varepsilon$. Due to the definition of $S(T, d, \varphi)$ this proves that

(14)
$$\int_0^{\eta(r)} g(\varepsilon) \, d\varepsilon \le \mathbf{E} \sup_{s,t \in T} |X(s) - X(t)| \le S(T, d, \varphi).$$

The convexity of φ, $\varphi(0) = 0$ and the Jensen inequality imply

$$\int_T \varphi\left(\frac{|X(s,\omega) - X(t,\omega)|}{d(s,t)}\right) \lambda(d\omega)$$

$$\le \int_T \varphi\left(\frac{|d(s,\omega) - d(t,\omega)|}{d(s,t)} \int_{d(s,\omega)}^{d(t,\omega)} g(\varepsilon) \, d\varepsilon\right) \lambda(d\omega)$$

$$\le \int_T \frac{|d(s,\omega) - d(t,\omega)|}{d(s,t)} \left|\int_{d(s,\omega)}^{d(t,\omega)} \varphi(g(\varepsilon)) \, d\varepsilon\right| \lambda(d\omega)$$

$$= \frac{1}{d(s,t)} \int_T \left|\int_{d(s,\omega)}^{d(t,\omega)} \varphi(g(\varepsilon)) \, d\varepsilon\right| \lambda(d\omega).$$

The Fubini theorem yields

$$\int_T \left|\int_{d(s,\omega)}^{d(t,\omega)} \varphi(g(\varepsilon)) \, d\varepsilon\right| \lambda(d\omega) = \int_0^{\eta(r)} \varphi(g(\varepsilon))\lambda(B(s,\varepsilon) \triangle B(t,\varepsilon)) \, d\varepsilon,$$

where $\triangle$ is the symmetric difference. Observe that if $d(s,t) \geq \varepsilon$, then

$$\lambda(B(s,\varepsilon) \triangle B(t,\varepsilon)) \leq \lambda(B(s,\varepsilon)) + \lambda(B(t,\varepsilon)) \leq 2\frac{\eta^{-1}(\varepsilon)^n}{r^n}.$$

On the other hand, if $\varepsilon \geq d(s,t)$, then

$$B_{\|\cdot\|}\left(\frac{s+t}{2}, \eta^{-1}(\varepsilon) - \frac{1}{2}\|s-t\|\right) \subset B_{\|\cdot\|}(s, \eta^{-1}(\varepsilon)) \cap B_{\|\cdot\|}(t, \eta^{-1}(\varepsilon)),$$

and thus

$$\frac{|B(s,\varepsilon) \triangle B(t,\varepsilon)|}{|B(0,\eta(r))|} \leq 2\left(\frac{\eta^{-1}(\varepsilon)^n}{r^n} - \frac{(\eta^{-1}(\varepsilon) - \frac{1}{2}\|s-t\|)^n}{r^n}\right)$$

$$\leq n\|s-t\|\frac{\eta^{-1}(\varepsilon)^{n-1}}{r^n}.$$

Hence, for $\varepsilon \geq d(s,t)$ we have $\lambda(B(s,\varepsilon) \triangle B(t,\varepsilon)) \leq n\|s-t\|\eta^{-1}(\varepsilon)^{n-1}/r^n$. Consequently, if $d(s,t) \geq \eta(r)$ then

$$(15) \quad \int_T \varphi\left(\frac{|X(s,\omega) - X(t,\omega)|}{d(s,t)}\right)\mu(d\omega) \leq \frac{2}{d(s,t)}\int_0^{\eta(r)} \frac{\eta^{-1}(\varepsilon)^n}{r^n}\varphi(g(\varepsilon))\,d\varepsilon$$

$$\leq \frac{2}{\eta(r)}\int_0^{\eta(r)} \frac{\eta^{-1}(\varepsilon)^n}{r^n}\varphi(g(\varepsilon))\,d\varepsilon,$$

while if $d(s,t) \leq \eta(r)$, then

$$(16) \quad \int_T \varphi\left(\frac{|X(s,\omega) - X(t,\omega)|}{d(s,t)}\right)\lambda(d\omega)$$

$$\leq \frac{2}{d(s,t)}\int_0^{d(s,t)} \frac{\eta^{-1}(\varepsilon)^n}{r^n}\varphi(g(\varepsilon))\,d\varepsilon + \frac{\|s-t\|}{d(s,t)}\int_{d(s,t)}^{\eta(r)} n\frac{\eta^{-1}(\varepsilon)^{n-1}}{r^n}\varphi(g(\varepsilon))\,d\varepsilon.$$

The construction of g is as follows:

$$g(\varepsilon) := K^{-1}\frac{\mathcal{S}_k\eta^{-1}(\varepsilon)^n}{r^n\varepsilon}\psi\left(\frac{r^n\varepsilon}{\mathcal{S}_k\eta^{-1}(\varepsilon)^n}\right) \quad \text{for } r_{k+1} < \varepsilon \leq r_k,$$

where the constant $K \geq 1$ will be chosen later. From the convexity of φ and Lemma 1 we deduce

$$(17) \quad \varphi(g(\varepsilon)) \leq K^{-1}\psi\left(\frac{r^n\varepsilon}{\mathcal{S}_k\eta^{-1}(\varepsilon)^n}\right) \quad \text{for } r_{k+1} < \varepsilon \leq r_k, \ k \geq 0.$$

We now show that the process X satisfies condition (3) for such g.

First we assume that $d(s,t) \leq \eta(r) = r_0$. Then there exists m such that $r_{m+1} < d(s,t) \leq r_m$. Pick $k > m$. The definition of $\mathcal{S}_k$ and (17) yield

$$\int_{r_{k+1}}^{r_k} \frac{\eta^{-1}(\varepsilon)^n}{r^n}\varphi(g(\varepsilon))\,d\varepsilon \leq K^{-1}r_k\int_{r_{k+1}}^{r_k} \frac{\eta^{-1}(\varepsilon)^n}{r^n\varepsilon}\psi\left(\frac{r^n\varepsilon}{\mathcal{S}_k\eta^{-1}(\varepsilon)^n}\right)d\varepsilon \leq K^{-1}r_k.$$

Similarly we obtain

$$\int_{r_{m+1}}^{d(s,t)} \frac{\eta^{-1}(\varepsilon)^n}{r^n} \, \varphi(g(\varepsilon)) \, d\varepsilon \le K^{-1} d(s,t).$$

Since (5) gives $2r_{k+1} \le r_k$, it is clear that $r_k \le 2^{-k+m+1} d(s,t)$ for $k > m$. Applying the above inequalities, we get

$$(18) \quad \frac{2}{d(s,t)} \int_0^{d(s,t)} \frac{\eta^{-1}(\varepsilon)^n}{r^n} \, \varphi(g(\varepsilon)) \, d\varepsilon \le 2K^{-1}\Big(1 + \sum_{k>m} 2^{-k+m+1}\Big) = 6K^{-1}.$$

It remains to estimate the second integral in (16). Let $0 \le k < m$. By (17), the definition of $\mathcal{S}_k$ and since $\eta^{-1}(y)/y$ is increasing, we have

$$(19) \quad \int_{r_{k+1}}^{r_k} n \, \frac{\eta^{-1}(\varepsilon)^{n-1}}{r^n} \, \varphi(g(\varepsilon)) \, d\varepsilon$$

$$\le nK^{-1} \int_{r_{k+1}}^{r_k} \frac{\eta^{-1}(\varepsilon)^{n-1}}{r^n} \, \psi\left(\frac{r^n \varepsilon}{\mathcal{S}_k \eta^{-1}(\varepsilon)^n}\right) d\varepsilon$$

$$\le nK^{-1} \frac{r_{k+1}}{\eta^{-1}(r_{k+1})} \int_{r_{k+1}}^{r_k} \frac{\eta^{-1}(\varepsilon)^n}{r^n \varepsilon} \, \psi\left(\frac{r^n \varepsilon}{\mathcal{S}_k \eta^{-1}(\varepsilon)^n}\right) d\varepsilon = nK^{-1} \frac{r_{k+1}}{\eta^{-1}(r_{k+1})}.$$

In the same way, we prove

$$(20) \quad \int_{d(s,t)}^{r_m} n \, \frac{\eta^{-1}(\varepsilon)^{n-1}}{r^n} \, \varphi(g(\varepsilon)) \, d\varepsilon \le nK^{-1} \frac{d(s,t)}{\|s-t\|}.$$

Notice that (5) gives

$$\frac{\|s-t\|}{d(s,t)} \frac{r_{k+1}}{\eta^{-1}(r_{k+1})} \le 2^{-m+k+1} \quad \text{for } 0 \le k < m.$$

Inequalities (19) and (20) imply

$$(21) \quad \frac{\|s-t\|}{d(s,t)} \int_{d(s,t)}^{\eta(r)} n \, \frac{\eta^{-1}(\varepsilon)^{n-1}}{r^n} \, \varphi(g(\varepsilon)) \, d\varepsilon$$

$$\le nK^{-1}\Big(1 + \sum_{k=0}^{m-1} 2^{-m+k+1}\Big) \le 3nK^{-1}.$$

If we plug estimates (18), (21) into (16), we obtain

$$\mathbf{E}\varphi\left(\frac{|X(s) - X(t)|}{d(s,t)}\right) \le 3K^{-1}(2+n).$$

The second case is when $d(s,t) \geq \eta(r)$. We use (18) to get

$$\frac{2}{\eta(r)} \int\limits_0^{\eta(r)} \frac{\eta^{-1}(\varepsilon)^n}{r^n} \varphi(g(\varepsilon)) \, d\varepsilon \leq 6K^{-1}.$$

The above inequality and (15) imply

$$\mathbf{E}\varphi\left(\frac{|X(s) - X(t)|}{d(s,t)}\right) \leq 6K^{-1} \leq 3K^{-1}(2+n).$$

Therefore for (3) to hold we need that $K := 3(2+n)$.

By the definition of g and the numbers $\mathcal{S}_k$ it is clear that for all $k \geq 0$ we have $\int_{r_{k+1}}^{r_k} g(\varepsilon) \, d\varepsilon = K^{-1}\mathcal{S}_k$. Consequently, (14) shows that $\sum_{k=0}^{\infty} \mathcal{S}_k \leq KS(T, d, \varphi)$. The theorem is proved with the constant $K = 3(2+n)$. ∎

5. Some basic tools. Before we prove the right-hand inequality of (6) we establish some helpful results. We start by proving a fact which allows us to consider processes with a finite number of different Lipschitz paths.

LEMMA 3. *Fix any point $t_0 \in T$. Let $F \subset T$ be a finite set. For each process $X(t)$, $t \in T$, which satisfies condition (3) there exists a sequence $(Y_k)_{k\geq1}$ of processes which satisfy (3), have a finite number of different Lipschitz trajectories each and*

$$(22) \qquad \lim_{k\to\infty} Y_k(t) = X(t) - X(t_0), \quad \text{a.s. and in } L_1, \quad \text{for } t \in F.$$

In particular (22) *implies*

$$\lim_{k\to\infty} \mathbf{E} \sup_{s,t\in F} |Y_k(s) - Y_k(t)| = \mathbf{E} \sup_{s,t\in F} |X(s) - X(t)|.$$

Proof. The process $X(t)$, $t \in T$, is defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. We set $Y(t) := X(t) - X(t_0)$. It is clear that $Y(t)$, $t \in T$, satisfies condition (3) and moreover $\|Y(t)\|_\varphi \leq d(t, t_0)$, which implies $\mathbf{E}|Y(t)| < \infty$ for $t \in T$.

First we assume that $\mathcal{F}$ is a finite σ-algebra. Due to (3) we have

$$|Y(s,\omega) - Y(t,\omega)| \leq d(s,t)\varphi^{-1}(1/\mathbf{P}(A)) \quad \text{for } s,t \in T, \ \omega \in A,$$

where A is any atom in $\mathcal{F}$. Hence the process Y has $\mathbf{P}$-a.s. a finite number of different Lipschitz trajectories.

In the general case we use the fact that F is a finite set. There exists an increasing sequence $(\mathcal{F}_k)_{k\geq1}$ of finite σ-algebras whose union generates $\sigma(Y(t) : t \in F)$. Notice that $\mathbf{E}|Y(t)| < \infty$ for $t \in T$ (since $\|Y(t)\|_\varphi < \infty$), thus we can define $Y_k(t) := \mathbf{E}(Y(t) \,|\, \mathcal{F}_k)$, $t \in T$. By the Jensen inequality we get

$$\mathbf{E}\varphi\left(\frac{|Y_k(s) - Y_k(t)|}{d(s,t)}\right) \leq \mathbf{E}\varphi\left(\frac{|Y(s) - Y(t)|}{d(s,t)}\right) \leq 1 \quad \text{for } s,t \in T.$$

The process Y_k satisfies (3), hence **P**-a.s. it has a finite number of different Lipschitz trajectories. Modifying Y_k on a set of measure 0 we may assume that Y_k has a finite number of different Lipschitz trajectories. Clearly $Y_k(t) \to Y(t)$ **P**-a.s. for $t \in F$. Since $\mathbf{E}|Y(t)| < \infty$, the convergence is also in L_1. ∎

The next step is to prove an approximation of the numbers $\mathcal{S}_k$.

LEMMA 4. *We have*

$$\frac{1}{4}\, r_k\varphi^{-1}\left(\frac{2r^n}{\eta^{-1}(r_k)^n}\right) \leq \mathcal{S}_k \qquad for\ k \geq 0,$$

$$\mathcal{S}_k \leq r_{k+1}\varphi^{-1}\left(\frac{r^n}{\eta^{-1}(r_{k+1})^n}\right) \quad for\ k \in I.$$

Proof. From (5) we know that $r_k - r_{k+1} \geq \frac{1}{2}r_k$. Lemma 2 says that $y\psi(1/y)$ is decreasing. Thus, for $k \geq 0$ we have

$$1 = \int_{r_{k+1}}^{r_k} \frac{\eta^{-1}(\varepsilon)^n}{r^n\varepsilon}\, \psi\left(\frac{r^n\varepsilon}{\mathcal{S}_k\eta^{-1}(\varepsilon)^n}\right) d\varepsilon$$

$$\geq (r_k - r_{k+1}) \frac{\eta^{-1}(r_k)^n}{r^n r_k}\, \psi\left(\frac{r^n r_k}{\mathcal{S}_k\eta^{-1}(r_k)^n}\right) \geq \frac{\eta^{-1}(r_k)^n}{2r^n}\, \psi\left(\frac{r^n r_k}{\mathcal{S}_k\eta^{-1}(r_k)^n}\right).$$

That means

$$\psi^{-1}\left(\frac{2r^n}{\eta^{-1}(r_k)^n}\right) \geq \frac{r^n r_k}{\mathcal{S}_k\eta^{-1}(r_k)^n}.$$

By Lemma 1 (that is, by the inequality $\varphi^{-1}(y)\psi^{-1}(y) \leq 2y$) we obtain

$$\mathcal{S}_k \geq \frac{1}{4}\, r_k\varphi^{-1}\left(\frac{2r^n}{\eta^{-1}(r_k)^n}\right).$$

We prove the second inequality. Since $y\psi(1/y)$ is decreasing and $r_k - r_{k+1} = r_{k+1}$ for $k \in I$, we have

$$1 = \int_{r_{k+1}}^{r_k} \frac{\eta^{-1}(\varepsilon)^n}{r^n\varepsilon}\, \psi\left(\frac{r^n\varepsilon}{\mathcal{S}_k\eta^{-1}(\varepsilon)^n}\right) d\varepsilon$$

$$\leq r_{k+1}\frac{\eta^{-1}(r_{k+1})^n}{r_{k+1}r^n}\, \psi\left(\frac{r^n r_{k+1}}{\mathcal{S}_k\eta^{-1}(r_{k+1})^n}\right) = \frac{\eta^{-1}(r_{k+1})^n}{r^n}\, \psi\left(\frac{r^n r_{k+1}}{\mathcal{S}_k\eta^{-1}(r_{k+1})^n}\right).$$

Hence

$$\psi^{-1}\left(\frac{r^n}{\eta^{-1}(r_{k+1})^n}\right) \leq \frac{r^n r_{k+1}}{\mathcal{S}_k\eta^{-1}(r_{k+1})^n}.$$

Again, using Lemma 1 (the inequality $y \leq \varphi^{-1}(y)\psi^{-1}(y)$), we get

$$\mathcal{S}_k \leq r_{k+1}\varphi^{-1}\left(\frac{r^n}{\eta^{-1}(r_{k+1})^n}\right). \quad \blacksquare$$

Recall that λ is the normalized Lebesgue measure on T. For $0 < \varepsilon \leq \eta(r)$, we set

$$B_\varepsilon(t) := B((1 - \eta^{-1}(\varepsilon)/r)t, \varepsilon), \qquad S_\varepsilon(t) := S((1 - \eta^{-1}(\varepsilon)/r)t, \varepsilon).$$

Observe that $B_\varepsilon(t) = B_{\|\cdot\|}((1 - \eta^{-1}(\varepsilon)/r)t, \eta^{-1}(\varepsilon)) \subset T$, hence $\lambda(B_\varepsilon(t)) = \eta^{-1}(\varepsilon)^n/r^n$. For each $f \in C(T)$ we define $f_\varepsilon(t) := \fint_{B_\varepsilon(t)} f(u)\,\lambda(du)$.

Assume that $0 < \varepsilon \leq \eta(r)$. We denote by $\sigma_{t,\varepsilon}$ the Lebesgue measure on the manifold $S_\varepsilon(t)$. For each $e \in \mathbb{R}^n$ we define

$$\triangle_{t,\varepsilon}^e := \{u \in S_\varepsilon(t) : \langle e, n(u)\rangle \geq 0\},$$

where $n(u)$ is the normal vector at $u \in S_\varepsilon(t)$ such that $\langle n(u), n(u)\rangle = 1$ ($n(u)$ is well defined $\sigma_{t,\varepsilon}$-a.s.). Observe that for $e \in \mathbb{R}^n$ and $f \in C(T)$,

$$(23) \qquad \lim_{h \to +0} \frac{1}{h} \int_{B_\varepsilon(t+he)\setminus B_\varepsilon(t)} f(u)\,du$$

$$= \lim_{h \to +0} \frac{1}{h} \int_{B_\varepsilon(t)\setminus B_\varepsilon(t-he)} f(u)\,du$$

$$= \left(1 - \frac{\eta^{-1}(\varepsilon)}{r}\right) \int_{\triangle_{t,\varepsilon}^e} f(u)\langle e, n(u)\rangle\,\sigma_{t,\varepsilon}(du).$$

Let $\sigma_{t,\varepsilon}^e$ denote the positive measure on $\triangle_{t,\varepsilon}^e$ given by the formula $\sigma_{t,\varepsilon}^e(du) := \langle e, n(u)\rangle\sigma_{t,\varepsilon}(du)$. Notice that if $f \in \mathrm{Lip}(T)$, then ∇f exists $|\cdot|$-a.s.

LEMMA 5. *Fix* $0 < \varepsilon \leq \eta(r)$. *For each* $e \in \mathbb{R}^n$ *with* $\|e\| = 1$ *and* $f \in \mathrm{Lip}(T)$ *the following equality holds* $|\cdot|$-*a.s. on* T:

$$\langle \nabla f_\varepsilon(t), e\rangle = \left(1 - \frac{\eta^{-1}(\varepsilon)}{r}\right)\beta(e) \fint_{\triangle_{t,\varepsilon}^e} \fint_{\triangle_{t,\varepsilon}^{-e}} \frac{f(u) - f(v)}{2\eta^{-1}(\varepsilon)}\,\sigma_{t,\varepsilon}^e(du)\sigma_{t,\varepsilon}^{-e}(dv)$$

for some $\beta(e) \leq n$.

Proof. First we assume $h > 0$. Observe that

$$(24) \qquad f_\varepsilon(t + he) - f_\varepsilon(t)$$

$$= \fint_{B_\varepsilon(t+he)} f(u)\,\lambda(du) - \fint_{B_\varepsilon(t)} f(u)\,\lambda(du)$$

$$= \frac{1}{|B_\varepsilon(t)|}\left(\int_{B_\varepsilon(t+he)\setminus B_\varepsilon(t)} f(u)\,du - \int_{B_\varepsilon(t)\setminus B_\varepsilon(t+he)} f(u)\,du\right).$$

By (23) we obtain

$$(25) \qquad \lim_{h \to +0} \frac{1}{h} \int_{B_\varepsilon(t+he)\setminus B_\varepsilon(t)} f(u)\,du = \left(1 - \frac{\eta^{-1}(\varepsilon)}{r}\right) \int_{\triangle_{t,\varepsilon}^e} f(u)\,\sigma_{t,\varepsilon}^e(du).$$

Define

$$\beta(e) := \frac{2\eta^{-1}(\varepsilon)\sigma^e_{t,\varepsilon}(\triangle^e_{t,\varepsilon})}{|B_\varepsilon(t)|} = \frac{2\eta^{-1}(\varepsilon)\int_{\triangle^e_{t,\varepsilon}} \langle e, n(u)\rangle\,\sigma_{t,\varepsilon}(du)}{|B_\varepsilon(t)|}.$$

The homogeneity and symmetry imply

$$\beta(e) = \frac{\int_{\triangle^e_{0,\eta(r)}} \langle re, n(u)\rangle\,\sigma_{0,\eta(r)}(du)}{|B_{\eta(r)}(0)|} + \frac{\int_{\triangle^{-e}_{0,\eta(r)}} \langle -re, n(u)\rangle\,\sigma_{0,\eta(r)}(du)}{|B_{\eta(r)}(0)|}.$$

By Remark 2, for each $t \in B_{\eta(r)}(0) = B(0, \eta(r))$ we have

$$\int_{S_{\eta(r)}(0)} \langle u - t, n(u)\rangle\,\sigma_{0,\eta(r)}(du) = n|B_{\|\cdot\|}(0,r)| = n|B_{\eta(r)}(0)|.$$

Applying this equality for $t = -re$, $t = re$ and $t = 0$, we get

$$\beta(e) = \frac{\int_{\triangle^e_{0,\eta(r)}} \langle u + re, n(u)\rangle\,\sigma_{0,\eta(r)}(du)}{|B_{\eta(r)}(0)|}$$
$$+ \frac{\int_{\triangle^{-e}_{0,\eta(r)}} \langle u - re, n(u)\rangle\,\sigma_{0,\eta(r)}(du)}{|B_{\eta(r)}(0)|} - \frac{\int_{\triangle^e_{0,\eta(r)}} \langle u, n(u)\rangle\,\sigma_{0,\eta(r)}(du)}{|B_{\eta(r)}(0)|}$$
$$- \frac{\int_{\triangle^{-e}_{0,\eta(r)}} \langle u, n(u)\rangle\,\sigma_{0,\eta(r)}(du)}{|B_{\eta(r)}(0)|} \le n + n - n = n.$$

We have used the fact that $\sigma_{0,\eta(r)}(S_{\eta(r)}(0)\backslash(\triangle^e_{0,\eta(r)} \cup \triangle^{-e}_{0,\eta(r)})) = 0$. Observe that by (23),

$$(26) \quad \lim_{h\to+0} \frac{1}{h} \int_{B_\varepsilon(t)\backslash B_\varepsilon(t+he)} f(u)\,\lambda(du) = \left(1 - \frac{\eta^{-1}(\varepsilon)}{r}\right) \int_{\triangle^{-e}_{t,\varepsilon}} f(u)\,\sigma^{-e}_{t,\varepsilon}(du).$$

Moreover $\beta(e) = \beta(-e)$, hence applying (24)–(26) we obtain

$$\lim_{h\to+0} \frac{1}{h}\left(f_\varepsilon(t + he) - f_\varepsilon(t)\right)$$
$$= \left(1 - \frac{\eta^{-1}(\varepsilon)}{r}\right)\beta(e) \fint_{\triangle^e_{t,\varepsilon}} \fint_{\triangle^{-e}_{t,\varepsilon}} \frac{f(u) - f(v)}{2\eta^{-1}(\varepsilon)}\,\sigma^e_{t,\varepsilon}(du)\,\sigma^{-e}_{t,\varepsilon}(dv).$$

The case of $h < 0$ can be treated in a similar way. ∎

6. Estimation from above. We assume that $\eta'(0) = \infty$. In this section we prove the right-hand inequality of (6) in Theorem 2.

Write $B_k(x) := B_{r_k}(x)$ for $x \in T$. Notice that

$$(27) \qquad \lambda(B_k(x)) = \eta^{-1}(r_k)^n/r^n \quad \text{for } x \in T.$$

For each $k \geq 0$ we define a linear operator $S_k : C(T) \rightarrow C(T)$ by the formula

$$S_k f(x) := f_{r_k}(x) = \fint_{B_k(x)} f(u) \, \lambda(du) \quad \text{for } x \in T.$$

If $f, g \in C(T)$, $k \geq 0$, then the following properties can be easily derived:

1. $S_k 1 = 1$;
2. if $f \leq g$, then $S_k f \leq S_k g$ and so $|S_k f| \leq S_k |f|$;
3. $S_0 f = \int_T f(u) \, \lambda(du)$, hence $S_k S_0 f = S_0 f$;
4. $\lim_{k \to \infty} S_k f(x) = f(x)$.

Observe that if $f \in \text{Lip}(T)$, then $S_k f$ is also Lipschitz and thus differentiable $|\cdot|$-a.s. Fix $m > 0$. For $0 \leq k \leq m$, we define operators $L_k, R_k, T_k : C(T) \rightarrow C(T)$. We put $L_m = \text{Id}$, $R_m = \text{Id}$, and for $k < m$,

$$L_k := L_{k+1} S_{k+1}, \quad R_k := R_{k+1}, \quad \text{if } k \in I;$$
$$L_k := L_{k+1}, \quad R_k := S_{k+1} R_{k+1}, \quad \text{if } k \in J.$$

Set also $T_k := L_k S_k R_k$ for $0 \leq k \leq m$. The properties of S_k imply corresponding properties of the operators L_k, R_k, T_k. For $0 \leq k \leq m$ and $f, g \in C(T)$, we have:

1. $L_k 1 = 1$, and if $f \leq g$, then $L_k f \leq L_k g$;
2. if $f \in \text{Lip}(T)$, then $R_k f$ is $|\cdot|$-a.s. differentiable on T;
3. the function $T_0 f$ is constant;
4. $|T_m f(t) - T_0 f(t)| \leq \sum_{k=0}^{m-1} |T_k f(t) - T_{k+1} f(t)|$.

Fix $f \in \text{Lip}(T)$ and $s, t \in T$. We will analyse $|T_{k+1} f(t) - T_k f(t)|$. There are two cases: either $k \in I$ or $k \in J$. In fact, we use two different methods.

CASE 1. Fix $k \in I$, $k < m$. By the definition, we have

$$T_{k+1} f(t) - T_k f(t) = L_{k+1} S_{k+1} (\text{Id} - S_k) R_{k+1} f(t).$$

Clearly $R_k = R_{k+1}$ for $k \in I$. For $g := R_k f$, it can be easily checked that

$$|S_{k+1} (\text{Id} - S_k) g(w)| \leq \fint_{B_{k+1}(w)} \fint_{B_k(u)} |g(u) - g(v)| \, \lambda(dv) \, \lambda(du).$$

For each Orlicz function φ we have

$$x \leq 1 + \frac{\varphi(xy)}{\varphi(y)} \quad \text{for } x \geq 0, \, y > 0.$$

Thus

$$\frac{|g(u) - g(v)|}{10 r_k \varphi^{-1}\left(\frac{1}{\lambda(B_{k+1}(w))}\right)} \leq 1 + \lambda(B_{k+1}(w)) \varphi\left(\frac{|g(u) - g(v)|}{10 r_k}\right).$$

Consequently, for $u \in B_{k+1}(w)$,

$$|g(u) - g(v)| \leq 10 r_k \varphi^{-1}\left(\frac{1}{\lambda(B_{k+1}(w))}\right) \left(1 + \lambda(B_{k+1}(w)) \varphi\left(\frac{|g(u) - g(v)|}{10 r_k}\right)\right).$$

Hence

$$|S_{k+1}(\text{Id} - S_k)g(w)|$$

$$\leq 10 r_k \varphi^{-1}\left(\frac{r^n}{\eta^{-1}(r_{k+1})^n}\right)\left(1 + \int_T \int_{B_k(u)} \varphi\left(\frac{|g(u) - g(v)|}{10 r_k}\right) \lambda(dv)\,\lambda(du)\right).$$

Notice that Lemma 4 and the equality $r_k = 2r_{k+1}$ yield

$$\frac{1}{8}\, r_k \varphi^{-1}\left(\frac{r^n}{\eta^{-1}(r_{k+1})^n}\right) \leq \frac{1}{4}\, r_{k+1} \varphi^{-1}\left(\frac{2r^n}{\eta^{-1}(r_{k+1})^n}\right) \leq \mathcal{S}_{k+1}.$$

Take $K_1 := 80$. Using property 1 of L_{k+1}, we obtain

$$(28)\quad |T_{k+1}f(t) - T_k f(t)|$$

$$\leq K_1 \mathcal{S}_{k+1}\left(1 + \int_T \int_{B_k(u)} \varphi\left(\frac{|R_k f(u) - R_k f(v)|}{10 r_k}\right) \lambda(dv)\,\lambda(du)\right).$$

CASE 2. Fix $k \in J$, $k < m$. It is clear that

$$T_{k+1}f(t) - T_k f(t) = L_{k+1}(\text{Id} - S_k)S_{k+1}R_{k+1}f(t).$$

For $k \in J$, we have $R_k = S_{k+1}R_{k+1}$. For $g := R_k f$, by the definition we get

$$|(\text{Id} - S_k)g(w)| = \left|g(w) - \int_{B_k(w)} g(u)\,du\right|.$$

Property 2 of R_k implies that g is Lipschitz. Moreover $w \in B_k(w)$, thus from Corollary 4, we obtain the crucial inequality: for any $A, B > 0$,

$$(29)\quad \left|g(w) - \int_{B_k(w)} g(u)\,du\right| \leq 6AB\left(\int_0^{\eta^{-1}(r_k)} \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon\right.$$

$$\left. + \frac{1}{n|B_{\|\cdot\|}(0,1)|} \int_{B_k(w)} \varphi\left(\frac{1}{B}\,\|\nabla g(u)\|_*\right) du\right).$$

We have used the fact that $B_k(w) = B_{\|\cdot\|}((1 - \eta^{-1}(r_k)/r)w, \eta^{-1}(r_k))$. It remains to choose the constants A and B. For $k \in J$, we have $2r_k/\eta^{-1}(r_k) = r_{k+1}/\eta^{-1}(r_{k+1})$, so we can take

$$B := 10\beta\, \frac{r_k}{\eta^{-1}(r_k)} = 5\beta\, \frac{r_{k+1}}{\eta^{-1}(r_{k+1})},$$

where $\beta > 0$ will be chosen later. Finding a suitable A is more difficult. First we observe that

$$(30)\quad \int_0^{\eta^{-1}(r_k)} \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon$$

$$= \left(\int_{2\eta^{-1}(r_{k+1})}^{\eta^{-1}(r_k)} + \int_0^{2\eta^{-1}(r_{k+1})}\right)\psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon.$$

Setting $\varepsilon = \frac{\eta^{-1}(r_k)}{r_k}\varepsilon'$ and applying $2\frac{r_k}{\eta^{-1}(r_k)} = \frac{r_{k+1}}{\eta^{-1}(r_{k+1})}$, we get

$$\int_{2\eta^{-1}(r_{k+1})}^{\eta^{-1}(r_k)} \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon = \int_{r_{k+1}}^{r_k} \psi\left(\frac{r_k^{n-1}}{A\eta^{-1}(r_k)^{n-1}\varepsilon^{n-1}}\right)\frac{\eta^{-1}(r_k)^n}{r_k^n}\varepsilon^{n-1}\,d\varepsilon.$$

Since $\eta^{-1}(y)/y$ is increasing $\eta^{-1}(\varepsilon)^n \le \frac{\eta^{-1}(r_k)^n}{r_k^n}\varepsilon^n$ for $r_{k+1} < \varepsilon \le r_k$. By Lemma 2 the function $y\psi(1/y)$ is decreasing, so

(31)
$$\int_{2\eta^{-1}(r_{k+1})}^{\eta^{-1}(r_k)} \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon$$

$$\le \int_{r_{k+1}}^{r_k} \psi\left(\frac{\eta^{-1}(r_k)}{r_k}\frac{\varepsilon}{A\eta^{-1}(\varepsilon)^n}\right)\frac{\eta^{-1}(\varepsilon)^n}{\varepsilon}\,d\varepsilon.$$

For the other integral in (30) we proceed differently. Put $\varepsilon = 2\eta^{-1}(\varepsilon')$. Then

$$\int_0^{2\eta^{-1}(r_{k+1})} \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon$$

$$= 2^n \int_0^{r_{k+1}} \psi\left(\frac{1}{A2^{n-1}\eta^{-1}(\varepsilon)^{n-1}}\right)\eta^{-1}(\varepsilon)^{n-1}\eta^{-1}(\varepsilon)'\,d\varepsilon.$$

We can assume that $\eta^{-1}(\varepsilon)'$ is right continuous with left limits. Since $\eta^{-1}(\varepsilon)$ is a convex function, we have $\eta^{-1}(\varepsilon)/\varepsilon \le \eta^{-1}(\varepsilon)'$. By the convexity of ψ we get

$$\psi\left(\frac{1}{A2^{n-1}\eta^{-1}(\varepsilon)^{n-1}}\right) \le \frac{\eta^{-1}(\varepsilon)}{\eta^{-1}(\varepsilon)'\varepsilon}\psi\left(\frac{\eta^{-1}(\varepsilon)'\varepsilon}{A2^{n-1}\eta^{-1}(\varepsilon)^n}\right)$$

$$\le \frac{\eta^{-1}(\varepsilon)}{2^{n-1}\eta^{-1}(\varepsilon)'\varepsilon}\psi\left(\frac{\eta^{-1}(\varepsilon)'\varepsilon}{A\eta^{-1}(\varepsilon)^n}\right).$$

Consequently,

(32)
$$\int_0^{2\eta^{-1}(r_{k+1})} \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon \le 2\int_0^{r_{k+1}} \psi\left(\frac{\eta^{-1}(\varepsilon)'\varepsilon}{A\eta^{-1}(\varepsilon)^n}\right)\frac{\eta^{-1}(\varepsilon)^n}{\varepsilon}\,d\varepsilon.$$

The derivative $\eta^{-1}(\varepsilon)'$ can be controlled on the interval $[0, r_{k+1}]$. Indeed, the convexity of η^{-1} implies

$$\eta^{-1}(\varepsilon)' \le \frac{\eta^{-1}(2\varepsilon) - \eta^{-1}(\varepsilon)}{\varepsilon} \le \frac{\eta^{-1}(2\varepsilon)}{\varepsilon}.$$

Using (5), we obtain $\eta^{-1}(\varepsilon)' \le 2^{k+2-i}\eta^{-1}(r_k)/r_k$ for $r_{i+1} \le \varepsilon \le r_i$ and $i > k$.

Plugging this estimate into (32), we obtain

$$(33) \quad \int_0^{2\eta^{-1}(r_{k+1})} \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon$$

$$\leq 2\sum_{i=k+1}^{\infty} \int_{r_{i+1}}^{r_i} \psi\left(2^{k-i}\frac{\eta^{-1}(r_k)}{r_k}\frac{4\varepsilon}{A\eta^{-1}(\varepsilon)^n}\right)\frac{\eta^{-1}(\varepsilon)^n}{\varepsilon}\,d\varepsilon.$$

We define the constant A by the formula

$$A := \frac{n\eta^{-1}(r_k)}{r^n r_k}\left(\mathcal{S}_k + 4\sum_{i>k}\alpha^{k-i}\mathcal{S}_i\right),$$

where α must satisfy the condition $1 < \alpha < 2$. The definition of $\mathcal{S}_k$ and the convexity of ψ give

$$(34) \quad \int_{r_{k+1}}^{r_k} \psi\left(\frac{\eta^{-1}(r_k)}{r_k}\frac{\varepsilon}{A\eta^{-1}(\varepsilon)^n}\right)\frac{\eta^{-1}(\varepsilon)^n}{\varepsilon}\,d\varepsilon$$

$$\leq \int_{r_{k+1}}^{r_k} \psi\left(\frac{r^n\varepsilon}{n\mathcal{S}_k\eta^{-1}(\varepsilon)^n}\right)\frac{\eta^{-1}(\varepsilon)^n}{\varepsilon}\,d\varepsilon \leq n^{-1}r^n.$$

In the same way, for $i > k$ we prove

$$(35) \quad \int_{r_{i+1}}^{r_i} \psi\left(2^{k-i}\frac{\eta^{-1}(r_k)}{r_k}\frac{4\varepsilon}{A\eta^{-1}(\varepsilon)^n}\right)\frac{\eta^{-1}(\varepsilon)^n}{\varepsilon}\,d\varepsilon$$

$$\leq \int_{r_{i+1}}^{r_i} \psi\left(\left(\frac{\alpha}{2}\right)^{i-k}\frac{r^n\varepsilon}{n\mathcal{S}_i\eta^{-1}(\varepsilon)^n}\right)\frac{\eta^{-1}(\varepsilon)^n}{\varepsilon}\,d\varepsilon$$

$$\leq n^{-1}\left(\frac{\alpha}{2}\right)^{i-k}\int_{r_{i+1}}^{r_i} \psi\left(\frac{r^n\varepsilon}{\mathcal{S}_i\eta^{-1}(\varepsilon)^n}\right)\frac{\eta^{-1}(\varepsilon)^n}{\varepsilon}\,d\varepsilon = n^{-1}r^n\left(\frac{\alpha}{2}\right)^{i-k}.$$

Inequalities (30), (31) and (33)–(35) yield

$$\int_0^{\eta^{-1}(r_k)} \psi\left(\frac{1}{A\varepsilon^{n-1}}\right)\varepsilon^{n-1}\,d\varepsilon \leq n^{-1}r^n\left(1 + 2\sum_{i>k}\left(\frac{\alpha}{2}\right)^{i-k}\right) = n^{-1}r^n\frac{2+\alpha}{2-\alpha}.$$

Define $\mathcal{S}_k' := \mathcal{S}_k + 4\sum_{i\geq k+1}^{\infty}\alpha^{k-i}\mathcal{S}_i$ and $K_2 := 60\beta$. By (29) we obtain

$$|(\mathrm{Id} - S_k)g(w)|$$

$$\leq K_2\mathcal{S}_k'\left(\frac{2+\alpha}{2-\alpha} + \frac{nr^{-n}}{n|B_{\|\cdot\|}(0,1)|}\int_{B_k(w)}\varphi\left(\frac{\eta^{-1}(r_{k+1})}{5\beta r_{k+1}}\|\nabla g(u)\|_*\right)du\right)$$

$$\leq K_2\mathcal{S}_k'\left(\frac{2+\alpha}{2-\alpha} + \fint_T \varphi\left(\frac{\eta^{-1}(r_{k+1})}{5\beta r_{k+1}}\|\nabla g(u)\|_*\right)\lambda(du)\right),$$

where we have used the fact that $|B_{\|\cdot\|}(0,1)| = r^{-n}|T|$. Property 1 of L_{k+1} implies

$$|T_{k+1}f(t) - T_kf(t)| \leq K_2 \mathcal{S}_k' \left(\frac{2+\alpha}{2-\alpha} + \int_T \varphi \left(\frac{\eta^{-1}(r_{k+1})}{5\beta r_{k+1}} \|\nabla R_k f(u)\|_* \right) \lambda(du) \right).$$

The last part. The estimates in the cases $k \in I$ and $k \in J$ give

$$|T_m f(t) - T_0 f(t)| \leq \sum_{k=0}^{m-1} |T_{k+1} f(t) - T_k f(t)|$$

$$\leq K_1 \sum_{k \in I, \, k < m} \mathcal{S}_{k+1} \left(1 + \int_T \int_{B_k(u)} \varphi \left(\frac{|R_k f(u) - R_k f(v)|}{10 r_k} \right) \lambda(dv) \, \lambda(du) \right)$$

$$+ K_2 \sum_{k \in J, \, k < m} \mathcal{S}_k' \left(\frac{2+\alpha}{2-\alpha} + \int_T \varphi \left(\frac{\eta^{-1}(r_{k+1})}{5\beta r_{k+1}} \|\nabla R_k f(u)\|_* \right) \lambda(du) \right).$$

Property 3 of the operators T_k implies that $T_0 f$ is a constant function. Hence

$$(36) \quad |S_m f(s) - S_m f(t)| = |T_m f(s) - T_m f(t)|$$

$$\leq 2K_1 \sum_{k \in I, \, k < m} \mathcal{S}_{k+1} \left(1 + \int_T \int_{B_k(u)} \varphi \left(\frac{|R_k f(u) - R_k f(v)|}{10 r_k} \right) \lambda(dv) \, \lambda(du) \right)$$

$$+ 2K_2 \sum_{k \in J, \, k < m} \mathcal{S}_k' \left(\frac{2+\alpha}{2-\alpha} + \int_T \varphi \left(\frac{\eta^{-1}(r_{k+1})}{5\beta r_{k+1}} \|\nabla R_k f(u)\|_* \right) \lambda(du) \right).$$

Notice that $B_k(u) \subset B(u, 2r_k)$. We use the above inequality to prove that for each process $X(t)$, $t \in T$, which satisfies (3) we have

$$\mathbf{E} \sup_{s,t \in T} |X(s) - X(t)| \leq K \sum_{k \geq 0} \mathcal{S}_k,$$

where the constant K depends only on n. By (2) and Lemma 3 we can assume that the process $X(t)$, $t \in T$, has a finite number of different Lipschitz trajectories.

LEMMA 6. *Let* $0 \leq k < m$. *For any* $u, v \in T$ *with* $d(u,v) \leq 2r_k$, *we have*

$$\mathbf{E}\varphi\left(\frac{|R_k X(u) - R_k X(v)|}{10 r_k} \right) \leq 1.$$

Proof. We set $X_k := R_k X$. If $R_k = \mathrm{Id}$, then condition (3) implies the lemma. Otherwise there exists $N > 0$ and a sequence $k = k_0 < k_1 < \cdots < k_N \leq m$ such that $X_k = R_k X = S_{k_1} \cdots S_{k_N} X$. For simplicity we write

$u_{k_0} := u$, $v_{k_0} := v$. We obtain the following equalities:

$$X_k(u) = \fint_{B_{k_1}(u_{k_0})} \fint_{B_{k_2}(u_{k_1})} \cdots \fint_{B_{k_N}(u_{k_{N-1}})} X(u_{k_N})\,\lambda(du_{k_N}) \cdots \lambda(du_{k_1}),$$

$$X_k(v) = \fint_{B_{k_1}(u_{k_0})} \fint_{B_{k_2}(v_{k_1})} \cdots \fint_{B_{k_N}(v_{k_{N-1}})} X(v_{k_N})\,\lambda(dv_{k_N}) \cdots \lambda(dv_{k_1}).$$

Take $u_{k_{i+1}} \in B_{k_i}(u_{k_i}) \subset B(u_{k_i}, 2r_{k_i})$. By the triangle inequality and (5) we have

$$d(u_{k_0}, u_{k_N}) \le \sum_{i=0}^{N-1} d(u_{k_i}, u_{k_{i+1}}) \le 2\sum_{i=0}^{N-1} r_{k_i} \le 4r_{k_0}, \quad d(v_{k_0}, v_{k_N}) \le 4r_{k_0}.$$

This means that for some probability measures ν_u, ν_v with supports respectively in $B(u, 4r_k)$, $B(v, 4r_k)$,

$$X_k(u) = \int_{B(u,4r_k)} X(w)\,\nu_u(dw), \quad X_k(v) = \int_{B(v,4r_k)} X(z)\,\nu_v(dz).$$

By the assumption $d(u, v) \le 2r_k$, hence $d(w, z) \le 10r_k$. The Jensen inequality, the Fubini theorem and (3) yield

$$\mathbf{E}\varphi\left(\frac{|X_k(u) - X_k(v)|}{10r_k}\right)$$

$$\le \fint_{B(u,4r_k)} \fint_{B(v,4r_k)} \mathbf{E}\varphi\left(\frac{|X(w) - X(z)|}{d(w,z)}\right) \nu_u(dw)\,\nu_v(dz) \le 1. \quad \blacksquare$$

The Auerbach lemma (for the proof see [10, Lemma 11, II.E]) states that there exists a biorthogonal system (b_i, b_i^*) in the space $\mathbb{R}^n \times \mathbb{R}^n$ such that $\|b_i\| = 1$, $\|b_i^*\|_* = 1$. Consequently, for each $v \in \mathbb{R}^n$,

$$(37) \qquad \|v\|_* = \sup_{u \in B_{\|\cdot\|}(0,1)} |\langle v, u \rangle|$$

$$\le \sup_{u \in B_{\|\cdot\|}(0,1)} \sum_{i=1}^{n} |\langle v, b_i \rangle|\,|\langle b_i^*, u \rangle| \le \sum_{i=1}^{n} |\langle v, b_i \rangle|.$$

LEMMA 7. *For* $|\cdot|$*-almost all* $t \in T$,

$$\mathbf{E}\varphi\left(\frac{\eta^{-1}(r_{k+1})}{5\beta r_{k+1}}\|\nabla R_k X(t)\|_*\right) \le 1, \quad 0 \le k < m,\ k \in J,$$

where $\beta := \sum_{i=1}^{n} \beta(b_i)$ $(\beta(b_i)$ *was defined in Lemma 5). Since* $\beta(b_i) \le n$, *we have* $\beta \le n^2$.

Proof. We put $X_k := R_k X$. The definition gives $X_k = S_{k+1} R_{k+1} X = (X_{k+1})_{r_{k+1}}$ for $k \in J$. Applying (37) and Lemma 5 we find that for $|\cdot|$-almost all $t \in T$,

$$\frac{\eta^{-1}(r_{k+1})}{5\beta r_{k+1}} \|\nabla X_k(t)\|_* \leq \sum_{i=1}^{n} \frac{1}{\beta} \frac{\eta^{-1}(r_{k+1})}{5r_{k+1}} |\langle (\nabla X_{k+1})_{r_{k+1}}(t), b_i\rangle|$$

$$\leq \sum_{i=1}^{n} \frac{\beta(b_i)}{\beta} \oint_{\triangle^{b_i}_{t,r_{k+1}}} \oint_{\triangle^{-b_i}_{t,r_{k+1}}} \frac{|X_{k+1}(u) - X_{k+1}(v)|}{10 r_{k+1}} \sigma_1^i(du)\, \sigma_2^i(dv),$$

where $\sigma_1^i(du) = \sigma_{t,r_{k+1}}^{b_i}(du)$, $\sigma_2^i(dv) = \sigma_{t,r_{k+1}}^{-b_i}(dv)$. The Jensen inequality yields

$$\varphi\left(\frac{\eta^{-1}(r_{k+1})}{5\beta r_{k+1}} \|\nabla X_k(t)\|_*\right)$$

$$\leq \sum_{i=1}^{n} \frac{\beta(b_i)}{\beta} \oint_{\triangle^{b_i}_{t,r_{k+1}}} \oint_{\triangle^{-b_i}_{t,r_{k+1}}} \varphi\left(\frac{|X_{k+1}(u) - X_{k+1}(v)|}{10 r_{k+1}}\right) \sigma_1^i(du)\, \sigma_2^i(dv).$$

Notice that $d(u,v) \leq 2r_{k+1}$ for $u \in \triangle^{b_i}_{t,r_{k+1}}$, $v \in \triangle^{-b_i}_{t,r_{k+1}}$. The Fubini theorem and Lemma 6 imply

$$\mathbf{E}\varphi\left(\frac{\eta^{-1}(r_{k+1})}{5\beta r_{k+1}} \|\nabla X_k(t)\|_*\right) \leq \sum_{i=1}^{n} \frac{\beta(b_i)}{\beta} = 1. \quad \blacksquare$$

By the Fubini theorem, Lemmas 6 and 7, and (36) we obtain

$$\mathbf{E}\sup_{s,t\in T} |S_m X(s) - S_m X(t)| \leq 4K_1 \sum_{k\in I,\, k<m} \mathcal{S}_{k+1} + 2K_2\left(1 + \frac{2+\alpha}{2-\alpha}\right) \sum_{k\in J,\, k<m} \mathcal{S}'_k.$$

Notice that $\sum_{k\in J,\, k<N} \mathcal{S}'_k \leq \left(1 + \frac{4}{\alpha-1}\right) \sum_{k\geq 0} \mathcal{S}_k$. We have proved that for some constant K which depends only n,

$$\mathbf{E}\sup_{s,t\in T} |S_m X(s) - S_m X(t)| \leq K \sum_{k\geq 0} \mathcal{S}_k.$$

Since $\lim_{m\to\infty} S_m X(t) = X(t)$, the Fatou lemma yields

$$\mathbf{E}\sup_{s,t\in T} |X(s) - X(t)| \leq \liminf_{m\to\infty} \mathbf{E}\sup_{s,t\in T} |S_m X(s) - S_m X(t)| \leq K \sum_{k\geq 0} \mathcal{S}_k.$$

This ends the proof of the theorem. $\blacksquare$

7. The case of $\eta'(0) < \infty$. We assume that $\eta'(0) < \infty$. Recall that there exists $m \geq 0$ such that $r_m > 0$ and $r_{m+1} = 0$. We have defined $\mathcal{S}_m$ as the infimum of all $c > 0$ such that

$$\int_0^{r_m} \frac{\lambda(B(0,\varepsilon))}{\varepsilon} \psi\left(\frac{\varepsilon}{c\lambda(B(0,\varepsilon))}\right) d\varepsilon = \int_0^{r_m} \frac{\eta^{-1}(\varepsilon)^n}{r^n \varepsilon} \psi\left(\frac{r^n \varepsilon}{c\eta^{-1}(\varepsilon)^n}\right) d\varepsilon \leq 1$$

and $\mathcal{S}_k = 0$ for $k > m$.

Proof of Theorem 2 in the case of $\eta'(0) < \infty$. We follow the proof of Theorem 2 in the case of $\eta'(0) = \infty$. The only difference is when $\mathcal{S}_m = \infty$. For all $0 < \delta \leq r_m/2$ we denote by $\mathcal{S}_m(\delta)$ numbers such that

$$\int\limits_{\delta}^{r_m} \frac{\lambda(B(0,\varepsilon))}{\varepsilon}\,\psi\left(\frac{\varepsilon}{\mathcal{S}_m(\delta)\lambda(B(0,\varepsilon))}\right) d\varepsilon$$

$$= \int\limits_{\delta}^{r_m} \frac{\eta^{-1}(\varepsilon)^n}{r^n\varepsilon}\,\psi\left(\frac{r^n\varepsilon}{\mathcal{S}_m(\delta)\eta^{-1}(\varepsilon)^n}\right) d\varepsilon = 1.$$

Since $\mathcal{S}_m = \infty$ we have $\lim_{\delta\to\infty} \mathcal{S}_m(\delta) = \infty$. The proof of the left-hand inequality of (6) in Theorem 2 implies $\mathcal{S}_m(\delta) \leq 3(n+2)S(T,d,\varphi)$. Hence $S(T,d,\varphi) = \infty$. This ends the proof. ∎

References

[1] P. Assouad, *Le mouvement brownien sur* [0, 1]; *applications* Φ-*sommantes et* (Φ, Ψ)-*sommantes*, in: Séminaire Laurent Schwartz 1969–1970: Applications radonifiantes, exp. 27, Centre de Math., École Polytech., Paris, 1970, 11 pp.

[2] W. Bednorz, *The majorizing measure theorem*, Ann. Probab. (2006), to appear.

[3] A. Cianchi, *Optimal Orlicz–Sobolev embeddings*, Rev. Mat. Iberoamericana 20 (2004), 427–474.

[4] H. Federer, *Geometric Measure Theory*, reprint of the 1969 edition, Classics in Math., Springer, Berlin, 1996.

[5] X. Fernique, *Caractérisation de processus à trajectoires majorées ou continues*, in: Séminaire de Probabilités XII, Lecture Notes in Math. 649, Springer, Berlin, 1978, 691–706.

[6] —, *Régularité de fonctions aléatoires non gaussiennes*, in: École d'Été de Probabilités de Saint-Flour XI-1981, Lecture Notes in Math. 976, Springer, Berlin, 1983, 1–74.

[7] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, reprint of the 1998 edition, Classics in Math., Springer, Berlin, 2001.

[8] M. Ledoux and M. Talagrand, *Probability in Banach Spaces. Isoperimetry and Processes*, Ergeb. Math. Grenzgeb. (3) 23, Springer, Berlin, 1991.

[9] M. Talagrand, *Sample boundedness of stochastic processes under increment conditions*, Ann. Probab. 18 (1990), 1–49.

[10] P. Wojtaszczyk, *Banach Spaces for Analysts*, Cambridge Stud. Adv. Math. 25, Cambridge Univ. Press, Cambridge, 1991.

Department of Mathematics
University of Warsaw
Banacha 2
02-097 Warszawa, Poland
E-mail: wbednorz@mimuw.edu.pl