ZOLOTAREV-TYPE DISTANCES

SERGEY G. BOBKOV

Abstract. We consider modified Zolotarev's ideal metrics for probability distributions on the real line and explore their connections with Lévy's, total variation, and transport distances. In particular, E. Rio's upper bound on the power transport (Kantorovich) distances is reversed under (necessary) moment assumptions and is sharpened for compactly supported measures. We also explain the role of such metrics in the study of the central limit theorem.

1. Introduction

Denote by $\mathfrak{P}_p$ the collection of all Borel probability measures μ on the real line $\mathbb{R}$ with finite absolute moments of order $p > 0$, i.e. such that $\int_{-\infty}^{\infty} |x|^p \, d\mu(x) < \infty$.

If two random variables X and Y have distributions μ and ν in $\mathfrak{P}_p$, the ideal Zolotarev distance of order $p > 0$ between μ and ν is defined as

$$\zeta_p(X,Y) = \zeta_p(\mu,\nu) = \sup \bigg| \int_{-\infty}^{\infty} u \, d\mu - \int_{-\infty}^{\infty} u \, d\nu \bigg|. \tag{1.1}$$

Putting $p = m + \alpha$ with $m \geq 0$ integer and $0 < \alpha \leq 1$, here the supremum is taken over all m times differentiable functions $u : \mathbb{R} \to \mathbb{R}$ whose m-th derivatives have Lipschitz semi-norms $\|u^{(m)}\|_{\mathrm{Lip}(\alpha)} \leq 1$, i.e.

$$|u^{(m)}(x) - u^{(m)}(y)| \leq |x-y|^\alpha, \quad x,y \in \mathbb{R}. \tag{1.2}$$

If p is a positive integer (and then $m = p-1$), this condition is equivalent to the pointwise bound $|u^{(p)}(x)| \leq 1$, $x \in \mathbb{R}$, assuming that $u(x)$ is p times differentiable.

Motivated by the problems of continuity and stability of stochastic models (such as approximation for distributions of sums of independent random vectors), these distances were introduced and studied by Zolotarev in a series of papers in the mid 70's [12], [13] (cf. also [14], [15]). They were considered for probability measures on an arbitrary Banach space using the notion of Frechet derivatives. Let us however focus on the one dimensional setting.

The ideal distances ζ_p have a number of remarkable properties, especially the following two properties. They are homogeneous with respect to (X,Y) of order p, i.e.

$$\zeta_p(cX, cY) = |c|^p \, \zeta_p(X,Y), \quad c \in \mathbb{R}. \tag{1.3}$$

Secondly, they are subadditive with respect to the convolution operation in the sense that

$$\zeta_p\Big(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i\Big) \leq \sum_{i=1}^n \zeta_p(X_i, Y_i), \tag{1.4}$$

which holds for any two collections of independent random variables (X_i) and (Y_i) with finite absolute moments of order p.

2010 *Mathematics Subject Classification.* Primary 60E, 60F.

Key words and phrases. Zolotarev ideal metrics, transport distances, central limit theorem.

School of Mathematics, University of Minnesota, Minneapolis, MN 55455 USA.

In particular, if X_i represent independent copies of a random variable X with mean zero and variance one, an application of (1.3)-(1.4) with $p = 3$ to the normalized sum $Z_n = \frac{1}{\sqrt{n}}\,(X_1 + \dots + X_n)$ yields a Berry-Esseen-type bound

$$\zeta_3(Z_n, Z) \le \frac{1}{\sqrt{n}}\,\zeta_3(X, Z), \tag{1.5}$$

where Z is a standard normal random variable. If additionally $\mathbb{E}X^3 = 0$, the standard rate is improved by involving the next index $p = 4$, in which case

$$\zeta_4(Z_n, Z) \le \frac{1}{n}\,\zeta_4(X, Z). \tag{1.6}$$

Similar relations hold true in the high dimensional situation. It is therefore not surprising that these distances have attracted a considerable attention in many respects, including the study of the relationship between ζ_p and other canonical metrics in the space of probability distributions. For example, it was shown by Senatov that ζ_p dominate the Lévy distance (and more generally – the Lévy-Prokhorov distance in high dimensions for the class of convex subsets of the Euclidean space, cf. [8], [9]). Similar assertions are proved by Rio [6], [7] for the power transport distances for probability distributions on the real line.

Note that any function u satisfying the Lipschitz property (1.2) admits a pointwise bound $|u(x)| \le c(1 + |x|^p)$ with a constant c depending on u, so that the integrals in (1.1) are finite for all μ, ν in $\mathfrak{P}_p$. However, for the finiteness of $\zeta_p(\mu, \nu)$ with $p > 1$ it is necessary and sufficient that the first m moments of μ and ν coincide:

$$\int_{-\infty}^{\infty} x^k\,d\mu(x) = \int_{-\infty}^{\infty} x^k\,d\nu(x), \quad k = 1, \dots, m. \tag{1.7}$$

This is a negative point in applications of Zolotarev distances, which motivates to modify ζ_p in order to deal with a real metric on $\mathfrak{P}_p$ possessing analogous properties like (1.3)-(1.4). Here is one natural choice to resolve the issue in the spirit of the original Zolotarev's approach to the class of the ideal distances.

Definition 1.1. Let $s > 0$. Given μ and ν in $\mathfrak{P}_p$, $p > 0$, put

$$\zeta_p^{(s)}(\mu, \nu) = \sup_{u \in U_p(s)} \bigg|\int_{-\infty}^{\infty} u\,d\mu - \int_{-\infty}^{\infty} u\,d\nu\bigg|, \tag{1.8}$$

where the supremum is taken over the collection $U_p(s)$ of all m times differentiable functions u on the real line satisfying the Lipschitz property (1.2) and such that

$$|u^{(k)}(x)| \le s^{p-k}, \quad x \in \mathbb{R},\ k = 1, \dots, m. \tag{1.9}$$

The latter is not a condition for the range $0 < p \le 1$, and thus $\zeta_p^{(s)}(\mu, \nu) = \zeta_p(\mu, \nu)$ for all $s > 0$. If $p > 1$, the situation is different, although

$$\zeta_p^{(s)}(\mu, \nu) \uparrow \zeta_p^{(\infty)}(\mu, \nu) \;=\; \zeta_p(\mu, \nu) \quad \text{as } s \to \infty.$$

The aim of this note is to explore basic properties of the distances $\zeta_p^{(s)}$, including the influence of the parameter s in applications to the central limit theorem and their connections with the Lévy metric, characteristic functions, weighted total variation, and especially transport distances. The power transport (also called Kantorovich or minimal) distance of order $p \geq 1$ is defined to be

$$W_p(\mu,\nu) = \inf_\pi \bigg(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x-y|^p \, d\pi(x,y)\bigg)^{1/p}, \quad p \geq 1, \tag{1.10}$$

where the infimum is taken over all Borel probability measures π on the plane $\mathbb{R}\times\mathbb{R}$ with marginal distributions μ and ν. For the range $0 < p \leq 1$, it is defined similarly as

$$W_p(\mu,\nu) = \inf_\pi \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x-y|^p \, d\pi(x,y), \quad 0 < p \leq 1. \tag{1.11}$$

The quantity W_p is finite and represents a metric in the space $\mathfrak{P}_p$. For a comprehensive discussion of this metric for probability measures on general spaces with abstract cost functions we refer the interested reader to [11], [1].

By the Kantorovich duality theorem (cf. e.g. [5]), applied on the real line with metric $d(x,y) = |x-y|^p$, the definition (1.11) leads to the identity

$$\zeta_p(\mu,\nu) = W_p(\mu,\nu), \quad 0 < p \leq 1.$$

As for the range $p \geq 1$, the main known tool connecting these two quantities is the following remarkable observation due to Rio (Theorem 3.1 in [7]).

Theorem 1.2. *Given μ and ν in $\mathfrak{P}_p$, $p \geq 1$, we have*

$$W_p(\mu,\nu) \leq \, C_p\, (\zeta_p(\mu,\nu))^{1/p} \tag{1.12}$$

with some constant C_p depending on p only.

It was also shown in [7] that the best constant in (1.12) satisfies $C_p \leq \frac{\pi}{\sqrt{e}}\, p$ for any integer $p \geq 4$. Actually, the assertion $C_p \leq Cp$ remains to hold for all real $p \geq 1$. Using the ideas in the original argument of E. Rio, one may replace the moment constraint (1.7) with the assumption on the support of the involved measures, and then we have:

Theorem 1.3. *If the measures μ,ν are supported on a finite interval of length s, then, for any $p \geq 1$,*

$$W_p^p(\mu,\nu) \leq \, 2\,(8p)^p\, \zeta_p^{(s)}(\mu,\nu). \tag{1.13}$$

Letting $s \to \infty$ in (1.13), we return to the relation (1.12). In this step, the assumption on the support of the measures can be relaxed to $\mu,\nu \in \mathfrak{P}_p$. It is in this sense (1.13) may be viewed as a sharpening of (1.12). On the other hand, the relations (1.12)-(1.13) can be easily reversed in a certain form.

Theorem 1.4. *Given μ and ν in $\mathfrak{P}_p$, $p \geq 1$,*

$$\zeta_p^{(s)}(\mu,\nu) \leq s^{p-1} W_p(\mu,\nu). \tag{1.14}$$

Moreover, if μ,ν are supported on a finite interval of length s and have equal moments up to order m, then

$$\zeta_p(\mu,\nu) \leq \frac{s^{p-1}}{\Gamma(p)}\, W_p(\mu,\nu). \tag{1.15}$$

In particular, $\zeta_p^{(s)}(\mu,\nu) \le W_p^p(\mu,\nu)$ for the range $s \le W_p(\mu,\nu)$ which is opposite to (1.13).

Returning to the scheme of i.i.d. copies of a random variable X with mean zero and variance one, let us also state the following sharpening of the central limit theorem (1.5)-(1.6) in terms of the Zolotarev-type distances.

Theorem 1.5. *If X has a finite absolute moment of order $p > 2$, then for any $s > 0$,*

$$\zeta_p^{(s)}(Z_n, Z) \,\le\, \sum_{k=3}^m \frac{s^{p-k}}{n^{\frac{k-2}{2}}\,k!}\,|\,\mathbb{E}X^k - \mathbb{E}Z^k| + \frac{1}{n^{\frac{p-2}{2}}}\,\frac{\Gamma(\alpha+1)}{\Gamma(p+1)}\,(\mathbb{E}\,|X|^p + \mathbb{E}\,|Z|^p). \tag{1.16}$$

For the range $2 < p \le 3$, the sum in (1.16) is vanishing, and this bound is simplified to

$$\zeta_p^{(s)}(Z_n, Z) \,\le\, \frac{1}{2n^{\frac{p-2}{2}}}\,(\mathbb{E}\,|X|^p + \mathbb{E}\,|Z|^p).$$

In the case $p > 3$, and if the moments of X and Z coincide up to order m, one may strengthen (1.16) by letting $s \to \infty$, and then we arrive at

$$\zeta_p(Z_n, Z) \,\le\, \frac{1}{n^{\frac{p-2}{2}}}\,\frac{\Gamma(\alpha+1)}{\Gamma(p+1)}\,(\mathbb{E}\,|X|^p + \mathbb{E}\,|Z|^p).$$

The paper is organized as follows. Basic properties of $\zeta_p^{(s)}$ including analogues of the relations (1.3)-(1.4) are recorded in Section 2. Then we turn to lower bounds for these metrics in terms of characteristic functions and the Lévy distance (which imply in particular the separation property for the Zolotarev-type distances). Upper bounds on $\zeta_p^{(s)}$ in terms of the weighted total variation distances are considered in Section 4, where we also include the derivation of the upper bound (1.16) in Theorem 1.5. In Section 5 we recall several representations for the one-dimensional formula (1.10) defining the transport distance W_p and then prove Theorem 1.4. Theorem 1.3 is proved in Section 6 except for a technical Lemma 6.2, which is proved in several steps in Sections 7-9.

2. Basic Properties of $\zeta_p^{(s)}$

Here we collect a few basic properties of Zolotarev-type distances. As before, write $p = m+\alpha$ with $m \ge 0$ integer and $0 < \alpha \le 1$.

First let us note that in the case $p \ge 1$ Definition 1.1 may be extended to all probability measures with finite first absolute moment (although the space $\mathfrak{P}_p$ is more natural in the study of various relations). Indeed, the condition (1.9) with $k = 1$ ensures that $|u(x)-u(0)| \le s^{p-1}|x|$ for all $x \in \mathbb{R}$. Hence the integrals in (1.8) are finite, and we have a uniform upper bound

$$\zeta_p^{(s)}(\mu,\nu) \le s^{p-1}\,(\mathbb{E}\,|X| + \mathbb{E}\,|Y|),$$

where we assume that the random variables X and Y have distributions μ and ν respectively. Hence, we obtain:

Proposition 2.1. $\zeta_p^{(s)}$ *represents a metric on $\mathfrak{P}_{\min(p,1)}$ for any parameter $s > 0$.*

The separation property follows from the lower bounds in Propositions 3.1-3.2.

The metrics $\zeta_p^{(s)}$ are equivalent to each other. If $p \leq 1$, there is no dependence on s, while for $p \geq 1$, it follows from (1.8) that

$$\zeta_p^{(s_1)}(\mu,\nu) \,\leq\, \zeta_p^{(s_2)}(\mu,\nu) \,\leq\, \Big(\frac{s_2}{s_1}\Big)^{p-1} \zeta_p^{(s_1)}(\mu,\nu), \quad s_1 < s_2. \tag{2.1}$$

Proposition 2.2. *Let the random variables X and Y have distributions in $\mathfrak{P}_{\min(p,1)}$. For any $b, c \in \mathbb{R}$, $c \neq 0$,*

$$\begin{aligned} \zeta_p^{(s)}(X+b, Y+b) &= \zeta_p^{(s)}(X,Y), \\ \zeta_p^{(s)}(cX, cY) &= |c|^p \, \zeta_p^{(s/|c|)}(X,Y). \end{aligned} \tag{2.2}$$

This is a modified variant of the homogeneity property (1.3), which in turn follows from (2.2) by letting $s \to \infty$.

Proof. For the first relation, just note that the class $U_p(s)$ is closed under the shifts $u \to u(x+b)$. Also, given u in $U_p(s)$, the function $u_c(x) = c^{-p}\, u(cx)$ has derivatives $u^{(k)}(x) = c^{-(p-k)}\, u(cx)$ which are bounded in absolute value by $(s/|c|)^{p-k}$ for $k = 1, \dots, m$. In addition,

$$\|u_c^{(m)}\|_{\mathrm{Lip}(\alpha)} = c^{-(p-m)}\, c^{\alpha}\, \|u^{(m)}\|_{\mathrm{Lip}(\alpha)} = \|u^{(m)}\|_{\mathrm{Lip}(\alpha)} \leq 1.$$

Hence u_c belongs to $U_p(s/|c|)$. By the same argument, if u_c belongs to $U_p(s/|c|)$, then u must belong to $U_p(s)$. This proves (2.2) according to (1.8). $\square$

Proposition 2.3. *For any two collections of independent random variables $(X_i)_{1 \leq i \leq n}$ and $(Y_i)_{1 \leq i \leq n}$ with finite absolute moments of order* $\min(p,1)$,

$$\zeta_p^{(s)}\Big(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i\Big) \leq \sum_{i=1}^n \zeta_p^{(s)}(X_i, Y_i). \tag{2.3}$$

Proof. Arguing by induction, one may assume that $n = 2$. Let μ_i and ν_i denote the distributions of X_i and Y_i, respectively, and let $\mu = \mu_1 \otimes \mu_2$, $\nu = \nu_1 \otimes \nu_2$. Given a function u in $U_p(s)$, we have

$$\begin{aligned} \int_{-\infty}^{\infty} u\, d(\mu - \nu) &= \int_{-\infty}^{\infty} \Big[\int_{-\infty}^{\infty} u(x_1 + x_2)\, d(\mu_1 - \nu_1)(x_1)\Big]\, d\mu_2(x_2) \\ &\quad + \int_{-\infty}^{\infty} \Big[\int_{-\infty}^{\infty} u(x_1 + x_2)\, d(\mu_2 - \nu_2)(x_2)\Big]\, d\nu_1(x_1). \end{aligned}$$

Since $U_p(s)$ is closed under the shifts, the inner integrals inside square brackets do not exceed $\zeta_p^{(s)}(\mu_1,\nu_1)$ and $\zeta_p^{(s)}(\mu_2,\nu_2)$. Hence

$$\int_{-\infty}^{\infty} u\, d(\mu - \nu) \leq \zeta_p^{(s)}(\mu_1,\nu_1) + \zeta_p^{(s)}(\mu_2,\nu_2).$$

It remains to take the supremum over all admissible u. $\square$

3. **Lower Bounds**

Introduce the Fourier-Stieltjes transforms (characteristic functions)

$$f(t) = \int_{-\infty}^{\infty} e^{itx}\, d\mu(x), \quad g(t) = \int_{-\infty}^{\infty} e^{itx}\, d\nu(x) \quad (t \in \mathbb{R}),$$

associated to probability measures μ and ν in $\mathfrak{P}_{\min(p,1)}$. The Zolotarev-type distances may be bounded from below in terms of these functions as follows.

Proposition 3.1. *For all $t \in \mathbb{R}$ and $s > 0$,*

$$\zeta_p^{(s)}(\mu,\nu) \ \geq \ \sup_{t \neq 0} \frac{|f(t)-g(t)|}{2\,|t|^p}, \qquad 0 < p \leq 1, \tag{3.1}$$

$$\zeta_p^{(s)}(\mu,\nu) \ \geq \ \sup_{t \neq 0} \frac{|f(t)-g(t)|}{2A_p^{(s)}(t)}, \qquad p > 1, \tag{3.2}$$

where

$$A_p^{(s)}(t) = \max\Big\{\frac{|t|^m}{s^\alpha}, \frac{|t|}{s^{p-1}}, |t|^p\Big\}.$$

Therefore, $\zeta_p^{(s)}(\mu,\nu) = 0$ implies that $f(t) = g(t)$ for all $t \in \mathbb{R}$, and thus $\mu = \nu$.

Letting $s \to \infty$ in (3.1)-(3.2), we obtain a simpler lower bound on the Zolotarev distance

$$\zeta_p(\mu,\nu) \geq \sup_{t \neq 0} \frac{|f(t)-g(t)|}{2\,|t|^p}, \quad p > 0,$$

which makes sense when μ and ν have equal moments up to order m (in the case $p > 1$).

Proof. Let $t \neq 0$. If $0 < p \leq 1$, the function $u_t(x) = \frac{1}{A}\cos(tx)$ has a Lipschitz semi-norm $\|u\|_{\mathrm{Lip}(p)} \leq 1$, if $A \geq |t|^p$. Choosing $A = |t|^p$, we obtain that

$$\frac{1}{A}\,|\operatorname{Re}(f(t)-g(t))| = \Big|\int_{-\infty}^{\infty} u_t\, d(\mu-\nu)\Big| \leq \zeta_p^{(s)}(\mu,\nu). \tag{3.3}$$

Similarly

$$\frac{1}{A}\,|\operatorname{Im}(f(t)-g(t))| = \Big|\int_{-\infty}^{\infty} u_t\, d(\mu-\nu)\Big| \leq \zeta_p^{(s)}(\mu,\nu), \tag{3.4}$$

thus proving (3.1). If $p > 1$, testing the conditions (1.2) and (1.9), we see that u_t belongs to $\mathfrak{P}_p$, if $A \geq A_p^{(s)}(t)$. Applying (3.3)-(3.4) with $A = A_p^{(s)}(t)$, we arrive at (3.2). $\square$

There is a similar relationship of $\zeta_p^{(s)}(\mu,\nu)$ with the Lévy distance $L(\mu,\nu)$, which is defined as the infimum over all $\varepsilon \geq 0$ such that

$$G(x-\varepsilon) - \varepsilon \leq F(x) \leq G(x+\varepsilon) + \varepsilon \quad \text{for all } x \in \mathbb{R}. \tag{3.5}$$

Here

$$F(x) = \mathbb{P}\{X \leq x\} = \mu\big((-\infty,x]\big), \quad G(x) = \mathbb{P}\{Y \leq x\} = \nu\big((-\infty,x]\big)$$

denote the distribution functions of random variables X and Y with distributions μ and ν.

Proposition 3.2. *We have*

$$\zeta_p^{(s)}(\mu,\nu) \geq c_p^{(s)} L^{p+1}(\mu,\nu) \tag{3.6}$$

with constants $c_p^{(s)}$ depending on (p,s) only. One may take $c_p^{(s)} = 1$ for $0 < p \leq 1$.

Proof. Fix $x \in \mathbb{R}$, $\varepsilon > 0$, and consider a non-decreasing function $u : \mathbb{R} \to \mathbb{R}$ such that $u(y) = 0$ for $y \leq x$ and $u(y) = 1$ for $y \geq x + \varepsilon$. If $\frac{1}{A} u$ belongs to $U_p(s)$ for a number $A > 0$, then

$$\begin{aligned} A\,\zeta_p^{(s)}(\mu,\nu) \ &\geq \ \int_{-\infty}^{\infty} u\,d(\nu-\mu) \ \geq \ \int_{x+\varepsilon}^{\infty} u(y)\,d\nu(y) - \int_{x}^{\infty} u(y)\,d\mu(y) \\ &\geq \ (1 - G(x+\varepsilon)) - (1 - F(x)). \end{aligned}$$

Hence $F(x) \leq G(x+\varepsilon) + A\,\zeta_p^{(s)}(\mu,\nu)$. One can interchange F and G, thus leading to

$$G(x-\varepsilon) - A\,\zeta_p^{(s)}(\mu,\nu) \leq F(x) \leq G(x+\varepsilon) + A\,\zeta_p^{(s)}(\mu,\nu). \tag{3.7}$$

If $0 < p \leq 1$, we choose the function

$$u(y) = \frac{1}{\varepsilon^p}\,(x-y)^p, \quad x \leq y \leq x+\varepsilon,$$

in which case $\|u\|_{\mathrm{Lip}(p)} \leq \varepsilon^{-p}$. The condition (1.2) is then fulfilled for the function u/A with $A = \varepsilon^{-p}$. Hence (3.7) holds with $A = \varepsilon^{-p}$, and we obtain (3.5) with $\varepsilon = (\zeta_p^{(s)}(\mu,\nu))^{\frac{1}{p+1}}$. As a result, we obtain (3.6) with $c_p^{(s)} = 1$.

In the case $p > 1$, let us take a C^∞-smooth non-decreasing function $h : \mathbb{R} \to [0,1]$ such that $h(t) = 0$ for $t \leq 0$, $h(t) = 1$ for $t \geq 1$, and define

$$u(y) = h\Big(\frac{y-x}{\varepsilon}\Big), \quad y \in \mathbb{R}.$$

Putting

$$B_p = \max_{1 \leq k \leq m+1}\ \max_{0 \leq t \leq 1} |h^{(k)}(t)|,$$

we have that the derivatives

$$u^{(k)}(y) = \frac{1}{\varepsilon^k}\,h^{(k)}\Big(\frac{y-x}{\varepsilon}\Big), \quad 1 \leq k \leq m,$$

are bounded in absolute value by $B_p\,\varepsilon^{-p}$ for $0 < \varepsilon \leq 1$. Hence the condition (1.9) is fulfilled for the function u/A, if

$$A \geq B_p\,\varepsilon^{-p}\ \max\Big\{\frac{1}{s^{p-1}}, \frac{1}{s^\alpha}\Big\}.$$

Turning to the condition (1.2), let us note that

$$\begin{aligned} \|u^{(m)}\|_{\mathrm{Lip}(\alpha)} \ &= \ \frac{1}{\varepsilon^m}\,\|h^{(m)}((y-x)/\varepsilon)\|_{\mathrm{Lip}(\alpha)} \\ &= \ \frac{1}{\varepsilon^m}\,\frac{1}{\varepsilon^\alpha}\,\|h^{(m)}\|_{\mathrm{Lip}(\alpha)} \ \leq \ \frac{1}{\varepsilon^p}\,\|h^{(m)}\|_{\mathrm{Lip}(1)} \ \leq \ B_p\,\varepsilon^{-p}. \end{aligned}$$

Hence the condition (1.2) is fulfilled for u/A, if $A \geq B_p\,\varepsilon^{-p}$. Consequently, this function belongs to $U_p(s)$ for

$$A = B_p\,\varepsilon^{-p}\ \max\Big\{\frac{1}{s^{p-1}}, \frac{1}{s^\alpha}, 1\Big\},$$

in which case we obtain (3.7). Hence we also obtain (3.5) by requiring that

$$\varepsilon^{p+1} = B_p \, \max\Big\{\frac{1}{s^{p-1}}, \frac{1}{s^\alpha}, 1\Big\}\, \zeta_p^{(s)}(\mu,\nu). \tag{3.8}$$

If the right-hand side does not exceed 1, we then arrive at the desired inequality (3.6) with

$$c_p^{(s)} = \Big(B_p \, \max\Big\{\frac{1}{s^{p-1}}, \frac{1}{s^\alpha}, 1\Big\}\Big)^{-\frac{1}{p+1}}.$$

In the other case where the quantity in (3.8) is greater than 1, there is nothing to prove since $L(\mu,\nu) \le 1$. $\square$

4. **Upper Bounds in Terms of Weighted Total Variation**

We now turn to upper bounds for the Zolotarev-type distances. Write $p = m+\alpha$ with $m \ge 0$ integer and $0 < \alpha \le 1$.

Proposition 4.1. *Let $s > 0$. If the random variables X and Y have distributions μ and ν in $\mathfrak{P}_p$, then*

$$\zeta_p^{(s)}(X,Y) \le \sum_{k=1}^m \frac{s^{p-k}}{k!}\, |\,\mathbb{E}X^k - \mathbb{E}Y^k| + \frac{\Gamma(\alpha+1)}{\Gamma(p+1)} \int_{-\infty}^\infty |x|^p \, |\mu-\nu|(dx). \tag{4.1}$$

In particular, if X and Y have equal moments up to order m, then

$$\zeta_p(X,Y) \le \frac{\Gamma(\alpha+1)}{\Gamma(p+1)} \int_{-\infty}^\infty |x|^p \, |\mu-\nu|(dx). \tag{4.2}$$

Here $|\mu-\nu|$ denotes the total variation measure for the signed measure $\mu-\nu$, so that the integral in (4.1) describes the weighted total variation distance between μ and ν with weight $|x|^p$. In particular, it follows from (4.1) that

$$\zeta_p^{(s)}(X,Y) \le \sum_{k=1}^m \frac{s^{p-k}}{k!}\, |\,\mathbb{E}X^k - \mathbb{E}Y^k| + \frac{\Gamma(\alpha+1)}{\Gamma(p+1)}\, (\mathbb{E}\,|X|^p + \mathbb{E}\,|Y|^p). \tag{4.3}$$

Note that the sum in (4.1) and (4.3) is vanishing for $0 < p \le 1$.

Proof. We follow a simple argument by Zolotarev [12], where the upper bound (4.2) is derived. Let u be an arbitrary function participating in the supremum (1.8). First suppose that $0 < p \le 1$. In this case, $m = 0$, $p = \alpha$, and the class $U_p(s)$ is described by the Lipschitz condition (1.2),

$$|u(x) - u(y)| \le |x-y|^\alpha, \quad x,y \in \mathbb{R}.$$

Choosing here $y = 0$, we get

$$\Big|\int_{-\infty}^\infty u\, d(\mu-\nu)\Big| \le \int_{-\infty}^\infty |u|\, d|\mu-\nu| \le \int_{-\infty}^\infty |x|^\alpha\, d|\mu-\nu|(x).$$

Taking the supremum over all u, we obtain (4.1).

In the case $p > 1$, we apply the integral Taylor formula around zero

$$u(x) = u(0) + \sum_{k=1}^m \frac{x^k}{k!}\, u^{(k)}(0) + \frac{x^m}{(m-1)!} \int_0^1 (1-t)^{m-1} \left(u^{(m)}(tx) - u^{(m)}(0)\right) dt.$$

This gives

$$\begin{aligned}\int_{-\infty}^\infty u\, d(\mu - \nu) \ &= \ \sum_{k=1}^m \frac{u^{(k)}(0)}{k!}\, (\,\mathbb{E}X^k - \mathbb{E}Y^k) \\ &+ \frac{1}{(m-1)!} \int_0^1 (1-t)^{m-1} \bigg(\int_{-\infty}^\infty x^m \left(u^{(m)}(tx) - u^{(m)}(0)\right) d(\mu-\nu)(x) \bigg)\, dt. \qquad (4.4)\end{aligned}$$

By the Lipschitz condition (1.2), $|u^{(m)}(tx) - u^{(m)}(0)| \le |tx|^\alpha$. Hence, the inner integral in (4.4) does not exceed in absolute value

$$t^\alpha \int_{-\infty}^\infty |x|^{m+\alpha}\, d|\mu - \nu|(x).$$

As a result, using the condition (1.9), we get

$$\begin{aligned}\bigg| \int_{-\infty}^\infty u\, d(\mu-\nu) \bigg| \ &\le \ \sum_{k=1}^m \frac{s^{p-k}}{k!}\, |\, \mathbb{E}X^k - \mathbb{E}Y^k| \\ &+ \frac{1}{(m-1)!} \int_0^1 (1-t)^{m-1} t^\alpha\, dt \int_{-\infty}^\infty |x|^p\, d|\mu-\nu|(x),\end{aligned}$$

where the second last integral is equal to $\Gamma(\alpha+1)/\Gamma(p+1)$. It remains to take the supremum over all admissible functions u. $\square$

We are prepared to apply these results to the central limit theorem.

Proof of Theorem 1.5. First, we appeal to Propositions 2.2-2.3 with $c = 1/\sqrt{n}$ and with (Y_i) being independent copies of a standard normal random variable Z. By (2.2)-(2.3), we then get

$$\begin{aligned}\zeta_p^{(s)}(Z_n, Z) \ &= \ n^{-p/2}\, \zeta_p^{(s\sqrt{n})} \Big(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i \Big) \\ &\le \ n^{-p/2} \sum_{i=1}^n \zeta_p^{(s\sqrt{n})}(X_i, Y_i) \ = \ n^{-\frac{p-2}{2}}\, \zeta_p^{(s\sqrt{n})}(X, Z).\end{aligned}$$

It remains to apply (4.3) which gives

$$\zeta_p^{(s\sqrt{n})}(X, Z) \ \le \ \sum_{k=1}^m \frac{n^{\frac{p-k}{2}}}{k!}\, |\, \mathbb{E}X^k - \mathbb{E}Z^k| + \frac{\Gamma(\alpha+1)}{\Gamma(p+1)}\, (\mathbb{E}\,|X|^p + \mathbb{E}\,|Z|^p).$$

5. **Upper Bounds in Terms of Power Transport Metrics**

Turning to the transport distances, first let us give a few remarks about the one-dimensional formula (1.10). In the case of the real line, the distance W_p may be related to the distribution functions

$$F(x) = \mu\big((-\infty, x]\big), \quad G(x) = \nu\big((-\infty, x]\big) \quad (x \in \mathbb{R})$$

associated to probability measures μ and ν. In particular, the definition of W_1 is simplified to

$$W_1(\mu,\nu) = \int_{-\infty}^{\infty} |F(x) - G(x)|\, dx,$$

which is also called the mean or L^1-distance. This formula was first obtained by Dall'Aglio [4] and later rediscovered by Vallander [10]. In the case $p = 2$, there is the following representation, apparently due to del Barrio (cf. [B-L], p. 13),

$$\begin{aligned} W_2^2(\mu,\nu) \ &= \ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \big(F(x \wedge y) - G(x \vee y)\big)^+ \, dx\, dy \\ &\quad + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \big(G(x \wedge y) - F(x \vee y)\big)^+ \, dx\, dy, \end{aligned}$$

where we adopt the standard notations $x \wedge y = \min\{x,y\}$, $x \vee y = \max\{x,y\}$, $x^+ = \max\{x,0\}$ for real numbers x, y.

But, if $p \geq 1$ is arbitrary, one has to use the inverse functions $F^{-1}(t) = \inf\{x \in \mathbb{R} : F(x) \geq t\}$, which allow one to give an explicit expression

$$W_p(\mu,\nu) = \Big(\int_0^1 |F^{-1}(t) - G^{-1}(t)|^p \, dt\Big)^{1/p}. \tag{5.1}$$

The latter implies (cf. [2])

$$W_p(\mu,\nu) = \sup \int_{-\infty}^{\infty} |u(F(x)) - u(G(x))|\, dx, \tag{5.2}$$

where the supremum is taken over all smooth functions $u : \mathbb{R} \to \mathbb{R}$ such that

$$\int_{-\infty}^{\infty} |u'(t)|^q \, dt \leq 1, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

The formula (5.2) may be extended to all functions F and G of bounded total variation such that $F(-\infty) = G(-\infty) = 0$ and $F(\infty) = G(\infty) = 1$. This turns out to be useful in the study of rates in the central limit theorem with respect to W_p, where for G one takes a Chebyshev-Edgeworth approximation of the distribution of the sum of independent random variables (and then in general the function G does not correspond to a positive measure). The representation (5.2) also provides a general upper bound

$$W_p(\mu,\nu) \leq \int_{-\infty}^{\infty} |F(x) - G(x)|^{1/p} \, dx.$$

One can also mention another dual representation

$$W_p^p(\mu,\nu) = \sup \Big[\int_{-\infty}^{\infty} u \, d\mu - \int_{-\infty}^{\infty} v \, d\nu\Big],$$

where the supremum is running over all couples of μ-integrable functions u, v such that $u(x) - v(y) \leq |x - y|^p$ for all $x, y \in \mathbb{R}$ (cf. [1]).

Let us finally turn to the upper bounds (1.14)-(1.15) on $\zeta_p^{(s)}$ and ζ_p in terms of the transport distances.

Proof of Theorem 1.4. In order to bound W_p from below for the range $p > 1$ in terms of Zolotarev-type distances, one may enlarge the class $U_p(s)$ in Definition 1.1 by requiring that the condition (1.9) is fulfilled for $k = 1$ only. In this case,

$$\begin{aligned}\zeta_p^{(s)} &\leq \sup_{|u'| \leq s^{p-1}} \Big| \int_{-\infty}^{\infty} u\,d\mu - \int_{-\infty}^{\infty} u\,d\nu \Big| \\ &= s^{p-1} \sup_{|u'| \leq 1} \Big| \int_{-\infty}^{\infty} u\,d\mu - \int_{-\infty}^{\infty} u\,d\nu \Big| \;=\; s^{p-1} \zeta_1(\mu,\nu).\end{aligned}$$

On the other hand,

$$\zeta_1(\mu,\nu) = W_1(\mu,\nu) \leq W_p(\mu,\nu).$$

This leads to the relation (1.14): $\zeta_p^{(s)} \leq s^{p-1} W_p$.

Turning to the next relation, one may assume that the measures μ and ν are supported on the finite interval $[a.b]$ containing the origin, with $s = b - a$. Let $p = m + \alpha$ with an integer $m \geq 1$ and $0 < \alpha \leq 1$, and let u be a function participating in the supremum (1.8). By the moment assumption, the difference of the integrals in (1.1) will not change, if we subtract from u any polynomial of degree at most m. Hence

$$\zeta_p(\mu,\nu) = \sup \Big| \int_{-\infty}^{\infty} u\,d\mu - \int_{-\infty}^{\infty} u\,d\nu \Big|, \tag{5.3}$$

where the supremum is taken over all m times differentiable functions $u : \mathbb{R} \to \mathbb{R}$ whose m-th derivatives have Lipschitz semi-norms

$$\|u^{(m)}\|_{\mathrm{Lip}(\alpha)} \leq 1 \tag{5.4}$$

and such that

$$u^{(k)}(0) = 0, \quad k = 1, \dots, m. \tag{5.5}$$

First consider the case $1 < p \leq 2$. Since $u'(0) = 0$, for all $a \leq y < x \leq b$,

$$|u(x) - u(y)| = \Big| \int_y^x (u'(z) - u'(0))\,dz \Big| \leq \int_y^x |z|^\alpha\,dz \leq s^\alpha\,(x-y).$$

where we used (5.4) with $m = 1$. This means that $\|u\|_{\mathrm{Lip}} \leq s^\alpha$. Since $\alpha = p - 1$, we get

$$\Big| \int_{-\infty}^{\infty} u\,d\mu - \int_{-\infty}^{\infty} u\,d\nu \Big| \leq s^{p-1} \zeta_1(\mu,\nu) \leq s^{p-1} W_p(\mu,\nu).$$

By (5.3), we therefore obtain (1.14) noting that $\Gamma(p) \leq 1$ for $1 \leq p \leq 2$.

In the case $p > 2$ when $m \geq 2$, we apply the integral Taylor formula to the derivative u' up to the order $m - 1$. By the assumption (5.5), it implies that

$$u'(x) = \frac{x^{m-1}}{(m-2)!} \int_0^1 (1-t)^{m-2} \left(u^{(m)}(tx) - u^{(m)}(0)\right) dt.$$

Hence, by (5.4),

$$|u'(x)| \;\leq\; \frac{|x|^{m-1}}{(m-2)!} \int_0^1 (1-t)^{m-2}\,|tx|^\alpha\,dt \;\leq\; s^{p-1}\,\frac{\Gamma(\alpha+1)}{\Gamma(p)} \;\leq\; \frac{s^{p-1}}{\Gamma(p)}.$$

This means that $\|u\|_{\mathrm{Lip}} \leq \frac{s^{p-1}}{\Gamma(p)}$, and as before, we get

$$\Big| \int_{-\infty}^{\infty} u\,d\mu - \int_{-\infty}^{\infty} u\,d\nu \Big| \leq \frac{s^{p-1}}{\Gamma(p)}\,\zeta_1(\mu,\nu) \leq \frac{s^{p-1}}{\Gamma(p)}\,W_p(\mu,\nu).$$

By (5.3), we therefore obtain (1.15). □

6. Proof of Theorem 1.3

For the proof of Theorem 1.3, one needs to construct functions on $[0,1]$ with special properties. As before, we represent a given number $p \geq 1$ as $p = m+\alpha$ with $m \geq 0$ integer and $0 < \alpha \leq 1$.

Definition 6.1. Let us call a non-decreasing function $h : [0,1] \to [0,\infty)$ a p-function with constant $c > 0$, if it is m times differentiable, $h(0) = 0$, $h^{(k)}(x) = 0$ at $x = 0$ and $x = 1$,

$$|h^{(k)}(x)| \leq 1 \quad (0 \leq x \leq 1, \quad k = 1,\dots,m), \tag{6.1}$$

and if for all $0 \leq x \leq y \leq 1$,

$$h^{(m)}(y) - h^{(m)}(x) \leq (y-x)^\alpha, \tag{6.2}$$

$$h(y) - h(x) \geq c\,(y-x)^p. \tag{6.3}$$

If h is $m+1$ times differentiable, and the derivative $h^{(m+1)}$ is continuous, the condition (6.2) is equivalent to $|h^{(m+1)}(x)| \leq 1$ on $[0,1]$ similarly to (6.1) with $k = m+1$.

For $p = 1$ ($m = 0$), the conditions (6.1)-(6.3) are simplified to

$$c\,(y-x) \leq h(y) - h(x) \leq y - x, \qquad 0 \leq x \leq y \leq 1.$$

One may choose $h(x) = x$ with $c = 1$ (which is best possible). As for larger values of p, we employ the following assertion whose proof is postponed to the next sections (Lemma 9.1).

Lemma 6.2. *For any $p > 1$, there exists a p-function with constant $c \geq (8p)^{-p}$.*

Proof of Theorem 1.3. Suppose that both μ and ν are supported on the interval $[a,b]$ of length $s = b-a$. One may assume that these measures have smooth densities on the real line, positive on (a,b) and vanishing on the endpoints of this interval.

The increasing bijection $T(x) = G^{-1}(F(x))$ from this interval onto itself pushes forward μ to ν. Moreover, as follows from (5.1),

$$W_p^p(\mu,\nu) = \int_a^b |T(x) - x|^p\,d\mu(x).$$

Hence, if there is a function $u \in U_p(s)$ such that

$$|T(x) - x|^p \leq C\,(u(T(x)) - u(x)), \quad x \in [a,b], \tag{6.4}$$

with some constant C, then after integration of this inequality over $d\mu(x)$ we would obtain the desired relation $W_p^p(\mu,\nu) \leq C\,\zeta_p^{(s)}(\mu,\nu)$.

The required function u can be constructed with the help of any p-function h. Note that $T(a) = a$ and $T(b) = b$. Without loss of generality, assume that the equation $T(x) = x$, that is, $F(x) = G(x)$, has finitely many solutions $a = x_0 < x_1 < \dots < x_n = b$, $n = 1,2,\dots$

Let us say that the intervals $\Delta_l = (x_{l-1}, x_l)$, $l = 1,\dots,n$, is of type (A) if $T(x) > x$ for $x \in \Delta_l$, and it is of type (B) if $T(x) < x$ for $x \in \Delta_l$. Then (6.4) holds true, as long as

$$(y-x)^p \leq C\varepsilon_l\,(u(y) - u(x)), \quad x < y,\ x,y \in \Delta_l, \tag{6.5}$$

where

$$\varepsilon_l = \begin{cases} \;\;\,1, & \text{if } \Delta_l \text{ is of type } (A), \\ -1, & \text{if } \Delta_l \text{ is of type } (B). \end{cases}$$

Using a p-function h, put $u(x) = 0$ for $x \le a$,

$$u(x) \;=\; \frac{1}{2}\, h(1) \sum_{i=1}^{l-1} \varepsilon_i (x_i - x_{i-1})^p + \frac{1}{2}\, \varepsilon_l (x_l - x_{l-1})^p \, h\Big(\frac{x - x_{l-1}}{x_l - x_{l-1}}\Big) \tag{6.6}$$

for $x \in \Delta_l$, $1 \le l \le n$, and $u(x) = u(b-)$ for $x \ge b$. By Definition 6.1, u is m-times differentiable, its derivatives up to order m are vanishing outside $[a, b]$, with

$$u^{(k)}(x) \;=\; \frac{1}{2}\, \varepsilon_l (x_l - x_{l-1})^{p-k} \, h^{(k)}\Big(\frac{x - x_{l-1}}{x_l - x_{l-1}}\Big), \quad x \in \Delta_l \;\; (1 \le l \le n),$$

for $k = 1, \dots, m$. By the property (6.1), we have

$$|u^{(k)}(x)| \;\le\; \frac{1}{2}\, (x_l - x_{l-1})^{p-k} \;\le\; \frac{1}{2}\, (b-a)^{p-k}.$$

Now,

$$u^{(m)}(x) \;=\; \frac{1}{2}\, \varepsilon_l (x_l - x_{l-1})^{\alpha} \, h^{(m)}\Big(\frac{x - x_{l-1}}{x_l - x_{l-1}}\Big), \quad x \in \Delta_l \;(1 \le l \le n).$$

If $x, y \in \Delta_l$, then, by the property (6.2),

$$|u^{(m)}(x) - u^{(m)}(y)| \le \frac{1}{2}\, |x-y|^{\alpha}. \tag{6.7}$$

If these points belong to different intervals, say if $x \in \Delta_i$, $y \in \Delta_j$, $1 \le i < j \le n$, then, by (6.7), and since $u^{(m)}(x_l) = 0$ for all $l = 1, \dots, n$, we get

$$\begin{aligned} |u^{(m)}(x) - u^{(m)}(y)| \;&\le\; |u^{(m)}(x) - u^{(m)}(x_i)| + |u^{(m)}(x_{j-1}) - u^{(m)}(y)| \\ &\quad + |u^{(m)}(x_i) - u^{(m)}(x_{j-1})| \\ &\le\; \frac{1}{2}\, |x - x_i|^{\alpha} + \frac{1}{2}\, |y - x_{j-1}|^{\alpha} \;\le\; |x-y|^{\alpha}. \end{aligned}$$

This means that u belongs to $U_p(s)$.

Finally, note that

$$u(y) - u(x) = \frac{1}{2}\, \varepsilon_l\, (x_l - x_{l-1})^p \left(h\Big(\frac{y - x_{l-1}}{x_l - x_{l-1}}\Big) - h\Big(\frac{x - x_{l-1}}{x_l - x_{l-1}}\Big) \right).$$

In particular, let $x \in \Delta_l$, $y = T(x)$. If Δ_l is of type (A), then $T(x) > x$, $\varepsilon_l = 1$, so, by (6.3),

$$\begin{aligned} u(T(x)) - u(x) \;&=\; \frac{1}{2}\, (x_l - x_{l-1})^p \left(h\Big(\frac{T(x) - x_{l-1}}{x_l - x_{l-1}}\Big) - h\Big(\frac{x - x_{l-1}}{x_l - x_{l-1}}\Big) \right) \\ &\ge\; \frac{c}{2}\, (T(x) - x)^p, \end{aligned}$$

where c is the parameter of the function h. In the other case where Δ_l is of type (B), we have $T(x) < x$, $\varepsilon_l = -1$, so that once more by (6.3),

$$\begin{aligned} u(T(x)) - u(x) \;&=\; \frac{1}{2}\, (x_l - x_{l-1})^p \left(h\Big(\frac{x - x_{l-1}}{x_l - x_{l-1}}\Big) - h\Big(\frac{T(x) - x_{l-1}}{x_l - x_{l-1}}\Big) \right) \\ &\ge\; \frac{c}{2}\, (x - T(x))^p. \end{aligned}$$

In both cases

$$|T(x) - x|^p \leq \frac{2}{c}\,(u(T(x)) - u(x)), \tag{6.8}$$

which means that the required relation (6.4) is fulfilled with $C = 2/c$ on the interval Δ_l. Since x and $T(x)$ belong to one of such intervals, this relation is fulfilled on the whole interval $[a,b]$. It remains to apply Lemma 6.2. □

A similar proof also yields a quantitative variant of Theorem 1.2.

Theorem 6.3. *For all $\mu, \nu \in \mathfrak{P}_p$, $p > 1$, we have*

$$W_p(\mu,\nu) \leq\ 8p\,\big(2\zeta_p(\mu,\nu)\big)^{1/p}.$$

Proof. As before, we may assume that μ and ν are supported on a finite interval (a,b), where these measures have smooth positive densities, vanishing on the endpoints of this interval. One may also assume that the equation $F(x) = G(x)$ has finitely many solutions $a = x_0 < x_1 < \dots < x_n = b$, $n = 1,2,\dots$ The remaining argument is based on the same p-function h and the function $u(x)$ as in (6.6). This function has the m-th derivative satisfying

$$|u^{(m)}(x) - u^{(m)}(y)| \leq |x-y|^\alpha$$

for all $x,y \in \mathbb{R}$. In addition, (6.8) is fulfilled, and we obtain the relation (6.4) with $C = 2/c$. It remains to apply Lemma 6.2. □

7. Construction of p-functions

As before, we write the number $p > 1$ as $p = m + \alpha$ with $m \geq 1$ integer and $0 < \alpha \leq 1$. One may look for a good p-function in the form

$$h(x) = \frac{1}{A}\int_0^x w(z)^{p-1}\,dz, \quad 0 \leq x \leq 1 \ \ (A > 0). \tag{7.1}$$

As a preliminary step towards Lemma 6.2, we first prove:

Lemma 7.1. *Suppose that the function w has m continuous derivatives on $[0,1]$ and possesses the following two properties:*

1) *w is increasing on $[0, \frac{1}{2}]$, symmetric about $\frac{1}{2}$, and $w(0) = w(1) = 0$;*
2) *$w(t)/t$ is non-increasing in $[0,1]$ (for example, if w is concave).*

Then the function h defined in (7.1) is a p-function with constant

$$c = \frac{1}{A}\int_0^1 w(z)^{p-1}\,dz, \tag{7.2}$$

where

$$A\ =\ \max\Big\{\max_{1 \leq k \leq m-1}\big\|(w^{p-1})^{(k)}\big\|_\infty, \big\|(w^{p-1})^{(m-1)}\big\|_{\mathrm{Lip}(\alpha)}\Big\}. \tag{7.3}$$

Proof. According to (7.1), the function h has the first m derivatives

$$h^{(k)}(x) = \frac{1}{A}\,\big(w(x)^{p-1}\big)^{(k-1)}, \quad k = 1,\dots,m,$$

which are vanishing at the endpoints of $[0,1]$ and satisfies the conditions (6.1)-(6.2) with optimal constant (7.3).

Let us now see that the condition (6.3) is met with the parameter c given by (7.2). Since $w(x)$ is symmetric about the point $x = \frac{1}{2}$ and is increasing on $[0,\frac{1}{2}]$, we have that, for each fixed $t \in (0,1)$, the function $v(x) = h(x+t) - h(x)$ is defined for $0 \le x \le 1-t$ and has derivative satisfying

$$v'(x) = \frac{1}{A}\,\big(w(x+t)^{p-1} - w(x)^{p-1}\big) = 0 \Longleftrightarrow x = \frac{1-t}{2}.$$

Since $v'(0) = \frac{1}{A}\,w(t)^{p-1} > 0$, necessarily v attains maximum at $x = \frac{1-t}{2}$ and minimum at $x = 0$ or at $x = 1-t$. Thus,

$$\min_{0 \le x \le 1-t} v(x) = \min\{v(0), v(1-t)\} = \min\{h(t), h(1) - h(1-t)\}.$$

But, by the symmetry of w about the point $1/2$,

$$h(1) - h(1-t) = \frac{1}{A}\int_{1-t}^{1} w(z)^{p-1}\,dz = \frac{1}{A}\int_0^t w(z)^{p-1}\,dz = h(t),$$

so that

$$h(x+t) - h(x) \ge h(t), \quad 0 \le x \le 1-t.$$

Now, by the assumption 2), $w(tz)/t$ is non-increasing in $t \in [0,1]$ for any fixed $z \in (0,1)$, so is the function

$$\frac{1}{t^p}\,h(t) = \frac{1}{A}\int_0^1 \Big(\frac{w(tz)}{t}\Big)^{p-1} dz.$$

Hence it is minimized at $t = 1$, that is, $h(t) \ge h(1)\,t^p$. In other words, the worst situation in the inequality (6.3) corresponds to $x = 0$ and $y = 1$, so that it is fulfilled with the optimal value $c = h(1)$. $\square$

Let us simplify the assertion of Lemma 7.1 for the range $1 < p \le 2$, when $m = 1$. In this case, (7.3) yields

$$\begin{aligned} A \ = \ \|w^\alpha\|_{\mathrm{Lip}(\alpha)} \ &= \ \sup_{0 \le x < y \le 1} \frac{|w(y)^\alpha - w(x)^\alpha|}{(y-x)^\alpha} \\ &\le \ \sup_{0 \le x < y \le 1} \frac{|w(y) - w(x)|^\alpha}{(y-x)^\alpha} \ = \ \|w\|_{\mathrm{Lip}}^\alpha, \end{aligned}$$

where we used a simple inequality $(a+b)^\alpha \le a^\alpha + b^\alpha$ $(a,b \ge 0)$. That is, we arrive at:

Corollary 7.2. *Given $p = 1+\alpha$, $0 < \alpha \le 1$, suppose that the function w has a continuous derivative on $[0,1]$ and possesses the properties $1-2$ from Lemma 7.1. Then h is a p-function with constant (7.2), where*

$$A \ = \ \max_{0 \le x \le 1} |w'(x)|^\alpha. \tag{7.4}$$

In the sequel, we follow Lemma 7.1 and are interested in the construction of p-functions h whose constants c as large as possible. To see what could be a good choice, especially when p is large, let us give a simple upper bound.

Proposition 7.3. *If h is a p-function with constant c, then necessarily*

$$c \leq \frac{\Gamma(\alpha+1)}{\Gamma(p+1)} \leq \frac{1}{\Gamma(p+1)}.$$

Proof. Since $h^{(k)}(0) = 0$ for $k = 0, 1, \dots, m$, by the integral Taylor formula, we have

$$h(y) = \frac{1}{(m-1)!} \int_0^y h^{(m)}(z)(y-z)^{m-1}\, dz, \quad y \in [0,1].$$

Applying in this formula the inequalities (6.2)-(6.3) with $x = 0$, we then get

$$\begin{aligned} c \,\leq\, h(1) &= \frac{1}{(m-1)!} \int_0^1 h^{(m)}(z)(1-z)^{m-1}\, dz \\ &\leq \frac{1}{(m-1)!} \int_0^1 z^{\alpha}(1-z)^{m-1}\, dz \;=\; \frac{\Gamma(\alpha+1)}{\Gamma(p+1)}. \end{aligned}$$

It remains to note that $\Gamma(\alpha+1) \leq 1$. □

Thus, a statement that h is a p-function with $c = (Cp)^{-p}$ with some positive absolute constant C would be a good choice (modulo an exponential factor for growing p).

8. The case $w(z) = z(1-z)$, $1 < p \leq 3$

In the proof of the inequality $W_p \leq C_p\, \zeta_p^{1/p}$, E. Rio used the p-function h based on the choice $w(z) = \sin(\pi z)$. To simplify technical computations, we prefer a slightly different choice based on the formula (7.1) with the function $w(z) = z(1-z)$, i.e.

$$h(x) = \frac{1}{A} \int_0^x (z(1-z))^{p-1}\, dz, \quad 0 \leq x \leq 1, \tag{8.1}$$

where $A = A(p)$ is described in (7.3) for all $p > 1$ or in (7.4) for the range $1 < p \leq 2$.

Since w is concave on $[0,1]$, all conditions of Lemma 7.1 are fulfilled, so we only need to estimate from below the integral

$$I(p) = \int_0^1 w(z)^{p-1}\, dz$$

and to estimate from above the derivatives of w^{p-1} of order up to $m-1$ together with the Lipschitz semi-norm $\|(w^{p-1})^{(m-1)}\|_{\mathrm{Lip}(\alpha)}$. Here, $p = m + \alpha$ with $m \geq 1$ integer and $0 < \alpha \leq 1$.

Note that when $1 < p \leq 2$, we have $m = 1$ and

$$I(p) \geq \int_0^1 z(1-z)\, dz = \frac{1}{6}.$$

Applying (7.4), we get

$$A \,=\, \max_{0\le x\le 1} |w'(x)|^\alpha \,=\, \max_{0\le x\le 1} |1-2x|^\alpha \,\le\, 1.$$

Thus, $\frac{I(p)}{A(p)} \ge 1/6$, which means that (8.1) defines a p-function with constant $c = 1/6$.

Next, consider the range $2 < p \le 3$ in which case $m = 2$. Then, according to (7.3),

$$A(p) \,=\, \max\Big\{\max_{0\le x\le 1} |(w^{p-1})'(x)|, \|(w^{p-1})'\|_{\mathrm{Lip}(\alpha)}\Big\}. \tag{8.2}$$

Since

$$(w(x)^{p-1})' \,=\, (p-1)\, w(x)^\alpha\, (1-2x) \tag{8.3}$$

and $w \le 1$, it follows that

$$|(w^{p-1})'(x)| \le p-1 \le 2, \quad x \in [0,1].$$

In order to estimate the Lipschitz semi-norm in (8.2), one may use the relation

$$\|fg\|_{\mathrm{Lip}(\alpha)} \le \|f\|_{\mathrm{Lip}(\alpha)} + \|g\|_{\mathrm{Lip}(\alpha)}, \tag{8.4}$$

holding whenever $|f| \le 1$ and $|g| \le 1$. Clearly, $\|1-2x\|_{\mathrm{Lip}(\alpha)} \le \|1-2x\|_{\mathrm{Lip}} \le 2$. Also,

$$\|w^\alpha\|_{\mathrm{Lip}(\alpha)} \le \|w\|_{\mathrm{Lip}}^\alpha = \max_{0\le x\le 1} |w'(x)|^\alpha = 1.$$

Applying (8.4) in (8.3), we therefore obtain that $\|(w^{p-1})'\|_{\mathrm{Lip}(\alpha)} \le 3$. Hence, the constant in (8.2) satisfies $A(p) \le 3$. In addition,

$$I(p) \ge \int_0^1 (z(1-z))^2\, dz = \frac{1}{30}.$$

Thus, $\frac{A(p)}{I(p)} \le 90$. One can summarize.

Lemma 8.1. *The function $h(x)$ in (8.1) is a p-function with constant $c = \frac{1}{6}$ for the range $1 \le p \le 2$ and with $c = \frac{1}{90}$ for $2 < p \le 3$. In both cases, $c^{-1} < (8p)^p$.*

9. The case $w(z) = z(1-z)$, $p > 3$

Keeping notations of the previous section, now let $p > 3$ so that $p = m+\alpha$, $m \ge 3$. To bound the derivatives of w^{p-1}, we apply the well-known chain rule formula

$$\frac{d^k}{dx^k}\, u(w(x)) = k! \sum \frac{d^N u(w)}{dw^N}\Big|_{w=w(x)} \prod_{r=1}^k \frac{1}{p_r!}\Big(\frac{1}{r!}\,\frac{d^r w(x)}{dx^r}\Big)^{p_r}, \tag{9.1}$$

where $N = p_1 + \dots + p_k$, and the summation is performed over all non-negative integer solutions $(p_1, \dots, p_k)$ to the equation

$$p_1 + 2p_2 + \dots + kp_k = k.$$

Choosing $w(x) = x(1-x)$, $u(t) = t^{p-1}$, for the k-th derivative

$$y_k(x) = \frac{d^k}{dx^k}\, w(x)^{p-1} = \frac{d^k}{dx^k}\, (x(1-x))^{p-1},$$

this formula gives the expression

$$y_k(x) = k! \sum (p-1)(p-2)\dots(p-N)\, w(x)^{p-1-N} \prod_{r=1}^{k} \frac{1}{p_r!}\Big(\frac{1}{r!}\,\frac{d^r w(x)}{dx^r}\Big)^{p_r}.$$

Since $w'(x) = 1-2x$, $w''(x) = -2$, and all other derivatives are vanishing, the above is simplified to

$$y_k(x) \ = \ k! \sum_{p_1+2p_2=k} (p-1)\dots(p-(p_1+p_2))\,\frac{(-1)^{p_2}}{p_1!p_2!}\, w_{p_1,p_2}(x), \tag{9.2}$$

where the summation is performed over all non-negative integers p_1, p_2 such that $p_1+2p_2 = k$, and where

$$w_{p_1,p_2}(x) = w(x)^{p-1-(p_1+p_2)}\,(1-2x)^{p_1}.$$

We need this formula with $1 \le k \le m-1$, in which case

$$p-1-(p_1+p_2) \ge p-1-k \ge p-m \ge \alpha.$$

Since $|w_{p_1,p_2}(x)| \le 1$ for all $x \in [0,1]$, from (9.2) it follows that

$$|y_k(x)| \ \le \ k! \sum_{p_1+2p_2=k} (p-1)\dots(p-(p_1+p_2))\,\frac{1}{p_1!p_2!}. \tag{9.3}$$

In order to further estimate the right-hand side of (9.3), note that if we applied (9.1) to the different function $w(x) = x^2$ in place of $x(1-x)$, we would obtain a similar formula

$$\frac{d^k}{dx^k}\, x^{2(p-1)} \ = \ k! \sum_{p_1+2p_2=k} (p-1)\dots(p-(p_1+p_2))\, x^{2(p-1-(p_1+p_2))}\,\frac{2^{p_1}}{p_1!p_2!}\, x^{p_1}.$$

In particular, at the point $x=1$ the latter gives

$$\begin{aligned} k! \sum_{p_1+2p_2=k} (p-1)\dots(p-(p_1+p_2))\,\frac{2^{p_1}}{p_1!p_2!} \ &= \ (2(p-1))\dots(2(p-1)-(k-1)) \\ &\le \ (2(p-1))^k. \end{aligned} \tag{9.4}$$

Therefore, according to (9.3), $|y_k(x)| \ \le \ (2(p-1))^k$, so that

$$\max_{1 \le k \le m-1} \|(w^{p-1})^{(k)}\|_\infty \ = \ \max_{1 \le k \le m-1} \|y_k\|_\infty \ \le \ (2(p-1))^{m-1}. \tag{9.5}$$

We also need to consider separately the case $k = m-1$, when (9.2) becomes

$$y_{m-1}(x) \ = \ (m-1)! \sum_{p_1+2p_2=m-1} (p-1)\dots(p-(p_1+p_2))\,\frac{(-1)^{p_2}}{p_1!p_2!}\, w_{p_1,p_2}(x). \tag{9.6}$$

Recall that $w_{p_1,p_2}(x) = w(x)^{p-n}\,(1-2x)^{p_1}$ with $n = p_1+p_2+1 \le m$. If $n \le m-1$, we have

$$\begin{aligned} \|w^{p-n}\|_{\mathrm{Lip}(\alpha)} \ &\le \ \|w^{p-n}\|_{\mathrm{Lip}} \\ &= \ (p-n) \max_{0 \le x \le 1} w(x)^{p-n-1} \ = \ \frac{p-n}{4^{p-n-1}} \ < \ 1. \end{aligned}$$

In the case $n = m$, as we have already noticed,

$$\|w^{p-n}\|_{\mathrm{Lip}(\alpha)} = \|w^\alpha\|_{\mathrm{Lip}(\alpha)} \le \|w\|_{\mathrm{Lip}}^\alpha < 1.$$

In both cases, $\|w^{p-n}\|_{\mathrm{Lip}(\alpha)} < 1$. Taking into account the inequality (8.4), we get

$$\|w_{p_1,p_2}\|_{\mathrm{Lip}(\alpha)} \leq \|w^{p-n}\|_{\mathrm{Lip}(\alpha)} + \|1-2x\|_{\mathrm{Lip}(\alpha)} < 3$$

for any choice of (p_1, p_2) participating in the expression for $y_{m-1}(x)$. Hence, from (9.6) and (9.4), we get

$$\|y_{m-1}\|_{\mathrm{Lip}(\alpha)} \ \leq \ 3\,(2(p-1))^{m-1}.$$

Comparing this bound with (9.5), for the constant $A = A(p)$ in (7.3), we thus obtain that

$$A(p) \ \leq \ 3\,(2(p-1))^{m-1}.$$

Finally,

$$I(p) \ \geq \ \int_0^1 (z(1-z))^m\,dz \ = \ \frac{m!^2}{(2m+1)!}.$$

Hence, by Lemma 7.1, the function $h(x) = \frac{1}{A}\int_0^x (z(1-z))^{p-1}\,dz$ is a p-function with constant c satisfying

$$c^{-1} \ \leq \ \frac{A(p)}{I(p)} \ \leq \ 3\,(2(p-1))^{m-1}\,\frac{(2m+1)!}{m!^2}.$$

By Stirling's approximation formula,

$$\frac{(2m+1)!}{m!^2} \leq (2m+1)\,\frac{4^m m^{2m} e^{-2m}\sqrt{4\pi m}\ e^{1/(24m)}}{m^{2m}e^{-2m}\cdot 2\pi m} = \frac{2m+1}{\sqrt{m}}\,4^m\,\frac{1}{\sqrt{\pi}}\ e^{1/(24m)}.$$

Hence, using $(1-\frac{1}{p})^p < 1/e$, we get

$$\begin{aligned} c^{-1} \ &\leq \ \frac{3e^{1/(24m)}}{2\sqrt{\pi}}\,8^m\,(p-1)^{m-1}\,\frac{2m+1}{\sqrt{m}} \\ &\leq \ \frac{3e^{1/(24m)}}{2\sqrt{\pi}}\,(8(p-1))^p\,\frac{2m+1}{(p-1)\sqrt{m}} \ \leq \ \frac{3e^{1/72}}{2e\sqrt{\pi}}\,\frac{2m+1}{(m-1)\sqrt{m}}\,(8p)^p. \end{aligned}$$

The sequence $\frac{2m+1}{(m-1)\sqrt{m}}$ is decreasing in $m \geq 3$, so it does not exceed $\frac{7}{2\sqrt{3}}$. But

$$\frac{3e^{1/72}}{2e\sqrt{\pi}}\,\frac{7}{2\sqrt{3}} < 0.638 < 1.$$

Hence, h is a p-function with constant c satisfying $c^{-1} < (8p)^p$. It remains to combine this statement about the range $p > 3$ with Lemma 8.1.

Lemma 9.1. *For all $p > 1$, the function $h(x) = \frac{1}{A}\int_0^x (z(1-z))^{p-1}\,dz$ is a p-function with constant $c^{-1} \leq (8p)^p$.*

References

[1] Bogachev, V. I.; Kolesnikov, A. V. The Monge-Kantorovich problem: achievements, connections, and prospects. Uspekhi Mat. Nauk 67 (2012), no. 5, 3–110. Russian Math. Surveys 67 (2012), no. 5, 785–890.

[2] Bobkov, S. G. Berry-Esseen bounds and Edgeworth expansions in the central limit theorem for transport distances. Probab. Theory Related Fields 170 (2018), no. 1–2, 229–262.

[3] Bobkov, S. G.; Ledoux, M. One-dimensional empirical measures, order statistics, and Kantorovich transport distances. Mem. Amer. Math. Soc. 261 (2019), no. 1259, v+126 pp.

[4] Dall'Aglio, G. Sugli estremi dei momenti delle funzioni di ripartizione doppia (Italian). Ann. Scuoloa Norm. Sup. Pisa (3) 10 (1956), 35–74.

[5] Dudley, R. M. Real analysis and probability. The Wadsworth & Brooks/Cole Mathematics Series, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1989.

[6] Rio, E. Minimal distances and ideal distances. C. R. Acad. Sci. Paris Sér. I Math. 326 (1998), no. 9, 1127–1130.

[7] Rio, E. Upper bounds for minimal distances in the central limit theorem. Ann. Inst. Henri Poincaré Probab. Stat. 45(3) (2009), 802–817.

[8] Senatov, V. V. On a relation between Lévy-Prohorov metrics and ideal metrics. Stability problems for stochastic models (Moscow, 1982), 213–216. Lecture Notes in Math., 982 Springer-Verlag, Berlin, 1983.

[9] Senatov, V. V. An estimate of the Lévy-Prokhorov metric. Teor. Veroyatnost. i Primenen. 29 (1984), no. 1, 108–113.

[10] Vallander, S. S. Calculations of the Vasserstein distance between probability distributions on the line (Russian, with English summary), Teor. Verojatnost. i Primenen. 18 (1973), 824–827.

[11] Villani, C. Topics in optimal transportation. Grad. Stud. Math., 58 American Mathematical Society, Providence, RI, 2003, xvi+370 pp.

[12] Zolotarev, V. M. Approximation of the distributions of sums of independent random variables with values in infinite-dimensional spaces. Teor. Verojatnost. i Primenen. 21 (1976), no. 4, 741–758.

[13] Zolotarev, V. M. Ideal metrics in the problem of approximating the distributions of sums of independent random variables. Teor. Verojatnost. i Primenen. 22 (1977), no. 3, 449–465.

[14] Zolotarev, V. M. Ideal metrics in the problems of probability theory and mathematical statistics. Austral. J. Statist. 21 (1979), no. 3, 193–208.

[15] Zolotarev, V. M. Modern theory of summation of random variables. Mod. Probab. Stat. VSP, Utrecht, 1997, x+412 pp.