Fourier analytic bounds for Zolotarev distances, and applications to empirical measures

Sergey G. Bobkov and Michel Ledoux

University of Minnesota and University of Toulouse

Abstract

Zolotarev distances between probability measures are natural analogues of the Kantorovich metric W_1 in which higher order derivatives of the determining family of functions are considered. This work develops Fourier analytic methods towards the investigation of families of Zolotarev distances on the d-dimensional torus, with applications to rates of convergence of empirical measures, in analogy with the known results for the Kantorovich metrics W_p, $p \geq 1$.

1 Introduction

In its simplest form, on the real line $\mathbb{R}$, the Zolotarev distance of integer order $p \geq 1$ between two probability measures μ and ν on the Borel sets of $\mathbb{R}$ with a p-th moment is defined as

$$\zeta_p(\mu,\nu) \,=\, \sup \bigg| \int_{\mathbb{R}} u\,d\mu - \int_{\mathbb{R}} u\,d\nu \bigg|$$

where the supremum runs over all $(p-1)$-times differentiable functions $u : \mathbb{R} \to \mathbb{R}$ such that $\|u^{(p-1)}\|_{\rm Lip} \leq 1$. Motivated by the continuity and stability of stochastic models, these metrics were introduced and discussed by Zolotarev in a series of papers in the mid 70's, including [17], [18] (cf. also [19]). These distances are of particular interest and importance in approximations of the distributions of sums of independent random variables and vectors in the central limit theorem.

For instance, it is well-known and easy to prove that ζ_3 directly produces the rate $\frac{1}{\sqrt{n}}$ in the central limit theorem for sums of independent identically distributed random variables. For a basic argument, write $\zeta_p(X,Y) = \zeta_p(\mu,\nu)$ when the random variables X and Y have distributions μ and ν respectively, and recall two remarkable properties of the ζ_p functionals. First, these distances are p-homogeneous with respect to (X,Y), i.e.

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

Secondly, they are sub-additive with respect to the convolution operation in the sense that

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

which holds for any two collections of independent random variables $X_1,\dots,X_n$ and $Y_1,\dots,Y_n$ with finite absolute moments of order p. In particular, if X_i, $i \geq 1$, represent independent copies of a random variable X with mean zero and variance one (and a third moment), an application of these properties with $p = 3$ to the normalized sum $Z_n = \frac{1}{\sqrt{n}}\,(X_1 + \cdots + X_n)$ immediately yields a Berry-Esseen-type bound

$$\zeta_3(Z_n,Z) \, \leq \, \frac{1}{\sqrt{n}}\,\zeta_3(X,Z),$$

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 $\zeta_4(Z_n,Z) \leq \frac{1}{n}\,\zeta_4(X,Z)$.

Zolotarev distances ζ_p may be defined for probability measures on $\mathbb{R}^d$, and for non-integer values of the parameter $p > 0$, as presented in the first sections of this work. The comparison of the family of Zolotarev probability distances with more classical metrics, such as the Kantorovich distances W_p, are not well understood, in particular in higher dimension. The classical Kantorovich distance W_p, $p \geq 1$, between probability measures μ and ν on $\mathbb{R}^d$ is defined as

$$W_p(\mu,\nu) \, = \, \inf_\pi \bigg(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |x-y|^p \, d\pi(x,y)\bigg)^{1/p} \tag{1.1}$$

where the infimum is taken over all probability measures π on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and ν, and with $|\cdot|$ the Euclidean norm (cf. e.g. [16], [1]). In the particular case $p = 1$, the Kantorovich duality expresses that

$$W_1(\mu,\nu) \, = \, \sup \bigg|\int_{\mathbb{R}^d} u\,d\mu - \int_{\mathbb{R}^d} u\,d\nu\bigg|$$

where the supremum runs over all 1-Lipschitz functions $u : \mathbb{R}^d \to \mathbb{R}$. As such $\zeta_1 = W_1$.

However, the relationship between ζ_p and W_p for $p > 1$ has not been clarified so far except in the one-dimensional case, where several explicit formulas for W_p in terms of distribution, and inverse distribution functions are available. Based on these representations, one main result connecting these two quantities is Rio's inequality [14]

$$W_p^p(\mu,\nu) \, \leq \, C_p^p \, \zeta_p(\mu,\nu) \tag{1.2}$$

with some constant $C_p > 0$ depending on p only. It is also shown in [14] that the optimal constant satisfies $C_p \leq Cp$ with some numerical C as long as p is an integer. This assertion about the growth of C_p remains to hold for all real numbers $p \geq 1$ [3]. The main motivating point in Rio's study was the central limit theorem and associated Berry-Esseen bounds for transport distances.

Besides a more formal investigation of Zolotarev metrics, the motivation for their study in this work is the issue about asymptotic problems for (random) empirical measures $\nu = \mu_n$

constructed on a sample $X_1, \dots, X_n$ drawn from a given probability measure μ. Various results about the behaviour of $W_p(\mu_n, \mu)$ on average as a function of the growing parameter n for samples on the real line can be found in [4].

However, such results for samples in dimension two and higher are either incomplete or require to involve different techniques. Towards this goal, Fourier analytic estimates for $W_1(\mu, \nu)$ were developed in [7] in terms of the Fourier-Stieltjes transforms of μ and ν, assuming that these measures are compactly supported on $\mathbb{R}^d$. The paper [5] (cf. [6] for corrections) continued this line of research with the extension of the previous results to the case where the metric participating in the definition of W_1 is not Euclidean.

Beyond the comparison itself with the Kantorovich distances, and a possible version of Rio's inequality (1.2) in higher dimension, this paper develops Fourier analytic bounds for Zolotarev-type metrics between probability measures on the d-dimensional torus $Q^d = (-\pi, \pi]^d$, with applications to rates of convergence of empirical measures (which may be compared to the rates under the Kantorovich metrics W_p). As a sample result, for μ on Q^d,

$$\mathbb{E}\big(\zeta_p^*(\mu_n, \mu)\big) \le C_{p,d} \, \frac{\log n}{n^{p/d}}$$

whenever $0 < p \le \frac{d}{2}$, where ζ_p^* is a modified periodic Zolotarev distance, closely related to ζ_p itself. In this range, the rates are comparable to the ones for Kantorovich distances [8], [10], showing the relevance of Zolotarev's metrics in this regard. When $p > \frac{d}{2}$, the optimal rate is of the order of $\frac{1}{\sqrt{n}}$, bigger than the ones in Kantorovich metrics (and perhaps in favour of a multi-dimensional version of (1.2)).

Turning to the content of the paper, the first sections 2–7 are devoted to the detailed description of families of Zolotarev metrics, and to their comparison. Sections 8–9 develop a first family of Fourier analytic estimates on Zolotarev distances, which are then illustrated towards rates of convergence of empirical measures when $p > \frac{d}{2}$ in Section 11. Before, Section 10 explains, for completeness, the difference between integer and fractional values of the parameter p in the definition of the Zolotarev distances, a fact deeply connected with embeddings of fractional Sobolev spaces. To address rates of empirical measures when $p \le \frac{d}{2}$ requires to investigate Fourier analytic bounds after suitable smoothings of the underlying probability measures, a technical task expanded, in a first approach, in Sections 12 and 13. Further sharpenings based on improved smoothing arguments with respect to signed measures are developed in Sections 14–17. The application to rates of convergence of empirical measures when $p \le \frac{d}{2}$ is finally presented in the last Section 18.

Throughout the paper, the parameter $p > 0$ of the Zolotarev distances will always be represented as $p = k + \alpha$ with an integer $k \ge 0$ and α real, $0 < \alpha \le 1$. Also $|x| = \big(\sum_{i=1}^d x_i^2\big)^{1/2}$ denotes the Euclidean norm of $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. $0 < c_{p,d}, C_{p,d} < \infty$ will denote constants, possibly changing from line to line, only depending on p and d.

2 Zolotarev distance ζ_p on Euclidean space

Zolotarev metrics were initially considered for (Borel) probability measures on an arbitrary Banach space using the notion of Fréchet derivatives. In the case of the Euclidean space

$\mathbb{R}^d$, such derivatives may be related to the usual partial derivatives, in terms of which the Zolotarev distances ζ_p can be defined within (p,d)-dependent factors as follows.

Denote by $\mathfrak{P}_p(\mathbb{R}^d)$, $p > 0$, the collection of all probability measures μ on the Borel sets of $\mathbb{R}^d$ such that $\int_{\mathbb{R}^d} |x|^p \, d\mu(x) < \infty$.

Definition 2.1 (Zolotarev distance ζ_p)**.** Let $p = k + \alpha > 0$. Given two probability measures μ and ν in $\mathfrak{P}_p(\mathbb{R}^d)$, the Zolotarev distance $\zeta_p(\mu,\nu)$ of order p between μ and ν is defined as

$$\zeta_p(\mu,\nu) \,=\, \sup \bigg| \int_{\mathbb{R}^d} u \, d\mu - \int_{\mathbb{R}^d} u \, d\nu \bigg| \tag{2.1}$$

where the supremum is taken over all k-times differentiable functions $u : \mathbb{R}^d \to \mathbb{R}$ whose partial derivatives of order k have $\mathrm{Lip}(\alpha)$-semi-norms at most 1 with respect to every coordinate.

Introducing the partial derivatives

$$D^\beta u \,=\, \frac{\partial^{|\beta|}}{\partial x_1^{k_1} \cdots \partial x_d^{k_d}} \, u$$

where $\beta = (k_1, \dots, k_d)$ is a multi-index such that $|\beta| = k_1 + \dots + k_d$, the Lipschitz property in Definition 2.1 expresses that the partial derivatives of order $|\beta| = k$ satisfy

$$|D^\beta u(x) - D^\beta u(y)| \,\le\, |x_i - y_i|^\alpha, \quad i = 1, \dots, d, \tag{2.2}$$

for all $x = (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_d)$ and $y = (x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_d)$ in $\mathbb{R}^d$. If p is an integer, this Lipschitz condition may be replaced by the requirement that the functions u participating in (2.1) are p-times differentiable and have all partial derivatives of order p bounded by 1 in absolute value.

The condition (2.2) ensures that $|u(x)| \le c\,(1 + |x|^p)$ for all $x \in \mathbb{R}^d$ with some constant $c > 0$ depending on u, so that the integrals in (2.1) are well-defined and finite. Nevertheless, for the finiteness of the supremum in (2.1) in the case $p > 1$, it is necessary to require that the measures μ and ν have equal mixed moments up to order k, that is,

$$\int_{\mathbb{R}^d} x_1^{k_1} \cdots x_d^{k_d} \, d\mu(x_1, \dots, x_d) \,=\, \int_{\mathbb{R}^d} x_1^{k_1} \cdots x_d^{k_d} \, d\nu(x_1, \dots, x_d) \tag{2.3}$$

for any collection of non-negative integers $k_1, \dots, k_d$ such that $k_1 + \dots + k_d \le k$. Indeed, fix such a collection of integers, and consider for example the functions of the form

$$u_a(x) \,=\, a\, x_1^{k_1} \cdots x_d^{k_d}, \quad x = (x_1, \dots, x_d) \in \mathbb{R}^d,$$

with parameter $a \in \mathbb{R}$. If a multi-index β has length $|\beta| = k$, then $D^\beta u_a(x) - D^\beta u_a(y) = 0$ for all $x, y \in \mathbb{R}^d$, so that the inequality (2.2) is satisfied for any a, and thus all u_a contribute in the supremum (2.1). Moreover, the difference of the integrals in (2.1) for $u = u_a$ is equal to the difference between the integrals in (2.3) multiplied by a. Since $a \in \mathbb{R}$ is arbitrary, the condition (2.3) is necessary for the finiteness of the supremum in (2.1).

As will be clarified in Remark 2.4 below, (2.3) is also sufficient for the finiteness of $\zeta_p(\mu,\nu)$.

It makes sense to generalize Definition 2.1 by assuming that μ and ν are signed measures on $\mathbb{R}^d$ with equal total masses. Equivalently, consider the collection of all signed measures λ on $\mathbb{R}^d$ such that $\lambda(\mathbb{R}^d) = 0$ and $\int_{\mathbb{R}^d} |x|^p \, d|\lambda|(x) < \infty$, where $|\lambda|$ denotes the variation measure on the real line associated to λ. For such measures, set

$$\|\lambda\|_{\zeta_p} \,=\, \sup \bigg| \int_{\mathbb{R}^d} u \, d\lambda \bigg|$$

where the supremum is taken over all functions u as in Definition 2.1 This quantity may be called the Zolotarev norm of order p in view of the relation $\zeta_p(\mu,\nu) = \|\mu - \nu\|_{\zeta_p}$ (this terminology is accepted in the case $p = 1$). As discussed above, the value $\|\lambda\|_{\zeta_p}$ is finite if and only if

$$\int_{\mathbb{R}^d} x_1^{k_1} \cdots x_d^{k_d} \, d\lambda(x_1,\dots,x_d) \,=\, 0$$

for any collection of non-negative integers $k_1,\dots,k_d$ such that $k_1 + \dots + k_d \le k$. This generalization is motivated by smoothing inequalities toward bounding of $\zeta_p(\mu,\nu)$ with the help of measures which are not necessarily positive as explained later on in this work.

Within d-dependent factors, the condition (2.2) can also be stated as a Lipschitz property of partial derivatives with respect to the Euclidean distance on $\mathbb{R}^d$. The following elementary assertion will be much useful in further considerations.

Lemma 2.2. *Let* $0 < \alpha \le 1$. *A function* u *on* $\mathbb{R}^d$ *has a* $\mathrm{Lip}(\alpha)$*-semi-norm at most* 1 *with respect to every coordinate if and only if it has a Lipschitz semi-norm* $\|u\|_{\mathrm{Lip}} \le 1$ *with respect to the metric on* $\mathbb{R}^d$ *defined by*

$$\rho_\alpha(x,y) \,=\, \sum_{i=1}^d |x_i - y_i|^\alpha, \quad x = (x_1,\dots,x_d),\ y = (y_1,\dots,y_d) \in \mathbb{R}^d. \tag{2.4}$$

In this case,

$$|u(x) - u(y)| \,\le\, \sum_{i=1}^d |x_i - y_i|^\alpha \,\le\, d^{1-\frac{\alpha}{2}} \, |x-y|^\alpha \quad (\le d\,|x-y|^\alpha). \tag{2.5}$$

The improved dependence $d^{1-\frac{\alpha}{2}}$ will be mostly ignored in the further developments, and only the simple bound d will be retained.

Proof. Without loss of generality, let $y = 0$ and $x_i \ge 0$, $i = 1,\dots,n$. Write

$$\begin{aligned} u(x) - u(0) \,=\, &\big(u(x_1,0,\dots,0) - u(0,0,\dots,0)\big) + \big(u(x_1,x_2,0,\dots,0) - u(x_1,0,\dots,0)\big) \\ &+ \dots + \big(u(x_1,\dots,x_{d-1},x_d) - u(x_1,\dots,x_{d-1},0)\big). \end{aligned}$$

If the functions $x_i \to u(x)$ have $\mathrm{Lip}(\alpha)$-semi-norm at most 1 with arbitrary fixed $(x_j)_{j \ne i}$, $i = 1,\dots,d$, then

$$|u(x) - u(0)| \,\le\, x_1^\alpha + x_2^\alpha + \dots + x_d^\alpha \,\le\, d^{1-\frac{\alpha}{2}} \, |x|^\alpha$$

by Hölder's inequality, which is (2.5). Conversely, the first inequality in (2.5) immediately implies the Lipschitz property of u with respect to every coordinate. $\square$

It will be of significant importance in the further developments to use smooth functions in Definition 2.1. This is the content of the next lemma.

Lemma 2.3. *In Definition 2.1, the functions u in the supremum* (2.1) *may be assumed to be C^∞-smooth.*

Proof. By (2.2) and the integral Taylor formula, any function u participating in (2.1) has partial derivatives up to order k satisfying a pointwise bound

$$\big|D^\beta u(x)\big| \,\leq\, c\,\big(1+|x|\big)^{p-|\beta|}, \quad x \in \mathbb{R}^d, \ \ 0 \leq |\beta| \leq k, \tag{2.6}$$

for some constant $c > 0$ which does not depend on x (when $|\beta| = 0$, put $D^\beta u = u$). Given $\varepsilon \in (0,1]$, consider the convolution by the Gaussian kernel

$$u_\varepsilon(x) \,=\, \frac{1}{(2\pi\varepsilon)^{d/2}} \int_{\mathbb{R}^d} u(x-y)\, e^{-|y|^2/2\varepsilon}\, dy, \quad x \in \mathbb{R}^d,$$

which represents a C^∞-smooth function. Thanks to (2.6), differentiation under the integral sign leads to a similar representation for the partial derivatives

$$D^\beta u_\varepsilon(x) \,=\, \frac{1}{(2\pi\varepsilon)^{d/2}} \int_{\mathbb{R}^d} D^\beta u(x-y)\, e^{-|y|^2/2\varepsilon}\, dy.$$

All $D^\beta u_\varepsilon$ with $|\beta| = k$ have therefore Lip(α)-semi-norms at most 1 with respect to every coordinate, so that the function u_ε is part of the supremum in (2.1). The Lipschitz condition (2.2) again ensures that $|u_\varepsilon(x)| \leq c\,(1+|x|)^p$. As an application of the Lebesgue dominated convergence theorem, $u_\varepsilon(x) \to u(x)$, and also $\int u_\varepsilon\, d\mu \to \int_{\mathbb{R}^d} u\, d\mu$, $\int u_\varepsilon\, d\nu \to \int_{\mathbb{R}^d} u\, d\nu$, as $\varepsilon \to 0$, for all probability measures μ and ν in $\mathfrak{P}_p(\mathbb{R}^d)$. In particular

$$\bigg|\int_{\mathbb{R}^d} u\, d\mu - \int_{\mathbb{R}^d} u\, d\nu\bigg| \,\leq\, \sup_{0<\varepsilon\leq 1} \bigg|\int_{\mathbb{R}^d} u_\varepsilon\, d\mu - \int_{\mathbb{R}^d} u_\varepsilon\, d\nu\bigg|$$

from which the conclusion follows. □

Remark 2.4. As announced, this remark explains why the mixed-moment condition (2.3) is sufficient for the finiteness of $\zeta_p(\mu,\nu)$. Let u be a function participating in the supremum (2.1) with $p > 1$ such that $u(0) = 0$. The multi-dimensional integral Taylor formula for $u(x)$ at zero indicates that

$$u(x) \,=\, \sum_{1\leq|\beta|\leq k} \frac{D^\beta u(0)}{\beta!}\, x^\beta + \sum_{|\beta|=k} \frac{|\beta|}{\beta!}\, x^\beta \int_0^1 (1-t)^{|\beta|-1}\, \big(D^\beta u(tx) - D^\beta u(0)\big) dt, \tag{2.7}$$

where for a multi-index $\beta = (k_1,\dots,k_d)$, $\beta! = k_1!\cdots k_d!$, and $x^\beta = x_1^{k_1}\cdots x_d^{k_d}$ for $x = (x_1,\dots,x_d) \in \mathbb{R}^d$. If $|\beta| = k$, then, by Lemma 2.2, for any $t \in [0,1]$ and $x \in \mathbb{R}^d$,

$$\big|D^\beta u(tx) - D^\beta u(0)\big| \,\leq\, d\,|tx|^\alpha \,\leq\, d\,|x|^\alpha.$$

In addition, $|x^\beta| \leq |x|^k$. Hence, the second sum in (2.7) does not exceed in absolute value

$$d\,|x|^p \sum_{|\beta|=k} \frac{1}{\beta!} \,=\, \frac{d^{k+1}}{k!}\, |x|^p.$$

Given probability measures μ and ν in $\mathfrak{P}_p(\mathbb{R}^d)$ satisfying (2.3), both sides of (2.7) may be integrated over these measures, and taking then the supremum over all admissible functions u, it follows that

$$\zeta_p(\mu,\nu) \,\leq\, \frac{d^{k+1}}{k!}\bigg(\int_{\mathbb{R}^d}|x|^p d\mu(x)+\int_{\mathbb{R}^d}|x|^p d\nu(x)\bigg) \,<\, \infty$$

which is the announced claim.

3 Relationship between Zolotarev and Kantorovich distances

Zolotarev distances may be compared to the more classical Kantorovich (transport) distances W_p as recalled in the introduction.

According to the notation $p = k+\alpha$, if $0 < p \leq 1$, then $k = 0$ and $\alpha = p$. In this case the Definition 2.1 is simplified into

$$\zeta_p(\mu,\nu) \,=\, \sup\bigg|\int_{\mathbb{R}^d} u\,d\mu - \int_{\mathbb{R}^d} u\,d\nu\bigg| \tag{3.1}$$

where the supremum is taken over all functions $u : \mathbb{R}^d \to \mathbb{R}$ which have a Lip(α)-semi-norm at most 1 with respect to every coordinate. It may additionally be required that $u(0) = 0$ so that, by Lemma 2.2, $|u(x)| \leq d|x|^p$, $x \in \mathbb{R}^d$. Hence

$$\zeta_p(\mu,\nu) \,\leq\, d\,\bigg(\int_{\mathbb{R}^d}|x|^p\,d\mu(x)+\int_{\mathbb{R}^d}|x|^p\,d\nu(x)\bigg),$$

which is finite for all probability measures μ,ν in $\mathfrak{P}_p(\mathbb{R}^d)$.

The quantity (3.1) may be recognized as a particular case of the transport Kantorovich distance. With respect to a given metric ρ on $\mathbb{R}^d$, the latter is defined to be

$$W_{1,\rho}(\mu,\nu) = \inf_\pi \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \rho(x,y)\,d\pi(x,y) \tag{3.2}$$

where the infimum is taken over all probability measures π on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and ν. By the duality Kantorovich theorem applied to the metric space $(\mathbb{R}^d,\rho)$ (cf. [9], [16]...),

$$W_{1,\rho}(\mu,\nu) \,=\, \sup\bigg|\int_{\mathbb{R}^d} u\,d\mu - \int_{\mathbb{R}^d} u\,d\nu\bigg|$$

where the supremum is taken over all functions $u : \mathbb{R}^d \to \mathbb{R}$ with Lipshitz semi-norms at most 1 with respect to the metric ρ. By Lemma 2.2, the latter expression is exactly (2.1) for the metric $\rho = \rho_p = \rho_\alpha$ defined in (2.4). As a conclusion:

Proposition 3.1 (Zolotarev and Kantorovich distances when $p \leq 1$)**.** *If $0 < p \leq 1$, for all probability measures μ and ν in $\mathfrak{P}_p(\mathbb{R}^d)$,*

$$\zeta_p(\mu,\nu) \,=\, W_{1,\rho_p}(\mu,\nu).$$

Alternatively, the equivalent metric $\widetilde{\rho}_p(x,y) = |x-y|^p$, $x,y \in \mathbb{R}^d$, may be considered in (3.2). However, $\widetilde{\rho}_p$ is not a metric in the case $p > 1$, leading to the standard definition (1.1) of the Kantorovich distances W_p.

4 Modified Zolotarev distance V_p

One negative property of the Zolotarev distance $\zeta_p(\mu,\nu)$ with index $p = k+\alpha > 1$ is that it is infinite when the mixed-moment condition (2.3) is violated (which is a typical situation for empirical measures, for example). This concerns any dimension; even on the real line, Rio's inequality (1.2) is useless when μ and ν do not have equal moments up to order k. In order to avoid this moment restriction, it is convenient to modify the definition of Zolotarev distances.

Definition 4.1 (Modified Zolotarev distance V_p)**.** Let $p = k+\alpha > 0$. Given μ and ν in $\mathfrak{P}_p(\mathbb{R}^d)$, set

$$V_p(\mu,\nu) \ = \ \sup_{u\in\Lambda_{p,d}} \bigg|\int_{\mathbb{R}^d} u\,d\mu - \int_{\mathbb{R}^d} u\,d\nu\bigg| \tag{4.1}$$

where $\Lambda_{p,d}$ is the collection of all k-times differentiable functions $u : \mathbb{R}^d \to \mathbb{R}$ whose partial derivatives up to order k are bounded by 1 in absolute value and such that the partial derivatives of order k have Lip(α)-semi-norm at most 1 with respect to every coordinate.

Given u in $\Lambda_{p,d}$, if $p \geq 1$, u itself is 1-Lipschitz, and if $p < 1$, Lemma 2.2 may be applied. In any case, with $p^* = \min(p,1)$,

$$|u(x) - u(y)| \ \leq \ d\,|x-y|^{p^*}, \quad x,y \in \mathbb{R}^d. \tag{4.2}$$

Since it may additionally be assumed that $u(0) = 0$ under the supremum sign in (4.1), it follows from (4.2) that $|u(x)| \leq d\,|x|^{p^*}$ for all $x \in \mathbb{R}^d$. Hence

$$V_p(\mu,\nu) \ \leq \ d\,\bigg(\int_{\mathbb{R}^d} |x|^{p^*}\,d\mu(x) + \int_{\mathbb{R}^d} |x|^{p^*}\,d\nu(x)\bigg),$$

which is finite on $\mathfrak{P}_p(\mathbb{R}^d)$. In addition, by the very definition,

$$V_p(\mu,\nu) \ \leq \ \zeta_p(\mu,\nu) \tag{4.3}$$

with an equality sign for $0 < p \leq 1$.

Similarly to Lemma 2.3, it may be assumed in Definition 4.1 that the functions u participating in the supremum (4.1) are C^∞-smooth.

In the case $p > 2$, Definition 4.1 can be simplified by involving partial derivatives of the first and maximal admissible orders, only. Namely, define

$$\widetilde{V}_p(\mu,\nu) \ = \ \sup_{u\in\widetilde{\Lambda}_{p,d}} \bigg|\int_{\mathbb{R}^d} u\,d\mu - \int_{\mathbb{R}^d} u\,d\nu\bigg|$$

where $\widetilde{\Lambda}_{p,d}$ is the collection of all k-times differentiable functions $u : \mathbb{R}^d \to \mathbb{R}$ whose partial derivatives of the first order are bounded by 1 in absolute value and such that the partial derivatives of the order k have Lip(α)-semi-norm at most 1 with respect to every coordinate.

Proposition 4.2. *If $p > 1$, for all probability measures μ and ν in $\mathfrak{P}_p(\mathbb{R}^d)$,*

$$V_p(\mu,\nu) \ \leq \ \widetilde{V}_p(\mu,\nu) \ \leq \ C_{p,d}\,V_p(\mu,\nu) \tag{4.4}$$

where $C_{p,d} > 0$ only depends on p, d.

Proof. It is only necessary to explain the second inequality in (4.4) for $p > 2$. For simplicity, consider the one-dimensional situation and the range $2 < p \leq 3$ (in the general case, the argument is similar to the one used in the proof of Lemma 5.2 below). Suppose that a function u belongs to $\widetilde{\Lambda}_{p,1}$, that is $|u'(x)| \leq 1$ and $|u''(x) - u''(y)| \leq |x-y|^\alpha$ for all $x, y \in \mathbb{R}$. The task is to show that the second derivative u'' is bounded. Given $a < b$,

$$\bigg|\int_a^b u''(x)\,dx\bigg| \,=\, |u'(b) - u'(a)| \,\leq\, 2.$$

Hence, for some point $x_0 \in [a,b]$, $|u''(x_0)| \leq \frac{2}{b-a}$. By the Lipschitz assumption, for every $x \in [a,b]$,

$$|u''(x)| \,\leq\, |u''(x_0)| + |x - x_0|^\alpha \,\leq\, \frac{2}{b-a} + (b-a)^\alpha.$$

Choosing $a = x-1$, $b = x+1$, it follows that $|u''(x)| \leq 3$. As a consequence, $\frac{1}{3}\,u \in \Lambda_{p,1}$, so that (4.4) holds true with $C_{p,1} = 3$. □

It may be mentioned to conclude this paragraph that, using the same ideas as in his work [14], it is possible to sharpen Rio's inequality (1.2) in terms of V_p for compactly supported measures on $\mathbb{R}$, in the form

$$W_p^p(\mu,\nu) \,\leq\, C_p D\, V_p(\mu,\nu)$$

for every probability measures $\mu, \nu \in \mathfrak{P}_p(\mathbb{R})$, $p \geq 1$, supported on a finite interval $[a,b]$, where $D = \max\{1, (b-a)^{p-1}\}$ with some constant $C_p > 0$ depending on p only [3].

5 Periodic Zolotarev distance ζ_p^*

The integrals in (2.1) of Definition 2.1 of the Zolotarev metric ζ_p will be bounded if it is additionally assumed that the functions u under the supremum sign are periodic. This is due to the property that all such functions (being multiplied by (p,d)-dependent constants) belong to the class $\Lambda_{p,d}$ of the modified Zolotarev distances and thus admit the pointwise bound (4.2) up to constants. As a result, another family of Zolotarev-type distances for the class of compactly supported measures, actually equivalent to V_p, may be emphasized.

Definition 5.1 (Periodic Zolotarev distance ζ_p^*)**.** Let $p = k + \alpha > 0$. Given two probability measures μ and ν on the cube $Q^d = (-\pi,\pi]^d$, define the periodic Zolotarev distance of order $p > 0$ between these measures as

$$\zeta_p^*(\mu,\nu) \,=\, \sup_{u \in M_{p,d}} \bigg|\int_{Q^d} u\,d\mu - \int_{Q^d} u\,d\nu\bigg| \tag{5.1}$$

where $M_{p,d}$ is the collection of all 2π-periodic functions $u : \mathbb{R}^d \to \mathbb{R}$ whose partial derivatives of order k have Lip(α)-semi-norms at most 1 with respect to every coordinate.

More generally, given a signed measure λ on Q^d with total mass $\lambda(\mathbb{R}^d) = 0$, set

$$\|\lambda\|_{\zeta_p^*} \,=\, \sup_{u \in M_{p,d}} \bigg|\int_{Q^d} u\,d\lambda\bigg|. \tag{5.2}$$

Here the periodicity property means that $u(x+2\pi m) = u(x)$ for all $x \in \mathbb{R}^d$ and $m \in \mathbb{Z}^d$.

Alternatively, ζ_p^* is the Zolotarev-type metric in the space of all probability measures on the d-dimensional torus $(\mathbb{S}^1)^d = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$, identified with $Q^d = (-\pi,\pi]^d$. It can be used in order to metrize the topology of weak convergence on this space, for any fixed positive value of p.

Similarly to Lemma 2.3, the functions u in the various suprema of Definition 5.1 may be assumed to be C^∞-smooth. Moreover, it can be set that $u(0) = 0$. Then, by Lemma 5.2 below, for any $x \in Q^d$,

$$\big|u(x)\big| \,\leq\, d\,|x|^p \,\leq\, d^{3/2}\pi \quad \text{in the case } p < 1,$$

whereas

$$\big|u(x)\big| \,\leq\, d\,(2\pi)^k |x| \,\leq\, d^{3/2}\pi\,(2\pi)^k \quad \text{in the case } p \geq 1.$$

Therefore

$$\|\lambda\|_{\zeta_p^*} \,\leq\, C_{p,d}\,\|\lambda\|_{\mathrm{TV}}. \tag{5.3}$$

If $p \geq 1$ is an integer, the requirement that a 2π-periodic function u belongs to the class $M_{p,d}$ is reduced to the point-wise bound $|D^\beta u(x)| \leq 1$ for all $x \in \mathbb{R}^d$ for any multi-index with $|\beta| = p$. In fact, this property can be extended to all multi-indices of smaller lengths for any value of $p = k + \alpha > 0$.

Lemma 5.2. *Given a function u in $M_{p,d}$, its partial derivatives up to order k are bounded by $(2\pi)^k$ in absolute value. More precisely,*

$$\big|D^\beta u(x)\big| \,\leq\, (2\pi)^{k-|\beta|+1}, \quad x \in \mathbb{R}^d, \tag{5.4}$$

for any multi-index β such that $1 \leq |\beta| \leq k$. Moreover,

$$\big|u(x) - u(y)\big| \,\leq\, d\,(2\pi)^k |x-y|^{p^*}, \quad x,y \in \mathbb{R}^d, \tag{5.5}$$

where it is recalled that $p^* = \min(p,1)$.

Proof. In view of the 2π-periodicity, assume that $x,y \in Q^d = (-\pi,\pi]^d$. Let then u be 2π-periodic and k-times differentiable with $\|u^\beta\|_{\mathrm{Lip}(\alpha)} \leq 1$ with respect to every coordinate, where $\beta = (k_1,\dots,k_d)$ has length $|\beta| = k$.

If $0 < p \leq 1$, then $\alpha = p$, $k = 0$, and there is no statement about the boundedness of the derivatives of u. Moreover, (5.5) coincides with (2.5) of Lemma 2.2.

Assume next that $p > 1$, so that $k \geq 1$. Note that all partial derivatives are 2π-periodic functions. Fix a multi-index $\beta = (k_1,\dots,k_d)$ with $|\beta| = k$, and choose an index $i = 1,\dots,d$ such that $k_i \geq 1$. By the Lipschitz property of the k-th order partial derivative along the variable x_i,

$$\begin{aligned}\big|D^\beta u(x_1,\dots,x_i,\dots,x_n) - D^\beta u(x_1,\dots,y_i,\dots,x_n)\big| \\ \leq\, |x_i - y_i|^\alpha \,\leq\, (2\pi)^\alpha \,\leq\, 2\pi, \quad x_i,y_i \in (-\pi,\pi],\end{aligned} \tag{5.6}$$

for every fixed collection $(x_j)_{j\neq i}$ of points in $(-\pi,\pi]$. Since $D^{\beta-e_i}u(x)$ is 2π-periodic as a function of x_i (where e_i denotes the i-th element of the canonical basis in $\mathbb{R}^d$),

$$\int_{-\pi}^{\pi} D^\beta u(x)\,dx_i \,=\, D^{\beta-e_i}u(x)\big|_{x_i=-\pi}^{x_i=\pi} \,=\, 0.$$

This implies that there is a point $y_i \in (-\pi, \pi]$ such that $D^\beta u(x_1, \dots, y_i, \dots, x_n) = 0$. Applying (5.6) with this choice of y_i, it follows that

$$\big|D^\beta u(x_1, \dots, x_i, \dots, x_n)\big| \, \leq \, 2\pi, \quad x \in Q^d. \tag{5.7}$$

Similarly, if $k \geq 2$, let $k_j \geq 1$ in the case $j \neq i$ or $k_j \geq 2$ if $j = i$. By periodicity,

$$\int_{-\pi}^{\pi} D^{\beta - e_i} u(x)\, dx_j \, = \, D^{\beta - e_i - e_j} u(x)\big|_{x_j = -\pi}^{x_j = \pi} \, = \, 0,$$

and there is a point $y_j \in (-\pi, \pi]$ such that $D^{\beta - e_i} u(x_1, \dots, y_j, \dots, x_n) = 0$. Hence

$$D^{\beta - e_i} u(x_1, \dots, x_j, \dots, x_n) \, = \, \int_{y_j}^{x_j} D^\beta u(x_1, \dots, z_j, \dots, x_n)\, dz_j.$$

By (5.7), the integrand here is bounded by 2π in absolute value, while the interval of integration has a length at most $|x_j - y_j| \leq 2\pi$. As a consequence,

$$\big|D^{\beta - e_i} u(x_1, \dots, x_j, \dots, x_n)\big| \, \leq \, (2\pi)^2, \quad x \in Q^d.$$

Repeating the process of decreasing of the components of β for any multi-index γ of length $|\gamma| = \ell = 1, \dots, k$ such that $\gamma \leq \beta$ component-wise, it holds true that

$$\big|D^\gamma u(x)\big| \, \leq \, (2\pi)^{k - \ell + 1}, \quad x \in Q^d.$$

Since γ may be an arbitrary multi-index for a suitable β, this gives the desired claim (5.4). In particular, u has first order partial derivatives bounded by $(2\pi)^k$ in absolute value. To reach (5.5), it remains to apply Lemma 2.2 with $\alpha = 1$ to the function $v = (2\pi)^{-k} u$. The proof of Lemma 5.2 is complete. $\square$

In the rest of this work, mostly the periodic Zolotarev metric ζ_p^* will be studied. The case $0 < p \leq 1$ was actually investigated before in [5], [6], [7] where the periodic Kantorovich metric

$$\widetilde{W}_\omega(\mu, \nu) \, = \, \sup_u \bigg| \int_{Q^d} u\, d\mu - \int_{Q^d} u\, d\nu \bigg|,$$

with the supremum running over all 2π-periodic functions u on $\mathbb{R}^d$ such that $\|u\|_{\mathrm{Lip}(\rho)} \leq 1$, was considered. Here the metric ρ was of the form

$$\rho(x, y) \, = \, \omega\big(\|x - y\|\big), \quad x, y \in Q^d,$$

for a given modulus of continuity ω, where $\|z\|$ denotes the shortest Euclidean distance from z to $2\pi\mathbb{Z}^d$. The particular case of the power function $\omega(t) = t^p$ then essentially leads to ζ_p^* (with a slight difference in definitions in dimension $d \geq 2$).

6 Relationship between ζ_p^* and V_p

Modified and periodic Zolotarev distances are closely related as demonstrated by the following statement.

Proposition 6.1 (Comparison between modified and periodic Zolotarev distances)**.** *Let* $p = k + \alpha > 0$. *For all probability measures* μ *and* ν *on* Q^d*,*

$$\zeta_p^*(\mu,\nu) \,\leq\, (2\pi)^k\, V_p(\mu,\nu). \tag{6.1}$$

Moreover, if these measures are supported on the cube $[0,\pi]^d$*, then*

$$V_p(\mu,\nu) \,\leq\, C_{p,d}\, \zeta_p^*(\mu,\nu) \tag{6.2}$$

for some constant $C_{p,d} > 0$ *depending on* (p,d) *only.*

Thus, ζ_p^* and V_p are equivalent in the class of all probability measures supported on $[0,\pi]^d$.

The first inequality (6.1) is a direct consequence of Lemma 5.2. The derivation of (6.2) will be a consequence of the following technical result, which will also be needed to compare these distances with ζ_p in the next section.

Lemma 6.2. *Given* u *in* $\Lambda_{p,d}$ *with* $u(0) = 0$*, there exists a function* $\widetilde u$ *on* $\mathbb{R}^d$ *such that*

(i) $\widetilde u(x) = u(x)$ *for all* $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]^d$*;*

(ii) for some $c_{p,d} > 0$ *depending on* (p,d) *only, the function* $c_{p,d}\,\widetilde u$ *belongs to* $M_{p,d}$*.*

Proof. For functions w on $Q^d = (-\pi,\pi]^d$ of class $\mathrm{Lip}(\alpha)$, $0 < \alpha \leq 1$, introduce the norm

$$\|w\|_\alpha \,=\, \max_{x \in Q^d} \big|w(x)\big| + \|w\|_{\mathrm{Lip}(\alpha)}.$$

It is well-known and readily verified that the space of all functions with finite $\|\cdot\|_\alpha$-norm represents an algebra satisfying the multiplicative relation: For functions w_1, w_2,

$$\|w_1 w_2\|_\alpha \,\leq\, \|w_1\|_\alpha \|w_2\|_\alpha. \tag{6.3}$$

As a first step, all partial derivatives of $u \in \Lambda_{p,d}$ up to order k have bounded $\|\cdot\|_\alpha$-norms

$$\|D^\beta u\|_\alpha \,\leq\, C_{p,d}, \quad 0 \leq |\beta| \leq k. \tag{6.4}$$

Namely, if $p < 1$, then $\|u\|_{\mathrm{Lip}(\alpha)} \leq d$ by Lemma 2.2, and therefore

$$\big|u(x)\big| \,\leq\, d\,|x| \,\leq\, d^{3/2}\pi, \quad x \in Q^d.$$

Hence $\|u\|_\alpha \leq d + d^{3/2}\pi$, thus proving (6.4). In the case $p \geq 1$, by the assumption,

$$\max_{x \in Q^d} |D^\beta u(x)| \,\leq\, 1 \tag{6.5}$$

for any multi-index with $1 \leq |\beta| \leq k$. Moreover, since $\|D^\beta u\|_{\mathrm{Lip}(\alpha)} \leq 1$ along every coordinate $i = 1,\dots,d$, when $|\beta| = k$, again by Lemma 2.2,

$$\|D^\beta u\|_{\mathrm{Lip}(\alpha)} \,\leq d\,, \quad |\beta| = k. \tag{6.6}$$

Thus $\|D^\beta u\|_\alpha \le d+1$ leading to (6.4) in this case.

It is now necessary to investigate (6.6) for any $0 \le |\beta| \le k$. By the multi-dimensional integral Taylor formula at zero as in (2.7),

$$\begin{aligned} u(x) \,&=\, u(0) + \sum_{1\le|\beta|\le k} \frac{D^\beta u(0)}{\beta!}\, x^\beta \\ &\quad + \sum_{|\beta|=k} \frac{|\beta|}{\beta!}\, x^\beta \int_0^1 (1-t)^{|\beta|-1} \big(D^\beta u(tx) - D^\beta u(0)\big)\,dt. \end{aligned} \tag{6.7}$$

If $|\beta| = k$, then, by (6.6), for any $t \in [0,1]$ and $x \in Q^d$,

$$\big|D^\beta u(tx) - D^\beta u(0)\big| \,\le\, d\,|tx|^\alpha \,\le\, d^{3/2}\pi.$$

Hence, using also (6.5) at $x = 0$ together with the assumption $u(0) = 0$, the Taylor formula (6.7) yields

$$\big|u(x)\big| \,\le\, \sum_{1\le|\beta|\le k} \frac{1}{\beta!}\,\pi^{|\beta|} + d^{3/2}\pi \sum_{|\beta|=k} \frac{1}{\beta!}\,\pi^{|\beta|}, \quad x \in Q^d,$$

which is bounded by a (p,d)-dependent constant. This extends (6.5) to the case $|\beta| = 0$.

Now, let $1 \le |\beta| \le k-1$. The Taylor formula (6.7) may be applied to $D^\beta u$ in place of u and with $k - |\beta|$ in place of k, which yields

$$\begin{aligned} D^\beta u(x) \,&=\, D^\beta u(0) + \sum_{1\le|\gamma|\le k-|\beta|} \frac{D^{\beta+\gamma} u(0)}{\gamma!}\, x^\gamma \\ &\quad + \sum_{|\gamma|=k-|\beta|} \frac{|\gamma|}{\gamma!}\, x^\gamma \int_0^1 (1-t)^{|\gamma|-1} \big(D^{\beta+\gamma} u(tx) - D^{\beta+\gamma} u(0)\big)\,dt. \end{aligned}$$

In view of (6.5) and (6.6), this expression is also bounded on Q^d by some (p,d)-dependent constant. Similarly, by Lemma 2.2 applied with $v = D^\beta u$, and using the general relation $\|v\|_{\mathrm{Lip}_i(\alpha)} \le 2\pi\,\|v\|_{\mathrm{Lip}_i(1)}$, where the Lipschitz property is understood along the i-th coordinate, $i = 1,\dots,d$, it holds true that

$$\begin{aligned} \|D^\beta u\|_{\mathrm{Lip}(\alpha)} \,&\le\, d\, \max_{1\le i\le d} \|D^\beta u\|_{\mathrm{Lip}_i(\alpha)} \\ &\le\, 2\pi d\, \max_{1\le i\le d} \|D^\beta u\|_{\mathrm{Lip}_i(1)} \\ &\le\, (2\pi)^2 d\, \max_{1\le i\le d}\, \max_{x\in Q^d} \big|D^{\beta+e_i} u(x)\big| \,\le\, C_{p,d}, \end{aligned}$$

which amounts to (6.4).

Now, take a C^∞-smooth function $\psi : (-\pi,\pi] \to [0,1]$ such that $\psi(t) = 1$ for $|t| \le \frac{\pi}{2}$ and $\psi(t) = 0$ for $\frac{3\pi}{4} \le |t| \le \pi$, and define $\psi_d(x) = \psi(x_1)\cdots\psi(x_d)$, $x = (x_1,\dots,x_d) \in Q^d$. This function also satisfies (6.4)

$$\|D^\beta \psi_d\|_\alpha \,\le\, C_{p,d}, \quad 0 \le |\beta| \le k, \tag{6.8}$$

for some $C_{p,d} > 0$. Define then $\widetilde{u}(x) = u(x)\,\psi_d(x)$, $x \in Q^d$, which satisfies the requirement (i) of the statement. As for (ii), first note that for $p < 1$, by (6.3),

$$\|\widetilde{u}\|_\alpha \,\leq\, \|u\|_\alpha\,\|\psi_d\|_\alpha \,\leq\, C_{p,d}$$

where (6.8) and (6.4) have been used with $|\beta| = 0$.

Finally, in the case $p \geq 1$, by the Newton binomial differentiation formula, for any multi-index $\beta = (k_1, \dots, k_d)$ with $|\beta| = k$,

$$\begin{aligned} D^\beta \widetilde{u}(x) \,&=\, \sum \binom{k_1}{\ell_1} \cdots \binom{k_d}{\ell_d}\, D^{(\ell_1,\dots,\ell_d)} u(x)\, D^{(k_1-\ell_1,\dots,k_d-\ell_d)} \psi_d(x) \\ &=\, \sum \binom{k_1}{\ell_1} \cdots \binom{k_d}{\ell_d}\, D^\gamma u(x)\, D^{\beta-\gamma} \psi_d(x) \end{aligned}$$

where the summation is performed over all integers $\ell_i = 0, 1, \dots, k_d$, that is, over all multi-indices $\gamma \leq \beta$. Applying the triangle inequality together with (6.3), (6.4) and (6.8), it follows that

$$\big\|D^\beta \widetilde{u}\big\|_\alpha \,\leq\, \sum \binom{k_1}{\ell_1} \cdots \binom{k_d}{\ell_d}\, \|D^\gamma u\|_\alpha\, \|D^{\beta-\gamma} \psi_d\|_\alpha \,\leq\, 2^k C_{p,d}^2.$$

Extending $\widetilde{u}$ from Q^d to the whole space $\mathbb{R}^d$ by 2π-periodicity, the extended function will be continuous and k-times differentiable with $\big\|D^\beta \widetilde{u}\big\|_\alpha \leq \widetilde{C}_{p,d} = 2^k C_{p,d}^2$ whenever $|\beta| = k$. The proof of the lemma is therefore complete (with $c_{p,d} = \frac{1}{\widetilde{C}_{p,d}}$). □

To conclude the section, it remains to prove (6.2) from Proposition 6.1. To this task, assume that μ and ν are supported on the cube $[-\frac{\pi}{2}, \frac{\pi}{2}]^d$ (since the classes of functions u in Definitions 4.1 and 5.1 are invariant under translation of the space variable). If $u \in \Lambda_{p,d}$, take a function $\widetilde{u}$ from Lemma 6.2 and then

$$\bigg|\int_{\mathbb{R}^d} u\,d\mu - \int_{\mathbb{R}^d} u\,d\nu\bigg| \,=\, \bigg|\int_{Q^d} u\,d\mu - \int_{Q^d} u\,d\nu\bigg| \,\leq\, \widetilde{C}_{p,d}\,\zeta^*(\mu,\nu).$$

Taking the supremum on the left-hand side over all $u \in \Lambda_{p,d}$ yields (6.2).

7 Relationship between ζ_p and ζ_p^*

This section is devoted to a comparison between the Zolotarev distance ζ_p of Definition 2.1 and the periodic Zolotarev distance ζ_p^* of Definition 5.1.

As a first observation, by the very definitions,

$$\zeta_p^*(\mu,\nu) \,\leq\, \zeta_p(\mu,\nu) \tag{7.1}$$

for all probability measures μ and ν with support in Q^d.

For the finiteness of $\zeta_p(\mu,\nu)$, it is necessary that all mixed moments of μ and ν coincide up to order k. This condition is therefore necessary to reverse the preceding inequality (7.1). The following analogue of Proposition 6.1 shows that ζ_p and ζ_p^* are equivalent in the class of all probability measures on $[0,\pi]^d$ under the mixed-moment condition (which is absent for $p \leq 1$).

Proposition 7.1 (Comparison between Zolotarev and periodic Zolotarev distances)**.** *Let $p = k + \alpha > 0$. If the probability measures μ and ν are supported on the cube $[0,\pi]^d$ and have equal mixed moments up to order k, then*

$$\zeta_p(\mu,\nu) \ \le \ C_{p,d}\, \zeta_p^*(\mu,\nu) \tag{7.2}$$

where $C_{p,d} > 0$ depends on (p,d) only. Equivalently, such a relation holds true for the norms $\|\lambda\|_{\zeta_p}$ and $\|\lambda\|_{\zeta_p^*}$, if the signed measure λ is supported on the cube $[0,\pi]^d$ and has zero mixed moments up to order k (including the order zero).

Proof. It may be assumed that μ and ν are supported on the cube $[-\frac{\pi}{2},\frac{\pi}{2}]^d$ (since the moment assumption is stable under translations of measures on $\mathbb{R}^d$). Let then $u : \mathbb{R}^d \to \mathbb{R}$ be a function participating in the supremum (2.1) of Definition 2.1 of the Zolotarev distance $\zeta_p(\mu,\nu)$, such that $u(0) = 0$.

If $p = k + \alpha \le 1$, and therefore $k = 0$, $\alpha = p$, the function u is supposed to have a Lipschitz semi-norm less than or equal to 1 with respect to every coordinate $i = 1,\dots,d$, so that $u \in \Lambda_{p,d}$. In this case, Lemma 6.2 may be applied to construct a function $\widetilde u$ with properties (i)-(ii), implying that

$$\Big|\int_{\mathbb{R}^d} u\,d\mu - \int_{\mathbb{R}^d} u\,d\nu\Big| \ = \ \Big|\int_{[-\frac{\pi}{2},\frac{\pi}{2}]^d} \widetilde u\,d\mu - \int_{[-\frac{\pi}{2},\frac{\pi}{2}]^d} \widetilde u\,d\nu\Big| \ \le \ C_{p,d}\,\zeta^*(\mu,\nu).$$

Taking the supremum over all u, the desired relation (7.2) is satisfied.

Now, let $p = k + \alpha > 1$, so that $k \ge 1$. Since u has continuous partial derivatives up to order k, it satisfies the integral Taylor formula (6.7) at zero. That is,

$$u(x) \ = \sum_{1 \le |\beta| \le k} \frac{D^\beta u(0)}{\beta!}\, x^\beta + \sum_{|\beta| = k} Q_\beta(x)\, x^\beta, \tag{7.3}$$

where

$$Q_\beta(x) \ = \ \frac{|\beta|}{\beta!} \int_0^1 (1-t)^{|\beta|-1} \big(D^\beta u(tx) - D^\beta u(0)\big)\,dt.$$

The first sum in (7.3) defines a polynomial $P = P(x)$ on $\mathbb{R}^d$ of degree at most k, so that $\int_{\mathbb{R}^d} P\,d\mu = \int_{\mathbb{R}^d} P\,d\nu$ by the moment assumption. Hence, the second sum defines the function

$$v(x) \ = \ u(x) - \sum_{1 \le |\beta| \le k} \frac{D^\beta u(0)}{\beta!}\, x^\beta$$

satisfying

$$\int_{\mathbb{R}^d} u\,d\mu - \int_{\mathbb{R}^d} u\,d\nu \ = \ \int_{\mathbb{R}^d} v\,d\mu - \int_{\mathbb{R}^d} v\,d\nu. \tag{7.4}$$

Necessarily, $D^{(\beta)} v(0) = 0$ for $|\beta| \le k$. In addition, for every multi-index γ with $|\gamma| = k$,

$$D^\gamma v(x) \ = \ D^\gamma u(x) - \sum_{1 \le |\beta| \le k} \frac{D^\beta u(0)}{\beta!}\, D^\gamma x^\beta \ = \ D^\gamma u(x) - D^\gamma u(0).$$

But, by the Lipschitz assumption on u in Definition 2.1 of ζ_p and by Lemma 2.2,

$$\big|D^\gamma u(x) - D^\gamma u(y)\big| \,\le\, d\,|x-y|^\alpha$$

for all $x,y \in \mathbb{R}^d$, whenever $|\gamma| = k$. Therefore

$$\big|D^\gamma v(x) - D^\gamma v(y)\big| \,\le\, d\,|x-y|^\alpha. \tag{7.5}$$

Thus, the Lip(α)-norms of partial derivatives of v of order k are bounded by d. In addition, choosing $y = 0$,

$$\big|D^\gamma v(x)\big| \,\le\, d\,|x|^\alpha \,\le\, \pi d^2, \quad x \in Q^d, \tag{7.6}$$

so that these derivatives are bounded on the cube.

This is also true for all partial derivatives of v of orders smaller than k, which is due to the property that $D^{(\beta)}v(0) = 0$ for $|\beta| \le k$. Indeed, fix β with $|\beta| < k$ and apply Taylor's formula (7.3) to $D^\beta v$ in place of u and to $k - |\beta|$ in place of k. This gives

$$D^\beta v(x) \,=\, \sum_{0 \le |\gamma| \le k-|\beta|} \frac{D^{\beta+\gamma}v(0)}{\gamma!}\, x^\gamma + \sum_{|\gamma| = k-|\beta|} T_\gamma(x)\, x^\gamma \,=\, \sum_{|\gamma| = k-|\beta|} T_\gamma(x)\, x^\gamma$$

where

$$T_\gamma(x) \,=\, \frac{|\gamma|}{\gamma!} \int_0^1 (1-t)^{|\gamma|-1} D^{\beta+\gamma} v(tx)\, dt.$$

Since $|\beta+\gamma| = k$, (7.6) yields $|T_\gamma(x)| \le \frac{1}{\gamma!}\,\pi d^2$ and the desired bound

$$\max_{|\beta| \le k} \big|D^\beta v(x)\big| \,\le\, C_{p,d}\,, \quad x \in Q^d,$$

up to some (p,d)-dependent constant $C_{p,d} > 0$. In view of (7.5), $\frac{1}{C_{p,d}}\, v$ belongs to $\Lambda_{p,d}$.

It remains to apply Lemma 6.2 to construct a function $\widetilde{v}$ such that $\frac{1}{C_{p,d}^2}\,\widetilde{v}$ belongs to $M_{p,d}$. Since $\widetilde{v} = v$ on $[-\frac{\pi}{2}, \frac{\pi}{2}]^d$, while μ and ν are supported on this cube,

$$\bigg|\int_{\mathbb{R}^d} v\, d\mu - \int_{\mathbb{R}^d} v\, d\nu\bigg| \,=\, \bigg|\int_{[-\frac{\pi}{2},\frac{\pi}{2}]^d} \widetilde{v}\, d\mu - \int_{[-\frac{\pi}{2},\frac{\pi}{2}]^d} \widetilde{v}\, d\nu\bigg| \,\le\, C_{p,d}^2\, \zeta_p^*(\mu,\nu).$$

Recalling (7.4), it follows that $\big|\int_{\mathbb{R}^d} u\, d\mu - \int_{\mathbb{R}^d} u\, d\nu\big| \le C_{p,d}^2\, \zeta_p^*(\mu,\nu)$, from which the requested claim follows. Proposition 7.1 is established. $\square$

8 Fourier analytic bound for ζ_p^* with integer p

This section introduces the Fourier analytic tools to control the periodic Zolotarev distances. Namely, the aim is to bound $\zeta_p^*(\mu,\nu)$ in terms of the closeness of the Fourier-Stieltjes transforms

$$f_\mu(m) \,=\, \int_{Q^d} e^{im\cdot x}\, d\mu(x), \quad f_\nu(m) \,=\, \int_{Q^d} e^{im\cdot x}\, d\nu(x), \tag{8.1}$$

of μ and ν on $Q_d = (-\pi,\pi]^d$, which will be sufficient to consider for integer vectors $m \in \mathbb{Z}^d$.

The section is concerned with the case of an integer index $p \ge 1$, which is somewhat simpler.

Theorem 8.1 (Fourier analytic bound for ζ_p^* with integer p)**.** *Given two probability measures μ and ν on Q^d, for any integer $p \geq 1$,*

$$\zeta_p^*(\mu,\nu) \, \leq \, d^{\frac{p-1}{2}} \bigg(\sum_{m \neq 0} \frac{1}{|m|^{2p}} \, \big|f_\mu(m) - f_\nu(m)\big|^2\bigg)^{1/2}. \tag{8.2}$$

Equivalently, for any signed measure λ on Q^d with total mass zero,

$$\|\lambda\|_{\zeta_p^*} \, \leq \, d^{\frac{p-1}{2}} \bigg(\sum_{m \neq 0} \frac{1}{|m|^{2p}} \, \big|f_\lambda(m)\big|^2\bigg)^{1/2}.$$

Proof. The task is to estimate the difference between the integrals in (5.1) of the Definition 5.1 of the periodic Zolotarev distance. Assuming that $u \in M_{p,d}$ is C^∞-smooth, it may be expanded into an absolutely convergent Fourier series

$$u(x) \, = \sum_{m \in \mathbb{Z}^d} a_m \, e^{im \cdot x}, \quad x \in Q^d. \tag{8.3}$$

Integrating this equality over the measures μ and ν,

$$\int_{Q^d} u \, d\mu - \int_{Q^d} u \, d\nu \, = \sum_{m \in \mathbb{Z}^d} a_m \, \big(f_\mu(m) - f_\nu(m)\big). \tag{8.4}$$

By the smoothness property, $\sum_{m \in \mathbb{Z}^d} |m|^s \, |a_m| < \infty$ for any $s > 0$. Therefore, the series in (8.3) may be differentiated term by term infinitely many times. So, for the p-th order partial derivative $v_j(x) = \partial^p_{x_j^p} u(x)$ along the variable x_j, $j = 1, \dots, d$, there is a similar series expansion

$$v_j(x) \, = \, i^p \sum_{m \in \mathbb{Z}^d} m_j^p \, a_m \, e^{im \cdot x}$$

which is absolutely convergent (where $m = (m_1, \dots, m_d)$). Hence, by the Parseval identity,

$$\frac{1}{(2\pi)^d} \int_{Q^d} \big|v_j(x)\big|^2 \, dx \, = \sum_{m \in \mathbb{Z}^d} m_j^{2p} \, |a_m|^2.$$

Next, the Lipschitz condition may be used, that is $|v_j(x)| \leq 1$ for all $x \in \mathbb{R}^d$, which yields

$$\sum_{m \in \mathbb{Z}^d} m_j^{2p} \, |a_m|^2 \, \leq \, 1$$

for any $j = 1, \dots, d$. Summing over all j and using that $m_1^{2p} + \dots + m_d^{2p} \geq \frac{1}{d^{p-1}} \, |m|^{2p}$, it follows that

$$\sum_{m \in \mathbb{Z}^d} |m|^{2p} \, |a_m|^2 \, \leq \, d^{p-1}. \tag{8.5}$$

Applying then Cauchy's inequality, it holds true that

$$\bigg|\sum_{m \in \mathbb{Z}^d} a_m \, \big(f_\mu(m) - f_\nu(m)\big)\bigg|^2 \, \leq \, d^{p-1} \sum_{m \neq 0} \frac{1}{|m|^{2p}} \, \big|f_\mu(m) - f_\nu(m)\big|^2.$$

It remains to recall (8.4) and to take the supremum over all admissible functions u to complete the argument. □

With the help of Proposition 7.1, the following consequence may be stated.

Corollary 8.2 (Fourier analytic bound for ζ_p with integer p)**.** *If the probability measures μ and ν are supported on the cube $[0,\pi]^d$ and have equal mixed moments up to the order $p-1$, then*

$$\zeta_p(\mu,\nu) \,\le\, C_{p,d} \bigg(\sum_{m \neq 0} \frac{1}{|m|^{2p}}\, \big|f_\mu(m) - f_\nu(m)\big|^2\bigg)^{1/2} \tag{8.6}$$

for some $C_{p,d} > 0$ depending on (p,d) only. Equivalently, for any signed measure λ on $[0,\pi]^d$ with zero mixed moments up to the order $p-1$,

$$\|\lambda\|_{\zeta_p} \,\le\, C_{p,d} \bigg(\sum_{m \neq 0} \frac{1}{|m|^{2p}}\, \big|f_\lambda(m)\big|^2\bigg)^{1/2}.$$

Since the Fourier-Stieltjes transforms represent bounded functions on $\mathbb{Z}^d$, the series (8.2) and (8.6) are convergent for $p > \frac{d}{2}$. These series may be divergent for $p \le \frac{d}{2}$, and a smoothing argument will then be required for a further analysis. This dichotomy will be developed in forthcoming sections and their applications to empirical measures.

9 Fourier analytic bound for ζ_p^* with fractional p

This paragraph expands upon the previous one in the case of fractional values $p = k + \alpha$ with $k \ge 0$ integer and $0 < \alpha < 1$, leading to several technical issues.

Keeping the notation (8.1) for the Fourier-Stieltjes coefficients, there is a similar series expansion for the k-th order partial derivative of u along the variable x_j, $j = 1, \dots, d$, in the form

$$v_j(x) \,=\, \frac{\partial^k}{\partial x_j^k}\, u(x) \,=\, i^k \sum_{m \in \mathbb{Z}^d} m_j^k\, a_m\, e^{im\cdot x}, \quad x \in Q^d,$$

which is absolutely convergent. Here as before, $m = (m_1, \dots, m_d)$. Hence, for any $h \in \mathbb{R}$,

$$v_j(x + he_j) - v_j(x - he_j) \,=\, 2i^k \sum_{m \in \mathbb{Z}^d} m_j^k\, a_m\, \sin(m_j h)\, e^{im\cdot x},$$

where $(e_1, \dots, e_d)$ denotes the canonical basis in $\mathbb{R}^d$. By the Parseval identity,

$$\frac{1}{(2\pi)^d} \int_{Q^d} \big|v_j(x + he_j) - v_j(x - he_j)\big|^2\, dx \,=\, 4 \sum_{m \in \mathbb{Z}^d} m_j^{2k}\, |a_m|^2\, \sin^2(m_j h).$$

Now, the Lipschitz condition $|v_j(x + he_j) - v_j(x - he_j)| \le |2h|^\alpha$ yields

$$\sum_{m \in \mathbb{Z}^d} m_j^{2k}\, |a_m|^2\, \sin^2(m_j h) \,\le\, \frac{1}{4}\, |2h|^{2\alpha}.$$

Choose here $h = 2^{-\ell-1}\,\pi$ with $\ell = 1, 2, \dots$ (fixed at this point), so that

$$\sum_{m \in \mathbb{Z}^d} m_j^{2k}\, |a_m|^2\, \sin^2(2^{-\ell-1}\, m_j \pi) \,\le\, \frac{1}{4}\, (2^{-\ell}\pi)^{2\alpha}.$$

Moreover, restricting the sum to the range $|m_j| \le 2^\ell$ and using $\sin(t) \ge \frac{2}{\pi} t$ for $0 \le t \le \frac{\pi}{2}$ yields the simpler bound

$$\sum_{|m_j| \le 2^\ell} m_j^{2k+2} \, |a_m|^2 \, 4^{-\ell} \;\le\; \frac{1}{4} \, (2^{-\ell}\pi)^{2\alpha}.$$

This inequality holds true for any fixed $j = 1, \dots, d$. Summing over all j's leads to

$$\sum_{j=1}^{d} \sum_{m \in \mathbb{Z}^d} \mathbb{1}_{\{|m_j| \le 2^\ell\}} \, m_j^{2k+2} \, |a_m|^2 \, 4^{-\ell} \;\le\; \frac{d}{4} \, (2^{-\ell}\pi)^{2\alpha}.$$

The double sum on the left-hand side of the preceding inequality may be further restricted to all vectors $m \in \mathbb{Z}^d$ with $\|m\|_\infty = \max_{1 \le j \le d} |m_j| \le 2^\ell$ in which case $\mathbb{1}_{\{|m_j| \le 2^\ell\}} = 1$. Then, interchanging the order of summation, it follows that

$$\sum_{\|m\|_\infty \le 2^\ell} \left(\sum_{j=1}^{d} m_j^{2k+2} \right) |a_m|^2 \, 4^{-\ell} \;\le\; \frac{d}{4} \, (2^{-\ell}\pi)^{2\alpha}.$$

Here the inner sum is greater than or equal to $d^{-k} \, |m|^{2k+2}$. Using that $\|m\|_\infty \ge |m|$ in order to further restrict the outer sum yields

$$\sum_{|m| \le 2^\ell} |m|^{2k+2} \, |a_m|^2 \, 4^{-\ell} \;\le\; \frac{d^{k+1}}{4} \, (2^{-\ell}\pi)^{2\alpha}.$$

As one more weakening,

$$\sum_{2^{\ell-1} \le |m| < 2^\ell} |m|^{2k+2} \, |a_m|^2 \, 4^{-\ell} \;\le\; \frac{d^{k+1}}{4} \, (2^{-\ell}\pi)^{2\alpha}.$$

Under the new restrictions on m inside the sum, necessarily

$$|m|^{2k+2} \, 4^{-\ell} \;\ge\; 2^{(\ell-1)(2k+2)} \, 4^{-\ell} \;=\; 4^{k\ell} \cdot 4^{-k-1}$$

so that

$$\sum_{2^{\ell-1} \le |m| < 2^\ell} |a_m|^2 \;\le\; 4^k \pi^{2\alpha} d^{k+1} \, 4^{-p\ell}.$$

To simplify the right-hand side, it may be used that $d^{k+1} \le d^{p+1}$ and

$$4^k \pi^{2\alpha} \;=\; 4^{p-\alpha} \pi^{2\alpha} \;=\; 4^p \left(\frac{\pi^2}{4} \right)^\alpha \;<\; 4^p \, \frac{\pi^2}{4}$$

so that

$$\sum_{2^{\ell-1} \le |m| < 2^\ell} |a_m|^2 \;\le\; 4^{p-1} \pi^2 d^{p+1} \, 4^{-p\ell}.$$

Hence, by Cauchy's inequality,

$$\sum_{2^{\ell-1}\le|m|<2^\ell} |a_m|\,\big|f_\mu(m)-f_\nu(m)\big| \;\le\; 2^{(p-1)/2}d^{(p+1)/2}\pi\,\frac{b_\ell}{2^{p\ell}}\,, \tag{9.1}$$

where, for every $\ell = 1, 2, \dots$,

$$b_\ell \;=\; \bigg(\sum_{2^{\ell-1}\le|m|<2^\ell} \big|f_\mu(m)-f_\nu(m)\big|^2\bigg)^{1/2}.$$

Now, let $q : [1,\infty) \to (0,\infty)$ be a non-decreasing function such that

$$0 \;<\; C_q \;=\; \bigg(\sum_{\ell=0}^{\infty} \frac{1}{q(2^\ell)}\bigg)^{1/2} \;<\; \infty. \tag{9.2}$$

Dividing and multiplying the right-hand side of (9.1) by $\sqrt{q(2^{\ell-1})}$, this inequality may be rewritten as

$$\sum_{2^{\ell-1}\le|m|<2^\ell} |a_m|\,\big|f_\mu(m)-f_\nu(m)\big| \;\le\; \frac{2^{(p-1)/2}d^{(p+1)/2}\pi}{\sqrt{q(2^{\ell-1})}}\,\frac{\sqrt{q(2^{\ell-1})}\,b_\ell}{2^{p\ell}}\,.$$

Performing summation over all $\ell \ge 1$ and applying Cauchy's inequality once more together with the definition (9.2) yields

$$\bigg(\sum_{m\neq 0} |a_m|\,\big|f_\mu(m)-f_\nu(m)\big|\bigg)^2 \le\; 2^{p-1}d^{p+1}\pi C_q^2 \sum_{\ell=1}^{\infty}\,\sum_{2^{\ell-1}\le|m|<2^\ell} \frac{q(2^{\ell-1})}{4^{p\ell}}\,\big|f_\mu(m)-f_\nu(m)\big|^2.$$

By the monotonicity with respect to ℓ, $q(2^{\ell-1}) \le q(|m|)$ whenever $2^{\ell-1} \le |m| < 2^\ell$, and similarly $4^{p\ell} \ge |m|^{2p}$.

As a result, the following bounds (which are also applicable to the case of integer values of p) may be stated. They are the analogue of Theorem 8.1 and Corollary 8.2 with the additional weight q.

Theorem 9.1 (Fourier analytic bound for ζ_p^* with fractional p)**.** *Given probability measures μ and ν on Q^d, for any $p > 0$ and any non-decreasing weight function q satisfying* (9.2)*,*

$$\zeta_p^*(\mu,\nu) \;\le\; A\,\bigg(\sum_{m\neq 0} \frac{q(|m|)}{|m|^{2p}}\,\big|f_\mu(m)-f_\nu(m)\big|^2\bigg)^{1/2} \tag{9.3}$$

with $A = 2^{\frac{p-1}{2}}d^{\frac{p+1}{2}}\pi C_q$. Equivalently, for any signed measure λ on Q^d with total mass zero,

$$\|\lambda\|_{\zeta_p^*} \;\le\; A\,\bigg(\sum_{m\neq 0} \frac{q(|m|)}{|m|^{2p}}\,\big|f_\lambda(m)\big|^2\bigg)^{1/2}.$$

Corollary 9.2 (Fourier analytic bound for ζ_p with fractional p)**.** *Let $p = k + \alpha > 0$. If the probability measures μ and ν are supported on the cube $[0,\pi]^d$ and have equal mixed moments up to the order k, then for any non-decreasing weight function q satisfying* (9.2)*,*

$$\zeta_p(\mu,\nu) \,\le\, C_q C_{p,d} \bigg(\sum_{m \neq 0} \frac{q(|m|)}{|m|^{2p}}\, \big|f_\mu(m) - f_\nu(m)\big|^2\bigg)^{1/2} \tag{9.4}$$

for some $C_{p,d} > 0$ depending on (p,d) only. Equivalently, for any signed measure λ on $[0,\pi]^d$ with zero mixed moments up to the order k,

$$\|\lambda\|_{\zeta_p} \,\le\, C_q C_{p,d} \bigg(\sum_{m \neq 0} \frac{q(|m|)}{|m|^{2p}}\, \big|f_\lambda(m)\big|^2\bigg)^{1/2}.$$

As for (8.2) and (8.6) in Theorem 8.1 and Corollary 8.2, the series (9.3) and (9.4) are convergent for $p > \frac{d}{2}$ when choosing, for example, $q(s) = \log^2(2s)$, $s \ge 1$.

10 Comparison between fractional and integer indices

The difference between the case of integer and fractional p in Theorems 8.1 and 9.1 of the two preceding sections is of course questionable, and one may wonder whether it is possible to remove the q-weight from the inequality (9.3) in Theorem 9.1? The answer turns out to be negative, and is deeply connected with embedding problems of fractional Sobolev spaces among which are the Lipschitz classes Lip(α). The following makes evidence of this obstruction, for simplicity in dimension one in the range $0 < p < 1$.

Proposition 10.1. *If $0 < p < 1$, the inequality* (8.2) *of Theorem 8.1 may not hold with a (finite) constant $C = C_p$ in the class of all probability measures on $Q = (-\pi,\pi]$.*

Some preparation and general facts will be helpful before addressing the proof itself.

The periodic fractional order Sobolev space $W^{p,r}$ with parameters $0 < p < 1$ and $r \ge 1$ is defined as the linear space of all complex-valued periodic functions u on the real line with finite norm

$$\|u\|_{W^{p,r}} \,=\, \bigg(\frac{1}{2\pi}\int_{-\pi}^{\pi} |u(x)|^r\,dx\bigg)^{1/r} + \bigg(\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \frac{|u(x+h)-u(x)|^r}{|h|^{pr+1}}\,dx\,dh\bigg)^{1/r}.$$

These norms are increasing for the increasing parameter r (hence getting stronger). That is,

$$\|u\|_{W^{p,s}} \,\le\, \|u\|_{W^{p,r}}, \quad \text{therefore} \quad W^{p,r} \subset W^{p,s} \quad \text{for } s < r.$$

The limit $r \to \infty$ yields the spaces $W^{p,\infty}$ with norm $\|u\|_{W^{p,\infty}} = \|u\|_\infty + \|u\|_{\mathrm{Lip}(p)}$, where

$$\|u\|_{\mathrm{Lip}(p)} \,=\, \mathrm{esssup}_{x,h} \frac{|u(x+h)-u(x)|}{|h|^p} \,=\, \mathrm{esssup}_{x,y} \frac{|u(x)-u(y)|}{|x-y|^p}.$$

Thus $u \in W^{p,\infty}$ if and only if u has a finite Lipschitz semi-norm $\|u\|_{\mathrm{Lip}(p)}$.

A natural question is whether $W^{p,r} = W^{p,s}$? A negative answer was given, only recently, by Mironescu and Sickel [13], who constructed several counter-examples. Some examples are based on the Fourier series representations with lacunary coefficients and appeal to the general theory described in [15]. Below, one such counter-example is described (the second example in [13]) for the case $s = \infty$ and $r = 2$, which is closely connected to Proposition 10.1.

Namely, in the special important case $r = 2$, the corresponding fractional Sobolev norm has a simple description in terms of the coefficients a_m in the Fourier series representation

$$u(x) \,=\, \sum_{m\in\mathbb{Z}} a_m e^{imx}, \quad x \in \mathbb{R}. \tag{10.1}$$

Assuming that this series is convergent in $L^2(Q)$, for any $h \in \mathbb{R}$, there is the Fourier series representation

$$u(x+h) - u(x-h) \,=\, 2\sum_{m\in\mathbb{Z}} a_m \sin(mh)\, e^{imx},$$

so that

$$\|u\|_{W^{p,2}} \,=\, \Big(\sum_{m\in\mathbb{Z}} |a_m|^2\Big)^{1/2} + \Big(\frac{2}{\pi}\sum_{m\in\mathbb{Z}} |a_m|^2 \int_{-\pi}^{\pi} \frac{\sin^2(mh)}{|h|^{2p+1}}\,dh\Big)^{1/2}.$$

Up to the factor $2\,|m|^{2p}$ for $m \neq 0$, the integral in the second sum equals $\int_0^{|m|\pi} t^{-2p-1}\,\sin^2(t)\,dt$, which is bounded away from zero and infinity by p-dependent constants. Hence, if $a_0 = 0$,

$$\|u\|_{W^{p,2}}^2 \,\sim\, \sum_{m\neq 0} |m|^{2p}\,|a_m|^2 \tag{10.2}$$

within p-dependent factors.

Lemma 10.2. *There exists a periodic Lipschitz function u in the class* Lip(p) *for which the norm $\|u\|_{W^{p,2}}$ is infinite.*

Proof. The series

$$u(x) \,=\, \sum_{m=2^\ell,\ell\geq 1} \frac{1}{m^p}\, e^{imx} \,=\, \sum_{\ell=1}^{\infty} \frac{1}{2^{p\ell}}\, e^{i\,2^\ell x}, \quad x \in \mathbb{R},$$

is absolutely convergent and therefore defines a continuous periodic function on $Q = (-\pi, \pi]$. However, the series (10.2) is divergent, meaning that u does not belong $W^{p,2}$.

In order to explore the Lipschitz property, by the very definition of u, for all $x, h \in \mathbb{R}$,

$$\big|u(x+h) - u(x)\big| \,\leq\, \sum_{m=2^\ell} \frac{1}{m^p}\,|e^{imh} - 1|. \tag{10.3}$$

Using the inequality $|e^{it} - 1| \leq 2\,\min(1,|t|)$, $t \in \mathbb{R}$, for any $h \in (0, \frac{1}{2})$,

$$\begin{aligned} \sum_{m=2^\ell\leq 1/h} \frac{1}{m^p}\,|e^{imh} - 1| \,&\leq\, 2h \sum_{m=2^\ell\leq 1/h} m^{1-p} \\ &=\, 2h \sum_{1\leq\ell\leq\log_2(1/h)} 2^{(1-p)\ell} \,\leq\, \frac{c}{1-p}\,h^p \end{aligned}$$

for some absolute constant $c > 0$. Similarly, summing the geometric progression,

$$\sum_{m=2^\ell>1/h} \frac{1}{m^p}\,|e^{imh}-1| \;\le\; 2 \sum_{m=2^\ell>1/h} m^{-p}$$

$$= \; 2 \sum_{\ell>\log_2(1/h)} 2^{-p\ell} \;\le\; \frac{1}{1-2^{-p}}\,h^p.$$

The two bounds applied in (10.3) yield the desired property $|u(x+h)-u(x)| \le c_p h^p$, $h \ge 0$, that is $u \in \mathrm{Lip}(p)$. $\square$

Proof (of Proposition 10.1). An equivalent formulation of (8.2) is the relation

$$\|\lambda\|_{\zeta_p^*} \;\le\; C \bigg(\sum_{m\neq 0} \frac{1}{|m|^{2p}}\,\big|f_\lambda(m)\big|^2\bigg)^{1/2} \tag{10.4}$$

in terms of the Fourier-Stieltjes transform $f_\lambda(m) = \int_Q e^{imx}\,d\lambda(x)$, $m \in \mathbb{Z}$, in the class of all signed measures λ on $Q = (-\pi,\pi]$ with total mass $\lambda(Q) = 0$.

Recall that in (5.2) of Definition 5.1, the supremum may be taken over all C^∞-smooth, 2π-periodic functions u with Lipschitz semi-norm $\|u\|_{\mathrm{Lip}(p)} \le 1$. In that case, the Fourier series (10.1) is absolutely convergent, that is $\sum_{m\in\mathbb{Z}} |a_m| < \infty$. Integrating (10.1) over λ and using $\lambda(Q) = 0$ yields

$$\int_Q u\,d\lambda \;=\; \sum_{m\neq 0} a_m f_\lambda(m).$$

Hence, (10.4) takes the form

$$\sup_{\|u\|_{\mathrm{Lip}(p)}\le 1} \Big|\sum_{m\neq 0} a_m f_\lambda(m)\Big| \;\le\; C\bigg(\sum_{m\neq 0}\frac{1}{|m|^{2p}}\,|f_\lambda(m)|^2\bigg)^{1/2}.$$

Equivalently (since $f_\lambda(-m) = \overline{f}_\lambda(m)$), both sums may be restricted to positive m. So, consider the relation

$$\sup_{\|u\|_{\mathrm{Lip}(p)}\le 1} \Big|\sum_{m=1}^\infty a_m f_\lambda(m)\Big| \;\le\; C\bigg(\sum_{m=1}^\infty \frac{1}{m^{2p}}\,|f_\lambda(m)|^2\bigg)^{1/2}$$

up to the constant $C > 0$. Putting $b_m = m^{-p}\,f_\lambda(m)$, $m \ge 1$, the latter may be rewritten as

$$\sup_{\|u\|_{\mathrm{Lip}(p)}\le 1} \Big|\sum_{m=1} m^p a_m b_m\Big| \;\le\; C\bigg(\sum_{m=1}^\infty |b_m|^2\bigg)^{1/2}. \tag{10.5}$$

It should be clear that all sequences of the form b_m are dense in ℓ^2 with respect to the ℓ^2-norm, since the same is true about the sequences $(f_\lambda(m))_{m\ge 1}$. Hence, the best constant in (10.5) is given by

$$C \;=\; \sup_{\|u\|_{\mathrm{Lip}(p)}\le 1} C(u), \quad C(u) \;=\; \bigg(\sum_{m=1}^\infty m^{2p}\,|a_m|^2\bigg)^{1/2}.$$

But $C(u)$ is equivalent to the fractional norm $\|u\|_{W^{p,2}}$ as was emphasized in (10.2). Moreover, $C = \infty$ according to Lemma 10.2. All together, this concludes the proof of Proposition 10.1. $\square$

11 Application to empirical measures on the torus ($p > \frac{d}{2}$)

This section initiates the applications of the Fourier analytic bounds of the preceding sections to empirical measures and their rates of convergence. Recall the periodic Zolotarev metric ζ_p^* from Definition 5.1.

The first statement is the result of the application of Theorems 8.1 and 9.1 to the empirical measures

$$\mu_n = \frac{1}{n}\sum_{j=1}^n \delta_{X_j}, \quad \nu_n = \frac{1}{n}\sum_{j=1}^n \delta_{Y_j},$$

constructed over samples $X = (X_1, \dots, X_n)$ and $Y = (Y_1, \dots, Y_n)$ of random variables (defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$) taking values in the torus Q^d.

Theorem 11.1 (Rate of convergence in ζ_p^* for $p > \frac{d}{2}$). *Let $X_1, \dots, X_n, Y_1, \dots, Y_n$ be random variables in Q^d such that the couples (X_j, Y_j) and (X_k, Y_k) are independent for $j \neq k$, and such that X_j and Y_j have equal distribution for every $j = 1, \dots, n$. If $p > \frac{d}{2}$, then*

$$\mathbb{E}\big(\zeta_p^*(\mu_n, \nu_n)\big) \le \frac{C_{p,d}}{\sqrt{n}} \tag{11.1}$$

where $C_{p,d} > 0$ depends on (p,d) only. If $X_1, \dots, X_n, Y_1, \dots, Y_n$ are independent and still X_j and Y_j are equidistributed for every $j = 1, \dots, n$, then a similar bound holds true for the ψ_2-norm of the distance $\zeta_p^*(\mu_n, \nu_n)$. If the random variables $X_1, \dots, X_n$ are independent with same law μ, the inequality (11.1) also holds true for $\mathbb{E}(\zeta_p^*(\mu_n, \mu))$.

Proof. The argument is similar to the one discussed in [5] and [7]. By the assumption, for any $m \in \mathbb{Z}^d$,

$$f_{\mu_n}(m) - f_{\nu_n}(m) = \frac{1}{n}\sum_{j=1}^n \big(e^{im\cdot X_j} - e^{im\cdot Y_j}\big),$$

which represents, by the assumptions on the random variables X_j, Y_j, $j = 1, \dots, n$, a normalized sum of n uncorrelated (complex-valued) random variables with mean zero. Hence

$$\mathbb{E}\big(|f_{\mu_n}(m) - f_{\nu_n}(m)|^2\big) = \frac{1}{n^2}\sum_{j=1}^n \mathbb{E}\big(|e^{im\cdot X_j} - e^{im\cdot Y_j}|^2\big) \le \frac{4}{n}.$$

An application of (8.2) with an integer p then leads to

$$\mathbb{E}\big(\zeta_p^*(\mu_n, \nu_n)\big) \le \frac{2d^{\frac{p-1}{2}}}{\sqrt{n}}\bigg(\sum_{m\neq 0}\frac{1}{|m|^{2p}}\bigg)^{1/2} \le \frac{C_{p,d}}{\sqrt{n}}.$$

A similar inequality is also obtained on the basis of (9.3) for non-integer values of p when choosing $q(s) = \log^2(2s)$, $s \ge 1$.

The statement about the ψ_2-norm means that

$$\mathbb{E}\Big(\exp\big\{c_{p,d}\, n\, \zeta_p^*(\mu_n, \nu_n)^2\big\}\Big) \le 2$$

with some constants $c_{p,d} > 0$ depending on (p,d) only. It is explained in detail in [5].

For the last claim of the statement, note that the functional $\mu \to \zeta_p^*(\mu_n, \mu)$ is convex on the space of all probability measures μ on Q^d. If the random variables $X_1, \dots, X_n, Y_1, \dots, Y_n$ are independent with same law μ, then $\mathbb{E}(\nu_n) = \mu$ and, by Jensen's inequality,

$$\mathbb{E}_Y\big(\zeta_p^*(\mu_n, \nu_n)\big) \, \geq \, \zeta_p^*(\mu_n, \mu)$$

where $\mathbb{E}_Y$ means partial integration with respect to the sample $Y = (Y_1, \dots, Y_n)$. Taking another expectation with respect to $X = (X_1, \dots, X_n)$ yields $\mathbb{E}(\zeta_p^*(\mu_n, \nu_n)) \geq \mathbb{E}(\zeta_p^*(\mu_n, \mu))$. It remains to apply (11.1). □

The rate (11.1) compares, within the range $p > \frac{d}{2}$, with the known rates of convergence of empirical measures in Kantorovich distances W_p [8], [10]. A specific feature however of Zolotarev metrics is that that the standard rate $\frac{1}{\sqrt{n}}$ in Theorem 11.1 is actually optimal with respect to the growing number n, in contrast with the picture for Kantorovich distances for which smaller rates are possible (cf. [10], [12]...).

Proposition 11.2 (Minimal rate in ζ_p^*). *Assume that the random variables* $X_1, \dots, X_n$, $Y_1, \dots, Y_n$ *with values in* Q^d *are independent with a common non-degenerate distribution* μ. *Then, for any* $p > 0$,

$$\mathbb{E}\big(\zeta_p^*(\mu_n, \nu_n)\big) \, \geq \, \frac{c}{\sqrt{n}} \tag{11.2}$$

for some $c > 0$ *only depending on* (p, d, μ).

Proof. By Definition 5.1,

$$\zeta_p^*(\mu_n, \nu_n) \, = \, \sup_{u \in M_{p,d}} \bigg| \int_{Q^d} u \, d\mu_n - \int_{Q^d} u \, d\nu_n \bigg| \, = \, \frac{1}{n} \sup_{u \in M_{p,d}} \bigg| \sum_{j=1}^n \big(u(X_j) - u(Y_j)\big) \bigg|.$$

For any function u on the right-hand side, the random variables $\xi_j = u(X_j) - u(Y_j)$, $j = 1, \dots, n$, are independent and identically distributed. In addition, they are bounded by the constant $C = \max_{x,y} |u(x) - u(y)| \leq d^2 \, (2\pi)^{k+1}$ and have mean zero, according to (5.5). By the Marcinkiewicz-Zygmund inequality,

$$\mathbb{E}\big(|\xi_1 + \dots + \xi_n|\big) \, \geq \, c \, \mathbb{E}(Z_n), \quad Z_n \, = \, \Big[\frac{1}{n}(\xi_1^2 + \dots + \xi_n^2)\Big]^{1/2},$$

actually holding with $c = \frac{1}{2\sqrt{2}}$ (cf. e.g. Lemma 3.4 in [4]). Pick up an arbitrary function u in $M_{p,d}$ such that

$$\sigma^2 \, = \, \mathbb{E}\big(|u(X_1) - u(Y_1)|^2\big) \, = \, 2 \, \mathrm{Var}(\xi_1) > 0$$

which is possible since μ is not a delta-measure. Thus, $\mathbb{E}(\zeta_p^*(\mu_n, \nu_n)) \geq \frac{c}{\sqrt{n}} \, \mathbb{E}(Z_n)$. By the strong law of large numbers, $Z_n \to \sigma$ as $n \to \infty$ with probability 1. Since also $Z_n \leq C$, by the Lebesgue dominated convergence theorem $\mathbb{E}(Z_n) \to \sigma$. Hence $b = \inf_{n \geq 1} \mathbb{E}(Z_n) > 0$ and (11.2) holds true with $c = \frac{1}{2\sqrt{2}} \, b$. □

Returning to the upper bound (11.1) of Theorem 11.1, ζ_p^* therein cannot be replaced by ζ_p for $p = k + \alpha > 1$, since in general μ_n and ν_n do not need to have equal moments up to order k.

12 Convolution of measures on the torus

This section, and the following ones, are concerned with smoothing of the Fourier analytic bounds of Theorems 8.1 and 9.1. Indeed, the series therein might be divergent, and smoothing arguments will allow for the restriction of the summation over m to a finite set (such as a large cube), or to insert into the series a decaying factor at the expense of a reasonably small error. This is achieved by applying the obtained Fourier analytic bounds to the convolved measures.

The convolution operation may be defined in a canonical way on every (locally compact) group including the torus $(\mathbb{S}^1)^d = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$, the product of the circle with itself d times. Identifying the torus with $Q^d = (-\pi, \pi]^d$, the convolution $\overline{\mu} = \mu * \kappa$ of two signed measures μ and κ on Q^d may be defined via the equality

$$\int_{Q^d} u \, d\overline{\mu} \;=\; \int_{Q^d} \int_{Q^d} u(x+y) \, d\mu(x) \, d\kappa(y) \tag{12.1}$$

for all 2π-periodic continuous functions u on $\mathbb{R}^d$. Equivalently, it is sufficient to consider (12.1) for all exponential functions $u(x) = e^{im \cdot x}$, in which case this equality becomes

$$f_{\overline{\mu}}(m) \;=\; f_\mu(m) \, f_\kappa(m), \quad m \in \mathbb{Z}^d, \tag{12.2}$$

in terms of the associated Fourier-Stieltjes transforms.

If κ is close to the delta measure δ_0 at the origin in a weak sense, then $\overline{\mu}$ is close to μ. This property can be quantified in terms of the periodic Zolotarev distances ζ_p^*.

Lemma 12.1. *Given two probability measures μ and κ on Q^d, for any $p = k + \alpha > 0$,*

$$\zeta_p^*(\overline{\mu}, \mu) \;\le\; d \, (2\pi)^k \int_{Q^d} |x|^{p^*} \, d\kappa(x), \tag{12.3}$$

where $p^* = \min(p, 1)$.

Proof. Given a function u in $M_{p,d}$, apply the convolution definition (12.1) together with Lemma 5.2 to get

$$\begin{aligned} \bigg| \int_{Q^d} u \, d\overline{\mu} - \int_{Q^d} u \, d\mu \bigg| \;&=\; \bigg| \int_{Q^d} \int_{Q^d} \big(u(x+y) - u(x) \big) \, d\mu(x) \, d\kappa(y) \bigg| \\ &\le\; \int_{Q^d} \int_{Q^d} \big| u(x+y) - u(x) \big| \, d\mu(x) \, d\kappa(y) \\ &\le\; d \, (2\pi)^k \int_{Q^d} |y|^{p^*} \, d\kappa(y). \end{aligned}$$

It remains to take the supremum on the left-hand side over all admissible functions u. □

This lemma can already be used for the smoothing of the measures in Theorems 8.1 and 9.1, and thus for the replacement of f_μ and f_ν in the bounds (8.2) and (9.3) by the Fourier-Stieltjes transforms $f_{\overline{\mu}}$ and $f_{\overline{\nu}}$ according to (12.2). In order to get rid of the constraint that κ is supported on the torus Q^d, and thus reach more flexibility, the following elementary statement, apparently part of the folklore, will be helpful.

Lemma 12.2. *For any probability measure λ on $\mathbb{R}^d$, there exists a unique probability measure κ on Q^d such that*

$$f_\lambda(m) \,=\, f_\kappa(m) \quad \textit{for all } m \in \mathbb{Z}^d. \tag{12.4}$$

Moreover, for any $r > 0$,

$$\int_{Q^d} |x|^r \, d\kappa(x) \,\le\, \int_{\mathbb{R}^d} |x|^r \, d\lambda(x). \tag{12.5}$$

The measure κ will be called "the measure λ screwed on the torus".

Proof. Define the map $j : \mathbb{R} \to \mathbb{Z}$ by $\pi(2j-1) < x \le \pi(2j+1)$, $j = j(x)$. The map $U : \mathbb{R} \to (-\pi, \pi]$, $U(x) = x - 2\pi j(x)$, satisfies $U(x) = x$ for $x \in (-\pi, \pi]$ and $|U(x)| \le |x|$ for all $x \in \mathbb{R}$. Next define the product maps $U_d : \mathbb{R}^d \to Q^d$, $j_d : \mathbb{R}^d \to \mathbb{Z}^d$, $j_d(x) = (j(x_1), \dots, j(x_d))$,

$$U_d(x) \,=\, \big(U(x_1), \dots, U(x_d)\big) \,=\, x - 2\pi j_d(x), \quad x = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Similarly, $|U_d(x)| \le |x|$ for all $x \in \mathbb{R}^d$.

Denote by $\kappa = \lambda\, U_d^{-1}$ the image of λ under the map U_d. This measure is supported on Q^d and its Fourier-Stieltjes coefficients are given by, for every $m \in \mathbb{Z}^d$,

$$\begin{aligned} f_\kappa(m) \,&=\, \int_{Q^d} e^{im\cdot x} \, d\kappa(x) \,=\, \int_{\mathbb{R}^d} e^{im\cdot U_d(x)} \, d\lambda(x) \\ &=\, \int_{\mathbb{R}^d} e^{im\cdot (x - 2\pi j_d(x))} \, d\lambda(x) \,=\, \int_{\mathbb{R}^d} e^{im\cdot x} \, d\lambda(x) \,=\, f_\lambda(m). \end{aligned}$$

Hence (12.4) holds true. Since $|U_d(x)| \le |x|$, $x \in \mathbb{R}^d$,

$$\int_{Q^d} |x|^r \, d\kappa(x) \,=\, \int_{\mathbb{R}^d} |U_d(x)|^r \, d\lambda(x) \,\le\, \int_{\mathbb{R}^d} |x|^r \, d\lambda(x),$$

and therefore (12.5) is also satisfied. Finally, the uniqueness of κ follows from (12.4) and the fact that the sequence $(f_\kappa(m))_{m\in\mathbb{Z}^d}$ determines this measure in a unique way. This is a particular case of the Stone-Weiesrtrass theorem specialized to the torus $(\mathbb{S}^1)^d$ and the family of functions $(z_1, \dots, z_d) \to z_1^{m_1} \dots z_d^{m_d}$, $m = (m_1, \dots, m_d) \in \mathbb{Z}^d$. The proof is complete. □

Combining Lemma 12.2 with Lemma 12.1 leads to the following further statement.

Lemma 12.3. *Given a probability measure μ on Q^d and a probability measure λ on $\mathbb{R}^d$, let $\overline{\mu} = \mu * \kappa$ where κ is λ screwed on the torus. For any $p = k + \alpha > 0$, with $p^* = \min(p, 1)$,*

$$\zeta_p^*(\overline{\mu}, \mu) \,\le\, d\,(2\pi)^k \int_{\mathbb{R}^d} |x|^{p^*} \, d\lambda(x). \tag{12.6}$$

For the further purposes, since $p^* \le 1$, (12.6) may be simplified by Jensen's inequality into

$$\zeta_p^*(\overline{\mu}, \mu) \,\le\, d\,(2\pi)^k \bigg(\int_{\mathbb{R}^d} |x| \, d\lambda(x) \bigg)^{p^*}. \tag{12.7}$$

13 Smoothed Fourier analytic inequalities

Choosing a random vector Z in $\mathbb{R}^d$ with distribution λ and characteristic function $h = f_\lambda$ in Lemmas 12.2 and 12.3, and applying the triangle inequality for ζ_p^*, the inequality (12.7) yields

$$\zeta_p^*(\mu,\nu) \,\le\, \zeta_p^*(\overline{\mu},\overline{\nu}) + \zeta_p^*(\overline{\mu},\mu) + \zeta_p^*(\overline{\nu},\nu) \,\le\, \zeta_p^*(\overline{\mu},\overline{\nu}) + 2d\,(2\pi)^k\,\mathbb{E}\big(|Z|\big)^p \tag{13.1}$$

for all probability measures μ and ν on Q^d, where it is recalled that $p^* = \min(p,1)$. Here, according to (12.2) and (12.4), the convolved measures have Fourier-Stieltjes transforms

$$f_{\overline{\mu}}(m) \,=\, f_\mu(m)\,h(m), \quad f_{\overline{\nu}}(m) \,=\, f_\mu(m)\,h(m) \quad m \in \mathbb{Z}^d.$$

Theorems 8.1 and 9.1 may then be applied to the couple $(\overline{\mu},\overline{\nu})$, and (13.1) then leads to more general upper bounds on $\zeta_p^*(\mu,\nu)$. As in Section 9, the function $q : [1,\infty) \to (0,\infty)$ is non-decreasing with

$$(0 <) \quad C_q \,=\, \bigg(\sum_{\ell=0}^\infty \frac{1}{q(2^\ell)}\bigg)^{1/2} \,<\, \infty.$$

Theorem 13.1 (Smoothed Fourier analytic inequalities for ζ_p^*)**.** *Let Z be a random vector in $\mathbb{R}^d$ with characteristic function h. Given two probability measures μ and ν on Q^d, for any integer $p \ge 1$,*

$$\zeta_p^*(\mu,\nu) \,\le\, d^{\frac{p-1}{2}}\bigg(\sum_{m\neq 0}\frac{|h(m)|^2}{|m|^{2p}}\,\big|f_\mu(m)-f_\nu(m)\big|^2\bigg)^{1/2} + 2d\,(2\pi)^{p-1}\,\mathbb{E}\big(|Z|\big). \tag{13.2}$$

If $p = k+\alpha > 0$ is not an integer, then for any non-decreasing positive function q on $[1,\infty)$,

$$\zeta_p^*(\mu,\nu) \,\le\, A\bigg(\sum_{m\neq 0}\frac{q(|m|)\,|h(m)|^2}{|m|^{2p}}\,\big|f_\mu(m)-f_\nu(m)\big|^2\bigg)^{1/2} + 2d\,(2\pi)^{p-1}\,\mathbb{E}\big(|Z|\big)^{p^*} \tag{13.3}$$

with $A = 2^{\frac{p-1}{2}} d^{\frac{p+1}{2}} \pi C_q$.

In the case $Z = 0$, the statement reduces to Theorems 8.1 and 9.1. If p is not an integer, a possible weight q in (13.3) is given by $q(s) = \log^2(2s)$, $s \ge 1$, which leads to an additional logarithmic factor inside the sum.

These bounds are also applicable to the Zolotarev distance ζ_p according to Corollaries 8.2 and 9.2.

Corollary 13.2 (Smoothed Fourier analytic inequalities for ζ_p)**.** *Let $p = k+\alpha > 0$, and let Z be a random vector in $\mathbb{R}^d$ with characteristic function h. If μ and ν are probability measures supported on the cube $[0,\pi]^d$ with equal mixed moments up to the order k, then for any integer $p \ge 1$,*

$$c_{p,d}\,\zeta_p(\mu,\nu) \,\le\, \bigg(\sum_{m\neq 0}\frac{|h(m)|^2}{|m|^{2p}}\,\big|f_\mu(m)-f_\nu(m)\big|^2\bigg)^{1/2} + \mathbb{E}\big(|Z|\big) \tag{13.4}$$

for some $c_{p,d} > 0$*. If* $p = k + \alpha > 0$ *is not an integer, then*

$$c_{p,d}\,\zeta_p(\mu,\nu) \,\leq\, C_q\bigg(\sum_{m \neq 0} \frac{q(|m|)\,|h(m)|^2}{|m|^{2p}}\,\left|f_\mu(m) - f_\nu(m)\right|^2\bigg)^{1/2} + \mathbb{E}\big(|Z|\big)^{p^*}. \tag{13.5}$$

The following addresses special choices of the random vector Z. One natural particular choice of Z is the normal distribution with mean zero and covariance matrix $\sqrt{t}\,\mathrm{I}_d$ in $\mathbb{R}^d$, $t > 0$. In this case, $\mathbb{E}(|Z|) \leq \sqrt{dt}$ and $h(m) = e^{-t|m|^2/2}$, $m \in \mathbb{Z}^d$.

Corollary 13.3. *Given two probability measures* μ *and* ν *on* Q^d*, for any integer* $p \geq 1$*, for any real* $t > 0$*,*

$$c_{p,d}\,\zeta_p^*(\mu,\nu) \,\leq\, \bigg(\sum_{m \neq 0} \frac{1}{|m|^{2p}}\,\left|f_\mu(m) - f_\nu(m)\right|^2 e^{-t|m|^2}\bigg)^{1/2} + \sqrt{t} \tag{13.6}$$

for some $c_{p,d} > 0$*. If* $p = k + \alpha > 0$ *is not an integer, for any real* $t > 0$*,*

$$c_{p,d}\,\zeta_p^*(\mu,\nu) \,\leq\, C_q\bigg(\sum_{m \neq 0} \frac{q(|m|)}{|m|^{2p}}\,\left|f_\mu(m) - f_\nu(m)\right|^2 e^{-t|m|^2}\bigg)^{1/2} + t^{p^*/2}. \tag{13.7}$$

If μ *and* ν *are supported on* $[0,\pi]^d$ *and have equal mixed moments up to order* k*, then similar bounds hold true for* ζ_p *in place of* ζ_p^**.*

Another natural choice is $Z = \frac{1}{T}\,\xi$ with parameter $T > 0$ where ξ is a random vector in $\mathbb{R}^d$ whose characteristic function is C^2-smooth and supported on the cube $[-1,1]^d$. Such distributions exist, and moreover, it may be required that $\mathbb{E}(|\xi|)^2 = 12\,d$. Since in this case $h(m)$ is supported on the cube $[-T,T]^d$, Theorem 13.1 yields the following corollary. When $m = (m_1,\dots,m_d) \in \mathbb{Z}^d$, $\|m\|_\infty = \max(|m_1|,\dots,|m_d|)$.

Corollary 13.4. *Given two probability measures* μ *and* ν *on* Q^d*, for any integer* $p \geq 1$ *and any real* $T > 0$*,*

$$c_{p,d}\,\zeta_p^*(\mu,\nu) \,\leq\, \bigg(\sum_{1 \leq \|m\|_\infty \leq T} \frac{1}{|m|^{2p}}\,\left|f_\mu(m) - f_\nu(m)\right|^2\bigg)^{1/2} + \frac{1}{T} \tag{13.8}$$

for some $c_{p,d} > 0$*. If* $p = k + \alpha > 0$ *is not an integer, for any real* $T > 0$*,*

$$c_{p,d}\,\zeta_p^*(\mu,\nu) \,\leq\, C_q\bigg(\sum_{1 \leq \|m\|_\infty \leq T} \frac{q(|m|)}{|m|^{2p}}\,\left|f_\mu(m) - f_\nu(m)\right|^2\bigg)^{1/2} + \frac{1}{T^{p^*}}. \tag{13.9}$$

If μ *and* ν *are supported on* $[0,\pi]^d$ *and have equal mixed moments up to order* k*, then similar bounds hold true for* ζ_p *in place of* ζ_p^**.*

The range $p < 1$ in (13.9) was considered in [5], [6].

In conclusion of this section, some special aspects, and simplifications, of the preceding bounds in dimensions $d = 1$ and $d = 2$ are emphasized, with respect to the corrections $\frac{1}{T}$ and $\frac{1}{T^{p^*}}$ in the smoothing inequalities (13.8) and (13.9) respectively.

Start with dimension $d = 1$, and (8.2) of Theorem 8.1. If $p \geq \frac{3}{2}$, the smoothing in the form (13.8) is actually not needed. In this case namely, for any $T \geq 1$,

$$\sum_{m>T} \frac{1}{m^{2p}} \left|f_\mu(m) - f_\nu(m)\right|^2 \leq 4 \sum_{m>T} \frac{1}{m^{2p}} \leq \frac{C}{T^{2p-1}}$$

with some absolute constant $C > 0$, so that (13.8) actually holds true with $\frac{1}{T^{p-\frac{1}{2}}}$ instead of $\frac{1}{T}$ on the right-hand side. A similar conclusion arises about (9.3) of Theorem 9.1 when $p \geq \frac{3}{2} + \varepsilon$, $\varepsilon > 0$, by choosing, for example $q(s) = \log^2(2s)$, $s \geq 1$, with a corresponding improvement of (13.9). If $p = 1$, the series in (8.2) is also convergent, but smoothing yields the improved dependence in $T > 0$ of the form $\frac{1}{T}$ instead of $\frac{1}{\sqrt{T}}$. If μ and ν are supported on $[0, \pi]$, a similar bound holds true for ζ_1 (without any moment restriction).

In the two-dimensional case, the series (8.2) of Theorem 8.1 is convergent as long as $p > 1$, and the same is true about (9.3) of Theorem 9.1 for the choice, for instance, of $q(s) = \log^2(2s)$, $s \geq 1$. If $p \geq 2$, the smoothing in the form (13.8) is not needed as well since, for any $T \geq 1$,

$$\sum_{\|m\|_\infty > T} \frac{1}{|m|^{2p}} \left|f_\mu(m) - f_\nu(m)\right|^2 \leq C_1 \sum_{m>T} \frac{1}{m^{2p-1}} \leq \frac{C_2}{T^{2p-2}}$$

with some absolute constants $C_1, C_2 > 0$. Hence (13.8) can be sharpened to the form with a factor $\frac{1}{T^{p-1}}$ instead of $\frac{1}{T}$. A similar conclusion can be developed on (9.3) of Theorem 9.1 when $p \geq 2 + \varepsilon$, $\varepsilon > 0$, by choosing $q(s) = \log^2(2s)$, $s \geq 1$, leading to a corresponding improvement. If $p = 1$, the series in (8.2) is not convergent, and smoothing is required in the form of (13.8). If μ and ν are supported on $[0, \pi]^2$, a similar bound holds true for ζ_1 (again without any moment restriction).

14 Exponentially decaying characteristic functions

Turning back to Lemma 12.3, the bound therein for $p > 1$ may or may not be satisfactory in applications, and it would be more natural to replace the power $p^* = \min(p, 1)$ with p. If so, this would potentially give a much better approximation of μ by the convolved measure $\overline{\mu}$. However, towards this goal, the smoothing measure κ should be allowed to be a signed measure. In dimension one, such smoothing measures were discussed in [2], and used there to derive some variants of Esseen's inequality for the Kolmogorov distance between distribution functions in terms of their Fourier-Stieltjes transforms.

A preliminary step within such an investigation is the (technical) construction of special compactly supported probability distributions H whose characteristic functions $h(t)$ decay exponentially fast "on average". This is despite the fact that a pointwise bound $|h(t)| \leq Ce^{-ct}$, $t \geq t_0$, is never possible with some positive constants c and C in this class of distributions.

Lemma 14.1. *For any $0 < \delta \leq \pi$ and $T \geq 0$, there exists a symmetric probability measure $H = H_{\delta T}$ on the interval $(-\delta, \delta)$ with characteristic function h satisfying*

$$\sum_{|m| \geq T/b} \left|h(bm)\right| \leq \frac{c}{b\delta} \, e^{-\delta T/5}, \quad \sum_{m \in \mathbb{Z}} |h(bm)| \leq \frac{c}{b\delta} \tag{14.1}$$

for all $0 < b \leq 1$ with some absolute constant $c > 0$.

Proof. The second inequality is a particular case of the first one with $T = 0$. Let $U_1, \dots, U_n$ be independent random variables uniformly distributed in the interval $(-1,1)$. The random variable $U = \frac{\delta}{n}\,(U_1 + \dots + U_n)$ takes values in $(-\delta,\delta)$, and its characteristic function

$$h(t) \,=\, \mathbb{E}(e^{itX}) \,=\, \Big(\frac{\sin(\delta t/n)}{\delta t/n}\Big)^n, \quad t \in \mathbb{R},$$

satisfies

$$\big|h(t)\big| \,\leq\, \Big(\frac{n}{\delta|t|}\Big)^n. \tag{14.2}$$

Assume first that $\delta T \geq 2\pi$ and choose $n = [\delta T/3]$. Put $T_0 = [T/b]$. Since $\delta \leq \pi$, necessarily $\frac{T}{b} \geq T \geq 2$, so $T_0 \geq \frac{2}{3} \cdot \frac{T}{b}$. It also holds true that

$$2 \,\leq\, n \,\leq\, \frac{\delta T}{3}\,, \qquad \frac{T_0}{n-1} \,\leq\, \frac{9}{b\delta}\,.$$

For the proof of the last inequality, note that, since $x = \delta T/3 > 2$,

$$\frac{T_0}{n-1} \,\leq\, \frac{T/b}{n-1} \,=\, \frac{3}{b\delta}\,\frac{x}{[x]-1}\,.$$

The latter fraction is maximized when $2 < x < 3$, $x \to 3$, so it does not exceed 3. In addition,

$$\frac{n}{b\delta\, T_0} \,\leq\, \frac{\delta T/3}{b\delta \cdot \frac{2}{3} \cdot \frac{T}{b}} \,=\, \frac{1}{2}\,.$$

Using these estimates in (14.2) implies that

$$\begin{aligned}
\sum_{|m| \geq T/b} \big|h(bm)\big| \,&=\, 2 \sum_{m=T_0}^\infty \big|h(bm)\big| \\
&=\, 2\,\big|h(bT_0)\big| + 2 \sum_{m=T_0+1}^\infty \big|h(bm)\big| \\
&\leq\, 2\,\Big(\frac{n}{b\delta T_0}\Big)^n + 2 \sum_{m=T_0+1}^\infty \Big(\frac{n}{b\delta m}\Big)^n \\
&\leq\, 2\,\Big(\frac{n}{b\delta T_0}\Big)^n + 2 \int_{T_0}^\infty \Big(\frac{n}{b\delta t}\Big)^n\,dt \\
&=\, 2\,\Big(1 + \frac{T_0}{n-1}\Big)\,\Big(\frac{n}{b\delta T_0}\Big)^n \\
&\leq\, 2\,\Big(1 + \frac{9}{b\delta}\Big)\,2^{-n} \,\leq\, 2\,\frac{\pi + 9}{b\delta}\,e^{-n\log 2},
\end{aligned}$$

using that $b\delta \leq \pi$. Since $n \geq \frac{\delta T}{3} - 1$ and $\frac{1}{3}\log 2 > \frac{1}{5}$, it follows that

$$\sum_{|m| \geq T/b} |h(bm)| \,\leq 4\,\frac{\pi + 9}{b\delta}\,e^{-\delta T/5} \,<\, \frac{50}{b\delta}\,e^{-\delta T/5}$$

which proves the first estimate in (14.1).

In the case $\delta T < 2\pi$, take $n = 2$ and note that

$$\sum_{|m| \geq T/b} \big|h(bm)\big| \, \leq \, \sum_{m \in \mathbb{Z}} \big|h(bm)\big| \, = \, 1 + 2S, \quad S \, = \, \sum_{m=1}^{\infty} \Big(\frac{\sin(b\delta\, m/2)}{b\delta\, m/2}\Big)^2.$$

Set $m_0 = [\pi/b\delta]$ and split the last sum into two parts. First

$$\sum_{m=1}^{m_0} \Big(\frac{\sin(b\delta\, m/2)}{b\delta\, m/2}\Big)^2 \, \leq \, \sum_{m=1}^{m_0} 1 \, = \, m_0 \leq \frac{\pi}{b\delta}\,.$$

Secondly, since $m_0 \geq 1$ and therefore $m_0 \geq \frac{\pi}{2b\delta}$,

$$\begin{aligned} \sum_{m_0+1}^{\infty} \Big(\frac{\sin(b\delta\, m/2)}{b\delta\, m/2}\Big)^2 \, &\leq \, \sum_{m_0+1}^{\infty} \Big(\frac{2}{b\delta m}\Big)^2 \\ &\leq \, \frac{4}{(b\delta)^2} \int_{m_0}^{\infty} \frac{1}{x^2}\, dx \\ &= \, \frac{4}{(b\delta)^2\, m_0} \, \leq \, \frac{8}{\pi b\delta}\,. \end{aligned}$$

Hence $S \leq (\pi + \frac{8}{\pi}) \frac{1}{b\delta} < \frac{6}{b\delta}$ and $1 + 2S < \frac{13}{b\delta}$. Therefore, using that $\delta T \leq 2\pi$, it follows that

$$\sum_{m \in \mathbb{Z}} \big|h(bm)\big| \, \leq \, \frac{13}{b\delta}\, e^{\delta T/5}\, e^{-\delta T/5} \, \leq \, \frac{13}{b\delta}\, e^{2\pi/5}\, e^{-\delta T/5} \, < \, \frac{50}{b\delta}\, e^{-\delta T/5}.$$

This concludes the proof of Lemma 14.1. □

A "continuous" variant of Lemma 14.1 where the sums are replaced with integrals is discussed in [2]. The distributions $H = H_T$ with $\delta = 1$ were used by Zolotarev in the proof of the Esseen-type inequality for the Lévy distance between distribution functions.

15 Signed measures with several zero moments

Using the probability distributions from Lemma 14.1, it is possible to construct compactly supported signed measures with additional special properties in terms of zero moments.

Lemma 15.1. *For any $0 < \delta \leq \pi$, $T \geq 0$, and an integer $L \geq 1$, there exists a symmetric signed measure κ on the interval $(-\delta, \delta)$ with total variation norm $\|\kappa\|_{\mathrm{TV}} \leq c^L$ for some absolute constant $c > 0$, and such that*

$$\kappa(-\delta, \delta) \, = \, 1, \quad \int_{-\delta}^{\delta} x^\ell \, d\kappa(x) \, = \, 0, \quad \ell = 1, \dots, L. \tag{15.1}$$

In addition, its Fourier-Stieltjes transform $\widehat{\kappa}$ satisfies

$$\sum_{|m| \geq T} \big|\widehat{\kappa}(m)\big| \, \leq \, \frac{c^L}{\delta}\, e^{-\delta T/5(L+1)}, \quad \sum_{m \in \mathbb{Z}} \big|\widehat{\kappa}(m)\big| \, \leq \, \frac{c^L}{\delta}. \tag{15.2}$$

Proof. Again, the second inequality in (15.2) is a particular case of the first one with $T = 0$.

If $L = 1$, the measure $\kappa = H$ from Lemma 14.1 does the job, due to its symmetry about the origin. In the general case $L \geq 1$, with the identification of distribution functions with measures, set

$$\kappa(x) \,=\, w_1 H\Big(\frac{x}{b_1}\Big) + \dots + w_{L+1} H\Big(\frac{x}{b_{L+1}}\Big)$$

with some weights $w_1, \dots, w_{L+1} \in \mathbb{R}$ and parameters $0 < b_1 < \dots < b_{L+1} \leq 1$. Then κ is supported on $(-\delta, \delta)$ as a measure and has total variation norm

$$\|\kappa\|_{\rm TV} \,\leq\, \sum_{i=1}^{L+1} |w_i|. \tag{15.3}$$

In addition, it has the Fourier-Stieltjes transform

$$\widehat\kappa(m) \,=\, \sum_{i=1}^{L+1} w_i\, h(b_i m), \quad m \in \mathbb{Z}. \tag{15.4}$$

Let $b = \min_{1 \leq i \leq L+1} b_i$, and take for H the measure from Lemma 14.1 with parameter bT in place of T and with the same δ. Hence, applying (14.1) with b_i in place of b, for every $i = 1, \dots, L+1$,

$$\sum_{|m| \geq bT/b_i} \big|h(b_i m)\big| \,\leq\, \frac{c}{b_i \delta}\, e^{-b\delta T/5}, \quad \sum_{m \in \mathbb{Z}} \big|h(b_i m)\big| \,\leq\, \frac{c}{b_i \delta}\,.$$

This may be weakened into

$$\sum_{|m| \geq T} \big|h(b_i m)\big| \,\leq\, \frac{c}{b\delta}\, e^{-b\delta T/5}, \quad \sum_{m \in \mathbb{Z}} \big|h(b_i m)\big| \,\leq\, \frac{c}{b\delta}.$$

Hence, according to (15.4),

$$\sum_{|m| \geq T} |\widehat\kappa(m)| \,\leq\, \frac{c}{b\delta}\, e^{-b\delta T/5} \sum_{i=1}^{L+1} |w_i|. \tag{15.5}$$

Turn then to the condition (15.1), which is the same as

$$\sum_{i=1}^{L+1} w_i \,=\, 1, \quad \Big(\sum_{i=1}^{L+1} w_i b_i^\ell\Big) \int_{-\delta}^{\delta} x^\ell\, dH(x) \,=\, 0$$

for all integers $\ell = 1, \dots, L$. Although the moments of H are vanishing for odd values of ℓ, the request may be strengthened into

$$\sum_{i=1}^{L+1} w_i \,=\, 1, \quad \sum_{i=1}^{L+1} w_i b_i^\ell \,=\, 0, \quad \ell = 1, \dots, L.$$

This is a linear system of $L+1$ equations in $L+1$ unknowns $w = (w_1, \dots, w_{L+1})$, which can be written in matrix form as $Vw = e_1$, where V is the Vandermonde matrix

$$V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ b_1^L & b_2^L & \cdots & b_{L+1}^L \end{pmatrix}$$

and $e_1 = (1, 0, \dots, 0)$ (as a column). It has a non-zero determinant $\det(V) = \prod_{i<j}(b_j - b_i)$, so that V is invertible and $w = V^{-1}e_1$. For the choice, for example, of

$$b_i = \frac{i}{L+1}, \quad i = 1, \dots, L+1, \tag{15.6}$$

there is a universal collection $w = (w_1, \dots, w_{L+1})$. In this case $b = \frac{1}{L+1}$, and (15.3), (15.5) yield the desired conclusions of Lemma 15.1 with some existing constants depending on L.

In order to quantify the dependence of the constants, the following result from [11] on the inverse of the Vandermonde matrix will be useful: if the norm of a $(L+1) \times (L+1)$ matrix $A = (a_{ij})$ is defined by

$$\|A\| = \max_{1 \le i \le L+1} \sum_{j=1}^{L+1} |a_{ij}|,$$

then

$$\|V^{-1}\| \le \max_{1 \le i \le L+1} \prod_{j \neq i} \frac{1 + |b_j|}{|b_i - b_j|}.$$

The choice (15.6) leads to

$$\begin{aligned} \|V^{-1}\| &\le \max_{1 \le i \le L+1} \prod_{j=1,\ j \neq i}^{L+1} \frac{L+j+1}{|j-i|} \\ &= \max_{1 \le i \le L+1} \frac{(L+2)\cdots(2L+2)}{(L+i+1)\,(i-1)!\,(L-i+1)!} \\ &= 2(L+1)\binom{2L+1}{L+1} \max_{1 \le i \le L+1} \frac{\binom{i-1}{L}}{L+i-1} \le 2^{3(L+1)} \end{aligned}$$

after the bound $\binom{n}{i} \le 2^n$ for the binomial coefficients. Since, for $i = 1, \dots, L+1$,

$$w_i = (V^{-1}e_1)_i = \sum_{j=1}^{L+1} (V^{-1})_{ij}\,\delta_{1j} = (V^{-1})_{i1},$$

it follows that $|w_i| \le \|V^{-1}\| \le 2^{3(L+1)}$, hence

$$\sum_{i=1}^{L+1} |w_i| \le (L+1)\,\|V^{-1}\| \le (L+1)\,2^{3(L+1)}.$$

It remains to apply this bound in (15.3) and (15.5). The proof is complete. $\square$

A multi-dimensional variant of the preceding Lemma 15.1 will be needed in a form properly adapted for applications. Denote by $\kappa_d = \kappa^d$ the product measure with marginal κ from Lemma 15.1 It has the Fourier-Stieltjes transform

$$\widehat{\kappa}_d(m) \, = \, \widehat{\kappa}(m_1) \cdots \widehat{\kappa}(m_d), \quad m = (m_1, \dots, m_d) \in \mathbb{Z}^d.$$

Let $0 < \delta \leq \pi$, $T > 0$, and let $L \geq 1$ be an integer.

Corollary 15.2. *The product signed measure κ_d is supported on the cube $(-\delta, \delta)^d$, has a total variation norm $\|\kappa_d\|_{\rm TV} \leq c^{dL}$ for some absolute constant $c > 0$ and*

$$\kappa_d\big((-\delta,\delta)^d\big) \, = \, 1, \quad \int_{-\delta}^{\delta} \cdots \int_{-\delta}^{\delta} x_1^{\ell_1} \cdots x_d^{\ell_d} \, d\kappa(x_1, \dots, x_d) \, = \, 0 \tag{15.7}$$

for all integers $\ell_i = 0, 1, \dots, L$ such that $\ell_i > 0$ for at least one index $i = 1, \dots, d$. In addition, for every $p \geq 0$,

$$\sum_{\|m\|_\infty \geq T} \frac{|\widehat{\kappa}(m)|^2}{|m|^{2p}} \, \leq \, \frac{c^{dL}}{T^{2p}} \, e^{-\delta T/5(L+1)}. \tag{15.8}$$

Proof. Only (15.8) requires some details. For every $i = 1, \dots, d$, the part of the sum in (15.8) covering the indices $\max_{j \neq i} |m_j| \leq |m_i|$ does not exceed, by (15.2),

$$\begin{aligned}
\sum_{|m_i| \geq T} \sum_{m_j \in \mathbb{Z}} \bigg(\frac{|\widehat{\kappa}(m_i)|^2}{|m_i|^{2p}} \prod_{j \neq i} |\widehat{\kappa}(m_j)|^2\bigg) \, &\leq \, \frac{1}{T^{2p}} \sum_{|m_i| \geq T} |\widehat{\kappa}(m_i)|^2 \prod_{j \neq i} \sum_{m_j \in \mathbb{Z}} |\widehat{\kappa}(m_j)|^2 \\
&\leq \, \frac{\|\kappa\|_{\rm TV}^d}{T^{2p}} \sum_{|m_i| \geq T} |\widehat{\kappa}(m_i)| \prod_{j \neq i} \sum_{m_j \in \mathbb{Z}} |\widehat{\kappa}(m_j)| \\
&\leq \, \frac{c^{dL}}{T^{2p}} \, \frac{c^{dL}}{\delta^d} \, e^{-\delta T/5(L+1)}
\end{aligned}$$

for some $c > 0$. For the final bound, the latter expression should be multiplied by d. But this factor can be absorbed in c^{dL} by choosing a larger constant c. □

16 Approximation by convolution with signed measures

On the basis of the analytic bounds produced in the previous sections, this paragraph addresses sharpenings of the convolution inequalities with signed measures.

Let μ be a probability measure on the torus Q^d, and recall the convolution $\overline{\mu} = \mu * \kappa$, understood in the periodic sense (12.1). Using the signed measure $\kappa = \kappa_d$ of Corollary 15.2, the inequality (12.3) of Lemma 12.1 may be improved in the case $p > 1$.

Lemma 16.1. *Let $p = k + \alpha > 0$. Let κ be a signed measure supported on $(-\delta, \delta)^d$, $0 < \delta \leq \pi$, with total mass $\kappa(Q^d) = 1$ and such that*

$$\int_{Q^d} x_1^{k_1} \cdots x_d^{k_d} \, d\kappa(x) \, = \, 0 \tag{16.1}$$

for any collection of integers $k_j \geq 0$ satisfying $1 \leq k_1 + \dots + k_d \leq k$. Then the signed measure $\overline{\mu}$ satisfies

$$\zeta_p^*(\overline{\mu}, \mu) \, \leq \, d e^d \, \delta^p \, \|\kappa\|_{\rm TV}. \tag{16.2}$$

Proof. Let u be a function in $M_{p,d}$ of class C^∞. By definition of the convolution $\overline{\mu} = \mu * \kappa$,

$$\int_{Q^d} u\, d\overline{\mu} - \int_{Q^d} u\, d\mu \;=\; \int_{Q^d}\int_{Q^d} \big(u(x+y) - u(x)\big)\, d\mu(x)\, d\kappa(y) \tag{16.3}$$

using that κ has total mass 1.

If $p \leq 1$, then $k = 0$, $\alpha = p$, and (16.1) is not a restriction. By the assumption, the function u has Lip(α) semi-norm at most 1 with respect to every coordinate. Hence, by Lemma 2.2, for all $x \in Q^d$ and $y = (y_1, \dots, y_d) \in (-\delta, \delta)^d$,

$$\big|u(x+y) - u(x)\big| \;\leq\; \sum_{i=1}^d |y_i|^\alpha \leq d\delta^\alpha.$$

Using this bound in (16.3) yields (16.2) without the factor e^d.

Now, let $p > 1$, so that $k \geq 1$. In order to bound the double integral in (16.3), apply a multi-dimensional integral Taylor formula at the point x as in the proof of Proposition 7.1. With the standard multi-dimensional notation,

$$u(x+y) - u(x) \;=\; \sum_{1 \leq |\beta| \leq k} \frac{D^\beta u(x)}{\beta!}\, y^\beta + \sum_{|\beta| = k} Q_\beta(y)\, y^\beta \tag{16.4}$$

where

$$Q_\beta(y) \;=\; \frac{|\beta|}{\beta!} \int_0^1 (1-t)^{|\beta|-1} \big(D^\beta u(x+ty) - D^\beta u(x)\big)\, dt.$$

Due to the moment assumption (16.1), the integration of the formula (16.4) over κ yields

$$\begin{aligned}\int_{Q^d} &\big(u(x{+}y) - u(x)\big) d\kappa(y) \\ &= \sum_{|\beta|=k} \frac{|\beta|}{\beta!} \int_0^1 (1-t)^{|\beta|-1} \bigg[\int_{Q^d} \big(D^\beta u(x+ty) - D^\beta u(x)\big)\, y^\beta\, d\kappa(y)\bigg]\, dt.\end{aligned}$$

Since $u \in M_{p,d}$, the partial derivatives $D^\beta u$ of the order $|\beta| = k$ have Lip(α) semi-norms at most 1 with respect to every coordinate. Hence, by Lemma 2.2, for all $t \in [0,1]$, $x \in Q^d$ and $y = (y_1, \dots, y_d) \in (-\delta, \delta)^d$,

$$\big|D^\beta u(x+ty) - D^\beta u(x)\big| \;\leq\; \sum_{i=1}^d |ty_i|^\alpha \;\leq\; d\delta^\alpha.$$

Since also $|y^\beta| = |y_1|^{k_1} \cdots |y_d|^{k_d} \leq \delta^k$ for any multi-index $\beta = (k_1, \dots, k_d)$ with $|\beta| = k$,

$$\bigg|\int_{Q^d} (u(x+y) - u(x))\, d\kappa(y)\bigg| \;\leq\; d\delta^p\, \|\kappa\|_{\rm TV} \sum_{|\beta|=k} \frac{1}{\beta!}\,.$$

Extending the summation in the last sum to all $k_i \geq 0$, it may be bounded by e^d. Finally, applying this estimate to (16.3), Fubini's theorem yields

$$\bigg|\int_{Q^d} u\, d\overline{\mu} - \int_{Q^d} u\, d\mu\bigg| \;\leq\; de^d\, \delta^p\, \|\kappa\|_{\rm TV}.$$

It remains to take the supremum over all admissible functions u in $M_{p,d}$ to conclude the argument. □

The following corollary is a simple consequence of (16.2) and the triangle inequality for ζ_p^*.

Corollary 16.2. *Under the assumptions of Lemma 16.1, the convolutions* $\overline{\mu} = \mu * \kappa$ *and* $\overline{\nu} = \nu * \kappa$ *of any two probability measures* μ *and* ν *on the torus* Q^d *with* κ *satisfy*

$$\zeta_p^*(\mu,\nu) \, \leq \, \zeta_p^*(\overline{\mu},\overline{\nu}) + 2de^d\,\delta^p\,\|\kappa\|_{\rm TV}. \tag{16.5}$$

17 Smoothing with the help of signed measures

This final technical section gathers all the preceding efforts to improve the remainder terms in the smoothed Fourier analytic inequalities as stated in Corollary 13.4.

Given two probability measures μ and ν on the torus Q^d, consider their convolutions with the signed measure $\kappa = \kappa_d$ from Corollary 15.2 with two parameters $0 < \delta \leq \pi$ and $T > 0$.

First assume that $p \geq 1$ is integer. Applying Theorem 8.1 to the measure $\lambda = \overline{\mu} - \overline{\nu}$ together with the smoothing inequality (16.5) from Corollary 16.2 yields

$$\zeta_p^*(\mu,\nu) \, \leq \, d^{\frac{p-1}{2}} \bigg(\sum_{m \neq 0} \frac{1}{|m|^{2p}}\, \big|f_{\overline{\mu}}(m) - f_{\overline{\nu}}(m)\big|^2\bigg)^{1/2} + 2de^d\,\delta^p\,\|\kappa\|_{\rm TV}.$$

Here, the convolved measures have Fourier-Stieltjes transforms

$$f_{\overline{\mu}}(m) \, = \, f_\mu(m)\,\widehat{\kappa}(m), \quad f_{\overline{\nu}}(m) = f_\mu(m)\,\widehat{\kappa}(m), \quad m \in \mathbb{Z}^d.$$

Splitting the summation to the integer points m with $\|m\|_\infty \leq T$ and $\|m\|_\infty > T$, and using $|\widehat{\kappa}(m)| \leq \|\kappa\|_{\rm TV}$ for the first region and $|f_\mu(m)| \leq 1$, $|f_\nu(m)| \leq 1$ for the second one, implies that

$$\begin{aligned}\zeta_p^*(\mu,\nu) \, &\leq \, d^{\frac{p-1}{2}}\,\|\kappa\|_{\rm TV} \bigg(\sum_{1 \leq \|m\|_\infty \leq T} \frac{1}{|m|^{2p}}\, \big|f_\mu(m) - f_\nu(m)\big|^2\bigg)^{1/2} \\ &\quad + d^{\frac{p-1}{2}}\,\|\kappa\|_{\rm TV} \bigg(\sum_{\|m\|_\infty > T} \frac{1}{|m|^{2p}}\, \big|f_{\widehat{\kappa}}(m)\big|^2\bigg)^{1/2} + 2de^d\,\delta^p\,\|\kappa\|_{\rm TV}.\end{aligned}$$

Next Corollary 15.2 may be applied leading to

$$c_{p,d}\,\zeta_p^*(\mu,\nu) \, \leq \, \bigg(\sum_{1 \leq \|m\|_\infty \leq T} \frac{1}{|m|^{2p}}\, \big|f_\mu(m) - f_\nu(m)\big|^2\bigg)^{1/2} + \frac{1}{T^p}\, e^{-\delta T/10(k+1)} + \delta^p$$

with some (p,d)-dependent constant $c_{p,d} > 0$. A natural choice here is given by $\delta = \frac{1}{T}$, leading to a sharpened form of the inequality (13.8) of Corollary 13.4 in which the remainder term $\frac{1}{T}$ is replaced with $\frac{1}{T^p}$. A similar argument based on the application of Theorem 9.1 also works when p is not an integer.

The following theorem summarized the main conclusions for integer and fractional values of $p > 0$. Recall that $q : [1,\infty) \to (0,\infty)$ is a non-decreasing weight function with $C_q = \big(\sum_{\ell=0}^\infty \frac{1}{q(2^\ell)}\big)^{1/2} < \infty$.

Theorem 17.1 (Improved smoothed Fourier analytic inequalities)**.** *Given two probability measures μ and ν on Q^d, for any integer $p \geq 1$ and any real $T > 0$,*

$$c_{p,d}\, \zeta_p^*(\mu,\nu) \, \leq \, \bigg(\sum_{1 \leq \|m\|_\infty \leq T} \frac{1}{|m|^{2p}} \, \big|f_\mu(m) - f_\nu(m)\big|^2\bigg)^{1/2} + \frac{1}{T^p} \tag{17.1}$$

for some $c_{p,d} > 0$. If $p = k + \alpha > 0$ is not an integer, for any real $T > 0$,

$$c_{p,d}\, \zeta_p^*(\mu,\nu) \, \leq \, C_q \bigg(\sum_{1 \leq \|m\|_\infty \leq T} \frac{q(|m|)}{|m|^{2p}} \, \big|f_\mu(m) - f_\nu(m)\big|^2\bigg)^{1/2} + \frac{1}{T^p}\,. \tag{17.2}$$

If μ and ν are supported on $[0,\pi]^d$ and have equal mixed moments up to order k, then similar bounds hold true for ζ_p in place of ζ_p^*.

Strictly speaking, it should be assumed in the preceding statement that $T \geq \frac{1}{\pi}$, in which case $\delta \leq \pi$. But (17.1) and (17.2) are automatically fulfilled for $T < \frac{1}{\pi}$ since the periodic Zolotarev distances between probability measures are bounded by (p,d)-constants (cf. (5.3)).

18 Application to empirical measures on the torus ($p \leq \frac{d}{2}$)

With the help of the improved smoothed inequalities of Theorem 17.1, the applications to empirical measures may finally be developed, with the extension of Theorem 11.1 of Section 11 to the range $p \leq \frac{d}{2}$.

The notation are the same as in Section 11, with the empirical measures

$$\mu_n \, = \, \frac{1}{n}\sum_{j=1}^n \delta_{X_j}, \quad \nu_n \, = \, \frac{1}{n}\sum_{j=1}^n \delta_{Y_j},$$

constructed over samples $X = (X_1,\dots,X_n)$ and $Y = (Y_1,\dots,Y_n)$ of random variables (defined on some probability space $(\Omega,\mathcal{A},\mathbb{P})$) taking values in the torus Q^d.

Theorem 18.1 (Rate of convergence in ζ_p^* for $p \leq \frac{d}{2}$ integer)**.** *Let $X_1,\dots,X_n,Y_1,\dots,Y_n$ be random variables in Q^d such that the couples (X_j,Y_j) and (X_k,Y_k) are independent for $j \neq k$ and such that X_j and Y_j have equal distribution for every $j = 1,\dots,n$. Suppose that p is an integer. If $1 \leq p < \frac{d}{2}$, then*

$$\mathbb{E}\big(\zeta_p^*(\mu_n,\nu_n)\big) \, \leq \, \frac{C_{p,d}}{n^{p/d}}\,, \tag{18.1}$$

where $C_{p,d} > 0$ depends on (p,d) only, and in the case $p = \frac{d}{2}$ (with even d),

$$\mathbb{E}\big(\zeta_p^*(\mu_n,\nu_n)\big) \, \leq \, C_d\, \frac{\log n}{\sqrt{n}}\,. \tag{18.2}$$

If $X_1, \dots, X_n, Y_1, \dots, Y_n$ are independent and still X_j and Y_j are equi-distributed for every j, then similar bounds hold for the ψ_2-norm of the distance $\zeta_p^*(\mu_n, \nu_n)$. If the random variables $X_1, \dots, X_n$ are independent with a common law μ, the same bounds hold true for $\mathbb{E}(\zeta_p^*(\mu_n, \mu))$.

Theorem 18.2 (Rate of convergence in ζ_p^* for $p \leq \frac{d}{2}$ fractional)**.** *Let $X_1, \dots, X_n, Y_1, \dots, Y_n$ be random variables in Q^d such that the couples (X_j, Y_j) and (X_k, Y_k) are independent for $j \neq k$ and such that X_j and Y_j have equal distribution for every $j = 1, \dots, n$. Suppose that p is not an integer. If $0 < p \leq \frac{d}{2}$, then*

$$\mathbb{E}\big(\zeta_p^*(\mu_n, \nu_n)\big) \, \leq \, C_{p,d} \, \frac{\log n}{n^{p/d}} \, , \tag{18.3}$$

where $C_{p,d} > 0$ depends on (p,d) only. If $X_1, \dots, X_n, Y_1, \dots, Y_n$ are independent and still X_j and Y_j are equi-distributed for every j, then similar bounds hold for the ψ_2-norm of the distance $\zeta_p^*(\mu_n, \nu_n)$. If the random variables $X_1, \dots, X_n$ are independent with same law μ, the inequality (18.3) *also holds true for $\mathbb{E}(\zeta_p^*(\mu_n, \mu))$.*

In the critical case $p = \frac{d}{2}$, the bounds (18.2) and (18.3) coincide, while in the general situation with $p < \frac{d}{2}$, a non-integer value of p leads to the additional logarithmic term. For the matter of comparison, the known estimates in Kantorovich distances (cf. [8], [10]) in the range $1 \leq p < \frac{d}{2}$ are similar,

$$\mathbb{E}\big(W_p^p(\mu_n, \nu_n)\big) \, \leq \, C_{p,d} \, \frac{1}{n^{p/d}} \, ,$$

with an additional $\log n$ factor when $p = \frac{d}{2}$.

Proofs (of Theorems 18.1 and 18.2). As in the proof of Theorem 11.1, for any $m \in \mathbb{Z}^d$,

$$f_{\mu_n}(m) - f_{\nu_n}(m) \, = \, \frac{1}{n} \sum_{j=1}^n \big(e^{im \cdot X_j} - e^{im \cdot Y_j}\big)$$

and

$$\mathbb{E}\big(\big|f_{\mu_n}(m) - f_{\nu_n}(m)\big|^2\big) \, = \, \frac{1}{n^2} \sum_{j=1}^n \mathbb{E}\big(\big|e^{im \cdot X_j} - e^{im \cdot Y_j}\big|^2\big) \, \leq \, \frac{4}{n} \, .$$

An application of (17.1) with $T \geq 1$ and an integer p leads to

$$c_{p,d} \, \mathbb{E}\big(\zeta_p^*(\mu_n, \nu_n)\big) \, \leq \, \frac{1}{\sqrt{n}} \bigg(\sum_{1 \leq \|m\|_\infty \leq T} \frac{1}{|m|^{2p}}\bigg)^{1/2} + \frac{1}{T^p} \, . \tag{18.4}$$

If $p < \frac{d}{2}$, the above sum is of the order

$$\int_{1 \leq |x| \leq T} \frac{dx}{|x|^{2p}} \, \sim \, \int_1^T r^{d-1} \, \frac{dr}{r^{2p}} \, \sim \, T^{d-2p},$$

leading to

$$c_{p,d} \, \mathbb{E}\big(\zeta_p^*(\mu_n, \nu_n)\big) \, \leq \, \frac{1}{\sqrt{n}} \, T^{\frac{d}{2}-p} + \frac{1}{T^p} \, .$$

The choice $T = n^{1/d}$ yields the required inequality (18.1) in this case.

If $p = \frac{d}{2}$, the sum in (18.4) is of the order of

$$\int_{1 \le |x|_\infty \le T} \frac{dx}{|x|^{2p}} \sim \int_1^T r^{d-1} \frac{dr}{r^{2p}} \sim \log T,$$

leading to

$$c_{p,d}\, \mathbb{E}\big(\zeta_p^*(\mu_n,\nu_n)\big) \le \frac{1}{\sqrt{n}} \log T + \frac{1}{T^{d/2}}.$$

The same choice $T = n^{1/d}$ yields the required inequality (18.2) in this case.

When p is not an integer, apply (17.2) with $T \ge 1$ to obtain a slightly weaker variant of (18.4) with a weight q, namely

$$c_{p,d}\, \mathbb{E}\big(\zeta_p^*(\mu_n,\nu_n)\big) \le \frac{C_q}{\sqrt{n}} \bigg(\sum_{1 \le \|m\|_\infty \le T} \frac{q(|m|)}{|m|^{2p}}\bigg)^{1/2} + \frac{1}{T^p}.$$

With the choice of $q(s) = \log^2(2s)$, $s \ge 1$, using $q(|m|) \le q(T) \le \log^2(2T)$ for the case $p < \frac{d}{2}$, the desired inequality (18.3) is established with the same arguments as in the case of integer values of p. When $p = \frac{d}{2}$, the above sum is of the order

$$\int_{1 \le |x| \le T} q(|x|)\, \frac{dx}{|x|^{2p}} \sim \int_1^T q(r)\, \frac{dr}{r}.$$

The choice of the constant function $q(s) = 1$ for $1 \le s \le T$ with $T \ge 2$ and $q(s) = \infty$ for $s > T$, in which case the last integral is of the order $\log T$, yields (cf. (9.2))

$$C_q^2 = \sum_{\ell=0}^\infty \frac{1}{q(2^\ell)} \le C \log T$$

with some absolute constant C. Choosing $T = 2n^{1/d}$ implies again (18.3).

The last claims of Theorems 18.1 and 18.2 follow as for Theorem 11.1. □

Acknowledgment. The research of S.B. was partially supported by NSF grant DMS-2154001.

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. Closeness of probability distributions in terms of Fourier-Stieltjes transforms. Uspekhi Mat. Nauk 71 (2016), no. 6, 37–98. Russian Math. Surveys 71 (2016), no. 6, 1021–1079.

[3] Bobkov, S. G. Zolotarev-type distances. Preprint (2023).

[4] 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.

[5] Bobkov, S. G.; Ledoux, M. Transport inequalities on Euclidean spaces for non-Euclidean metrics. J. Fourier Anal. Appl. 26 (2020), no. 4, Paper No. 60, 27 pp.

[6] Bobkov, S. G.; Ledoux, M. Erratum to the article "Transport inequalities on Euclidean spaces for non-Euclidean metrics J. Fourier Anal. Appl. 26 (2020), no. 4, paper no. 60, 27 pp." Submitted to: J. Fourier Anal. Appl. (2023).

[7] Bobkov, S. G.; Ledoux, M. A simple Fourier analytic proof of the AKT optimal matching theorem. Ann. Appl. Probab. 31 (2021), no. 6, 2567–2584.

[8] Dereich, S.; Scheutzow, M. and Schottstedt, R. Constructive quantization: approximation by empirical measures. Ann. Inst. H. Poincaré Probab. Statist. 49 (2013), 1183–1203.

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

[10] Fournier, N.; Guillin, A. On the rate of convergence in Wasserstein distance of the empirical measure. Probab. Theory Related Fields 162 (2015), 707–738.

[11] Gautschi, W. On inverses of Vandermonde and confluent Vandermonde matrices. Numer. Math. 4 (1962), 117–123.

[12] Ledoux, M. On optimal matching of Gaussian samples. Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 457, Veroyatnost i Statistika. 25 (2017), 226–264; translation in J. Math. Sci. (N.Y.) 238 (2019), no. 4, 495–522.

[13] Mironescu, P; Sickel, W. A Sobolev non-embedding. Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl. 26 (2015), no. 3, 291–298.

[14] Rio, E. Upper bounds for minimal distances in the central limit theorem. Ann. Inst. Henri Poincaré Probab. Stat. 45 (2009), no. 3, 802–817.

[15] Schmeisser, H-J.; Triebel, H. Topics in Fourier analysis and function spaces. Wiley-Intersci. Publ. John Wiley & Sons, Ltd., Chichester, 1987, 300 pp.

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

[17] 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.

[18] 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.

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

SB: School of Mathematics, University of Minnesota, Minneapolis, MN 55455 USA, bobkov@math.umn.edu

ML: Institut de Mathématiques de Toulouse, Université de Toulouse – Paul-Sabatier, F-31062 Toulouse, France, ledoux@math.univ-toulouse.fr