HIGH-DIMENSIONAL PHENOMENA: DILATIONS, TENSOR PRODUCTS AND GEOMETRY OF L_1

Tomasz Tkocz

A thesis submitted for the degree of

Doctor of Philosophy

University of Warwick

Mathematics Institute

March 2015

Contents

Acknowledgements

I am deeply indebted to my supervisor, Keith Ball, for his invaluable advice and encouragement. He has constantly demonstrated great passion, joy and enthusiasm for mathematics, which has been motivating.

I am also grateful to Rafał Latała who supervised the first year of my PhD at Warsaw as well as Neil O'Connell who co-supervised my first year at Warwick.

I would like to thank Gideon Schechtman and Ofer Zeitouni for their great help, especially during my summer projects at Weizmann.

I wish to express my gratitude to Marek Kuś and Karol Życzkowski for introducing me to problems concerning tensor products of random unitary matrices. I very much appreciate all the illuminating conversations I have had with Radosław Adamczak, Matthieu Fradelizi, Olivier Guédon, Stanisław Kwapień and Krzysztof Oleszkiewicz.

I would also like to thank the Polish National Science Centre (grant NCN 2011/01/N/ST1/05960) as well as the Warwick Mathematics Institute for their financial support without which I would not have been able to attend many conferences.

Of many people to whom credit is due, I would like to especially thank Karol Baron for his great help and guidance at the very beginning of my mathematical journey.

Many thanks to my collaborator and friend, Piotr Nayar, with whom doing mathematics is such good fun!

Finally, I wish to offer my sincere thanks to my parents and fiancée.

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

Parts of this thesis have been published: [JNST14], [NT13], [NT14a], [NT14b], [Tko13], [TSK$^+$12]. All but one of these articles is written jointly with other researchers. As is customary in mathematics, no details are provided within these articles of which author contributed which ideas. However, we have included some comments of this type at the end of each chapter.

Abstract

The purpose of this dissertation is to study several problems related to high-dimensional phenomena in analysis, geometry and probability.

The first problem examines the behaviour of probability measures of dilations of sets possessing certain symmetries. We show that for the standard Gaussian measure on complex vector space, cylinders are optimal in the sense that, under dilations, the Gaussian measure grows no more rapidly for cylinders than for other domains possessing enough symmetries. We also prove an analogous result in the real case for Weibull and Gamma distributions. As a consequence, we derive optimal comparison of moments for these distributions.

The second problem stems from the study of composite periodic quantum systems. It asks about the behaviour of certain random matrices when their size tends to infinity. We show that the spectrum of the tensor product of two large random unitary matrices is asymptotically Poissonian; what we would expect for diagonal matrices. The same conclusion is established for the tensor product of a large number of 2×2 random unitary matrices.

The third problem concerns the invertibility of operators on L_1. We construct an example of a locally invertible operator with kernel of arbitrarily large dimension. The construction is combinatorial, relying on expander graphs and recent results from computer science about the restricted isometry property on ℓ_1. We also establish some Sobolev-type inequalities and find a certain large class of convolution operators which are globally invertible on large subspaces.

Preface

This thesis concerns several topics lying at the intersection of analysis and geometry, each of which has a probabilistic flavour. We are interested in two types of situation: (A) one wants to find generic properties of objects that hold in any dimension, and (B) one wants to describe asymptotic properties of objects if the dimension tends to infinity.

The first problem under consideration examines the behaviour of probability measures of dilations of convex sets and falls into paradigm (A). In 1969, L. Shepp made, and in 1999, R. Latała and K. Oleszkiewicz proved a conjecture saying that under dilations, a Gaussian measure on a *real* Banach space grows no more rapidly for strips than for other symmetric convex sets. This result has its roots in the study of Gaussian measures on Banach spaces, where one often needs precise estimates for tails of distributions of norms of Gaussian vectors. How to use the assumption of convexity on Gaussian space is a challenge, since the intrinsic properties of the standard Gaussian measure, the rotational invariance as well as the product structure at the same time, do not seem to be intimately related to convexity. Natural questions that arise are these: what can be said for Gaussian measures on *complex* Banach spaces, and what sets are optimal for other measures? We address both of them, prompted by the idea that for product measures, it is often possible to reduce the problem to low dimensions via an inductive sort of argument. In the case of Gaussian measures on complex space, after linearisation and dimension-reduction, we conclude the argument by proving a simple, new, entropic functional inequality. In the real case, for symmetric Weibull and Gamma distributions, our strategy is the same — first we reduce the problem to a problem in 2 dimensions, and then we show how it can be solved using some inequalities for one dimensional functionals involving convexity.

Our second problem emerges from composite quantum systems modelled by the tensor products of the spaces relating to their subsystems. Since in quantum mechanics, dynamics is governed by unitary operators, for a composite system, the relevant object is a tensor product of unitary matrices. If the dynamics is generic, we take random unitary matrices to model the evolution, so we study the spectra of tensor products of random unitary matrices. We focus on asymptotic properties, as the size of the matrix becomes large (paradigm (B)). We mainly look at two situations: 1) the tensor product of two independent random unitary matrices with the size of at least one of them tending to infinity, and 2) the tensor product of a large number of independent random unitary matrices of a fixed size. The main goal is to show how the spectra of such matrices behave asymptotically.

The third problem concerns the geometry of L_1 spaces. A bounded linear operator $T\colon L_1\big([0,1]\big) \longrightarrow L_1\big([0,1]\big)$ is said to be ϵ*-locally invertible* if for every measurable subset A of $[0,1]$ of Lebesgue measure at most a half, the restricted operator $T|_{L_1(A)}$ is invertible with $\big\|(T|_{L_1(A)})^{-1}\big\| \geq \epsilon^{-1}$. The main question we are concerned with reads as follows. *Suppose that T is a bounded linear operator on $L_1\big([0,1]\big)$ which is ϵ-locally invertible. Is T invertible when restricted to a subspace of finite codimension?* We answer this question in the negative. Our approach hinges very much on studying related aspects for finite dimensional ℓ_1 spaces. We make use of magical combinatorial properties of expanders that have recently been discovered in the context of sparse signal recovery. In a sense, our work draws attention to a beautiful interplay between combinatorial properties of finite-dimensional objects and their infinite-dimensional counterparts, with the emphasis on what happens when the dimension tends to infinity. This closely follows the spirit of the local theory of Banach spaces, a branch of geometric functional analysis that has developed rapidly in the last few decades. Having said that, of course, we should also remark that the problem discussed here can be seen as of type (B).

This thesis comprises three chapters, each devoted to one of the problems mentioned above. The chapters are independent and can be regarded as individual pieces of work. What binds them together are paradigms (A) and (B) in the study of high-dimensional phenomena.

Throughout the text we try to use standard notation. Unclear or ambiguous symbols are explained as they appear. Each chapter finishes with a *notes & comments* section which gives the origin of every theorem stated as well as the author's contribution.

Chapter 1

Measures of dilations

1.1 Introduction

Let γ_n be the standard Gaussian measure on $\mathbb{R}^n$, i.e. the measure with density at $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ given by $\frac{1}{\sqrt{2\pi}^n} e^{-|x|^2/2}$, where $|x| = \sqrt{x_1^2 + \ldots + x_n^2}$ denotes the Euclidean norm. Take a Borel subset A of $\mathbb{R}^n$ and expand it by dilating, that is consider the set $tA = \{ta,\ a \in A\}$, $t \geq 1$. We shall ask vaguely: *how fast does the function $t \mapsto \gamma_n(tA)$ grow?* It is fairly easy to see that one extreme case is a Euclidean ball. Indeed, let B be the closed Euclidean ball centred at the origin with radius r chosen so that $\gamma_n(B) = \gamma_n(A)$. In particular, $\gamma_n(A \setminus B) = \gamma_n(B \setminus A)$ and, by simply moving mass to where the density is bigger, we get that for $t \geq 1$,

$$\begin{aligned}
\gamma_n\big(t(A \setminus B)\big) &= \int_{t(A\setminus B)} e^{-|x|^2/2} \frac{\mathrm{d}x}{\sqrt{2\pi}^n} = t^n \int_{A\setminus B} e^{-t^2|x|^2/2} \frac{\mathrm{d}x}{\sqrt{2\pi}^n} \\
&= t^n \int_{A\setminus B} e^{-(t^2-1)|x|^2/2} \mathrm{d}\gamma_n(x) \leq t^n \int_{A\setminus B} e^{-(t^2-1)r^2/2} \mathrm{d}\gamma_n(x) \\
&= t^n \int_{B\setminus A} e^{-(t^2-1)r^2/2} \mathrm{d}\gamma_n(x) \leq t^n \int_{B\setminus A} e^{-(t^2-1)|x|^2/2} \mathrm{d}\gamma_n(x) \\
&= \gamma_n\big(t(B \setminus A)\big).
\end{aligned}$$

Therefore,

$$\gamma_n(tA) \leq \gamma_n(tB), \qquad t \geq 1.$$

A similar argument also shows that

$$\gamma_n(tA) \geq \gamma_n(tB), \qquad t \leq 1.$$

These two inequalities show that under dilations, the Gaussian measure grows no faster for any set than it does for a ball. The natural question that arises is this: what is the other extreme case? In other words, the Gaussian measure of which sets grows most slowly. In 1969, in an unpublished preprint, L. Shepp made the following conjecture.

1.1 Conjecture (Shepp)**.** *Let K be a convex symmetric ($K = -K$) subset in $\mathbb{R}^n$ and let $P = \{x \in \mathbb{R}^n,\ |x_1| \leq p\}$ be a strip with width p chosen so that $\gamma_n(P) = \gamma_n(K)$; then*

$$\begin{aligned} \gamma_n(tK) &\geq \gamma_n(tP), \qquad t \geq 1, \\ \gamma_n(tK) &\leq \gamma_n(tP), \qquad t \leq 1. \end{aligned} \tag{1.1}$$

It should be remarked here that strips are no longer optimal in the wider class of symmetric sets — the above inequalities are not true after dropping the assumption of convexity. For instance, they do not hold for the *cross* $K = ([-1,1] \times \mathbb{R}) \cup (\mathbb{R} \times [-1,1])$ in the plane.

In the form stated above, the conjecture was first published in [Sza91] (see Remark 2.7 therein). V. Zalgaller and V. Sudakov showed that it holds for $n = 3$ (see [ZS74]). S. Kwapień and J. Sawa proved the conjecture under the additional assumption that K is symmetric with respect to every hyperplane $\{x_i = 0\}$, $i = 1, \ldots, n$ (see [KS93]). Not until 30 years after it had been stated, was Shepp's conjecture proved in full generality, by R. Latała and K. Oleszkiewicz (see [LO99]). Their result is sometimes referred to as the *S-inequality.*

There are some natural further directions of research. Shepp's preprint was concerned with the existence of strong exponential moments of a Gaussian measure on a Banach space. Suppose that X is a (centred) Gaussian vector on a Banach space $(F, \|\cdot\|)$ distributed according to a Gaussian measure μ on F. One often needs precise estimates on the quantity $\mathbb{P}(\|X\| > t)$ which is simply $1 - \mu(tK)$, where K denotes the unit ball in F. Such a ball is a convex and symmetric set. Now, if F is a real Banach space, certain approximation techniques allow one to reduce the situation to the simplest case of the standard Gaussian measure γ_n on $\mathbb{R}^n$ and estimates like the S-inequality yield optimal bounds for $\mathbb{P}(\|X\| > t)$ (for details see, e.g., [KS93]). The same paradigm applies when F is a complex Banach space. Hence one of the interesting questions is to find a version of the S-inequality for the standard Gaussian distribution on $\mathbb{C}^n$.

Another natural question is this: are there any other measures for which the S-inequality holds (strips are optimal)? R. Latała made the following conjecture (see survey [Lat02])

1.2 Conjecture (Latała)**.** *Let ν be a rotationally invariant measure on $\mathbb{R}^n$, absolutely continuous with respect to the Lebesgue measure with a density of the form $f(|x|)$ for some nonincreasing function $f\colon [0,\infty) \longrightarrow [0,\infty)$. Then for any convex symmetric set A in $\mathbb{R}^n$ and any symmetric strip P in $\mathbb{R}^n$ such that $\nu(A) = \nu(P)$ the inequality $\nu(tA) \geq \nu(tP)$ is satisfied for $t \geq 1$.*

It was proved that the conjecture holds for $n \leq 3$ (see [ZS74]). To the best of our knowledge, this is the only known result addressing Latała's conjecture in its full generality.

The next sections are devoted to the complex counterpart of the S-inequality as well as its extensions to some other measures in the real case (products of symmetric Gamma and Weibull distributions). We also present applications of S-inequalities to the derivation of optimal comparison of moments.

1.2 Gaussian measures on complex space

1.2.1 Preliminaries

We define the standard Gaussian measure ν_n on the space $\mathbb{C}^n$ via the formula

$$\nu_n(A) = \gamma_{2n}(\tau(A)), \qquad \text{for any Borel set } A \subset \mathbb{C}^n,$$

where $\mathbb{C}^n \xrightarrow{\tau} \mathbb{R}^{2n}$ is the bijection given by

$$\tau(z_1, \ldots, z_n) = (\mathfrak{Re} z_1, \mathfrak{Im} z_1, \ldots, \mathfrak{Re} z_n, \mathfrak{Im} z_n).$$

We say that a closed subset K of $\mathbb{C}^n$ *supports the complex S-inequality*, *$S\mathbb{C}$-inequality* for short, if for every $s > 0$ the dilation $L = sK$ and every *cylinder* $C = \{z \in \mathbb{C}^n,\ |z_1| \leq R\}$ satisfy

$$\nu_n(L) = \nu_n(C) \quad \Longrightarrow \quad \nu_n(tL) \geq \nu_n(tC), \quad \text{for } t \geq 1. \tag{1.2}$$

A subset K of $\mathbb{C}^n$ is called *circled* if $e^{i\theta}K = K$ for every $\theta \in \mathbb{R}$. A natural counterpart of the S-inequality, (1.1), in the complex case is the following conjecture due to A. Pełczyński.

1.3 Conjecture (Pełczyński)**.** *All convex subsets K of $\mathbb{C}^n$ which are* circled *support the $S\mathbb{C}$-inequality.*

Following the methods from [LO99], the author in his master's thesis obtained a partial result saying that all convex circled sets support the $S\mathbb{C}$-inequality as long as they are not too *big*. More precisely, he showed in [Tko11] that there exists a universal constant $c > 0.64$ such that for every convex circled subset K of $\mathbb{C}^n$ with $\nu_n(K) < c$, if C is a cylinder of the same ν_n measure as K, then

$$\nu_n(tK) \geq \nu_n(tC), \qquad \text{for every } t \in [1, t_0],$$

where t_0 is determined by the condition $\nu_n(t_0 K) = c$.

Note that a unit ball with respect to a norm on $\mathbb{C}^n$ is a convex and circled set. We are interested in the class $\mathfrak{R}$ of all closed sets in $\mathbb{C}^n$ which are *Reinhardt complete*, i.e. along with each point $(z_1, \ldots, z_n)$ such a set contains all points $(w_1, \ldots, w_n)$ for which $|w_k| \leq |z_k|$, $k = 1, \ldots, n$ (consult for instance the textbook [Sha92, I.1.2, pp. 8–9]). Sets from the class $\mathfrak{R}$ are not necessarily convex (e.g. $\{(z_1, z_2) \in \mathbb{C}^2,\ |z_1|^{1/2} + |z_2|^{1/2} \leq 1\}$). For us, it is important that this class contains all unit balls with respect to unconditional norms on $\mathbb{C}^n$. Recall that a norm $\|\cdot\|$ is said to be *unconditional* if $\|(e^{i\theta_1}z_1, \ldots, e^{i\theta_n}z_n)\| = \|z\|$ for all $z \in \mathbb{C}^n$ and $\theta_1, \ldots, \theta_n \in \mathbb{R}$. Our main result reads

1.4 Theorem. *Every set from the class $\mathfrak{R}$ supports the $S\mathbb{C}$-inequality.*

Now we establish some simple general observations which allow us to reduce the problem to a one-dimensional entropy inequality. This inequality, which may be of independent interest, is proved in the next subsection. Then we prove the main theorem. In the last subsection we discuss its corollaries.

1.5 Proposition. *A closed subset K of $\mathbb{C}^n$ supports the $S\mathbb{C}$-inequality if and only if for every $s > 0$ the dilation $L = sK$ and every cylinder C satisfy*

$$\nu_n(L) = \nu_n(C) \quad \Longrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}\nu_n(tL)\Big|_{t=1} \geq \frac{\mathrm{d}}{\mathrm{d}t}\nu_n(tC)\Big|_{t=1}. \tag{1.3}$$

Proof. We only show the interesting part that (1.3) implies (1.2) following the proof of [KS93, Lemma 1]. Fix a dilation L of K and a cylinder C such that $\nu_n(L) = \nu_n(C)$. Let a function h be given by $\nu_n(tL) = \nu_n(h(t)C)$, $t \geq 1$. Then, by the assumption, we find

$$h(t)\frac{\mathrm{d}}{\mathrm{d}s}\nu_n(sC)\Big|_{s=h(t)} = \frac{\mathrm{d}}{\mathrm{d}s}\nu_n(sh(t)C)\Big|_{s=1} \leq \frac{\mathrm{d}}{\mathrm{d}s}\nu_n(stL)\Big|_{s=1} = t\frac{\mathrm{d}}{\mathrm{d}s}\nu_n(sL)\Big|_{s=t}.$$

Differentiating the equation which defines the function h yields $\frac{\mathrm{d}}{\mathrm{d}s}\nu_n(sL)\big|_{s=t} = h'(t)\frac{\mathrm{d}}{\mathrm{d}s}\nu_n(sC)\big|_{s=h(t)}$, thus $h(t) \leq th'(t)$. This means that the function $h(t)/t$ is nondecreasing, so $1 = h(1) \leq h(t)/t$ for $t \geq 1$. $\square$

For any closed set A the derivative of the function $t \mapsto \nu_n(tA)$ is easy to compute. Indeed,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\nu_n(tA)\Big|_{t=1} &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{tA} e^{-|z|^2/2}\mathrm{d}z\Big|_{t=1} = \frac{\mathrm{d}}{\mathrm{d}t}\int_A t^{2n}e^{-t^2|w|^2/2}\mathrm{d}w\Big|_{t=1} \\ &= 2n\nu_n(A) - \int_A |z|^2\mathrm{d}\nu_n(z).\end{aligned}$$

Moreover, the integral of $|z|^2$ over a cylinder C may be expressed explicitly in terms of the measure $\nu_n(C)$. Namely,

$$\int_C |z|^2\mathrm{d}\nu_n(z) = 2(1-\nu_n(C))\ln(1-\nu_n(C)) + 2n\nu_n(C).$$

Combining these two remarks with the preceding proposition we obtain an equivalent formulation of the problem.

1.6 Proposition. *A closed subset K of $\mathbb{C}^n$ supports the $S\mathbb{C}$-inequality if and only if for every $s > 0$ the dilation $L = sK$ satisfies*

$$\int_L |z|^2 \mathrm{d}\nu_n(z) \le 2np + 2(1-p)\ln(1-p), \tag{1.4}$$

where $p = \nu_n(L)$ is the measure of the dilation L.

1.2.2 A one-dimensional entropy inequality

Observe that the quantity $x \ln x$ with $x = 1 - p$ appears in (1.4). It is not surprising that entropy will play a role in the rest of our proof. Recall that the *entropy* of a function $f\colon X \longrightarrow [0, \infty)$ with respect to a probability measure μ on a measurable space X is defined by

$$\begin{aligned} \mathrm{Ent}_\mu f = &\int_X f(x) \ln f(x) \mathrm{d}\mu(x) \\ &- \left(\int_X f(x)\mathrm{d}\mu(x)\right) \ln \left(\int_X f(x)\mathrm{d}\mu(x)\right). \end{aligned} \tag{1.5}$$

We adopt the standard convention that $0 \ln 0 = 0$.

The following simple one-dimensional entropy inequality is an important ingredient in the proof of the main theorem, Theorem 1.4.

1.7 Lemma. *Let μ be a Borel probability measure on $[0, \infty)$ and suppose $f\colon [0, \infty) \longrightarrow [0, \infty)$ is a bounded and nondecreasing function. Then*

$$\mathrm{Ent}_\mu f \le -\int_0^\infty f(x)\Big(1 + \ln \mu\left((x, \infty)\right)\Big)\mathrm{d}\mu(x). \tag{1.6}$$

Proof. Using homogeneity of both sides of (1.6), we can assume without loss of generality that $\int_0^\infty f \mathrm{d}\mu = 1$. Then we may rewrite the assertion of the lemma as follows

$$\int_0^\infty \ln\left(f(x) \int_{(x,\infty)} \mathrm{d}\mu(t)\right) f(x)\mathrm{d}\mu(x) \le -1. \tag{1.7}$$

Introduce the probability measure ν on $[0, \infty)$ with the density f with respect to μ. Thanks to the monotonicity of f we can bound the left hand side of the last inequality by

$$\int_0^\infty \ln\Big(\nu\left((x, \infty)\right)\Big)\mathrm{d}\nu(x) = -\int_0^\infty \int_0^1 \frac{\mathrm{d}u}{u} \mathbf{1}_{\{u \ge \nu((x,\infty))\}}(u, x)\mathrm{d}\nu(x).$$

Define the function

$$H(y) = \inf \{t,\ \nu((t,\infty)) \le y\},$$

which is the *inverse* tail function, and observe that

$$\{(u,x),\ u \ge \nu((x,\infty))\} \supset \{(u,x),\ H(u) \le x\},$$

as $u \ge \nu((H(u),\infty)) \ge \nu((x,\infty))$. This leads to

$$\begin{aligned} -\int_0^\infty \int_0^1 \frac{\mathrm{d}u}{u} \mathbf{1}_{\{u \ge \nu((x,\infty))\}}(u,x)\mathrm{d}\nu(x) &\le -\int_0^\infty \int_0^1 \frac{\mathrm{d}u}{u} \mathbf{1}_{\{H(u)\le x\}}(u,x)\mathrm{d}\nu(x) \\ &= -\int_0^1 \nu([H(u),\infty)) \frac{\mathrm{d}u}{u}. \end{aligned}$$

Since $u \le \nu([H(u),\infty))$, we finally get the desired estimate. □

1.8 Remark. If μ has a density, say g, the proof can be rewritten as follows. For $t \ge x$ the monotonicity of f yields $f(x) \le f(t)$, so $f(x)\int_{(0,\infty)} \mathrm{d}\mu(t) \le \int_x^\infty f(t)g(t)\mathrm{d}t = F(x)$ and the right-hand side of inequality (1.7) can be bounded above by

$$\begin{aligned} \int_0^\infty \big(\ln F(x)\big)\big(-F'(x)\big)\mathrm{d}x &= -F\ln F\Big|_0^\infty + \int_0^\infty \frac{F'(x)}{F(x)}F(x)\mathrm{d}x \\ &= F\Big|_0^\infty = -1, \end{aligned}$$

as $F(0) = \int_0^\infty f\mathrm{d}\mu = 1$ and F vanishes at infinity.

1.2.3 Proof of the main result

We shall need a multidimensional version of Lemma 1.7 for product measures. To establish it, we shall exploit the product structure. For simplicity, we formulate this result for the Gaussian measure.

1.9 Lemma. *Let $g\colon \mathbb{C}^n \longrightarrow [0,\infty)$ be a bounded function satisfying*

1) $g((e^{i\theta_1}z_1,\ldots,e^{i\theta_n}z_n)) = g(z)$ for any $z \in \mathbb{C}^n$ and $\theta_1,\ldots,\theta_n \in \mathbb{R}$,

2) for any $w,z \in \mathbb{C}^n$ the condition $|w_k| \le |z_k|$, $k = 1,\ldots,n$ implies $g(w) \le g(z)$.

Then

$$\operatorname{Ent}_{\nu_n} g \leq \int_{\mathbb{C}^n} g(z) \left(\frac{|z|^2}{2} - n \right) \mathrm{d}\nu_n(z). \tag{1.8}$$

Proof. For a fixed vector $r = (r_1, \dots, r_n) \in [0, \infty)^n$ we denote

$$r^{(k)} = (r_1, \dots, r_{k-1}, r_{k+1}, \dots, r_n) \in [0, \infty)^{n-1},$$

and then define the functions

$$g_k^r(x) = g(r_1, \dots, r_{k-1}, x, r_{k+1}, \dots, r_n), \qquad k = 1, \dots, n.$$

Notice that for a function $h \colon \mathbb{C} \longrightarrow [0, \infty)$ obeying property 1) we get

$$\int_{\mathbb{C}} h(z) \mathrm{d}\nu_1(z) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} h(te^{i\theta}) e^{-t^2/2} t \mathrm{d}t \mathrm{d}\theta = \int_0^{\infty} h(t) \mathrm{d}\mu(t),$$

where μ denotes the probability measure on $[0, \infty)$ with the density at t given by $te^{-t^2/2}$. Therefore,

$$\begin{aligned} \int_{\mathbb{C}^n} g(z) \left(\frac{|z|^2}{2} - n \right) \mathrm{d}\nu_n(z) &= \int_{[0,\infty)^n} g(r) \left(\frac{\sum_{k=1}^n r_k^2}{2} - n \right) \mathrm{d}\mu^{\otimes n}(r) \\ &= \int_{[0,\infty)^n} \sum_{k=1}^{n} \left[\int_{[0,\infty)} g_j^r(x) \left(\frac{x^2}{2} - 1 \right) \mathrm{d}\mu(x) \right] \mathrm{d}\mu^{\otimes n}(r), \end{aligned}$$

where $\mu^{\otimes n} = \mu \otimes \ldots \otimes \mu$ denotes the product measure. Applying Lemma 1.7 for the function g_j^r and the measure μ we obtain the estimate

$$\begin{aligned} \int_{\mathbb{C}^n} g(z) \left(\frac{|z|^2}{2} - n \right) \mathrm{d}\nu_n(z) &\geq \int_{[0,\infty)^n} \sum_{k=1}^{n} \operatorname{Ent}_{\mu}(g_j^r) \mathrm{d}\mu^{\otimes n}(r) \\ &\geq \operatorname{Ent}_{\mu^{\otimes n}} g = \operatorname{Ent}_{\nu_n} g, \end{aligned}$$

where the last inequality follows from the subadditivity of entropy (see, e.g., [Led01, Proposition 5.6]). □

Proof of Theorem 1.4. Fix $K \in \mathfrak{R}$. We want to show (1.4), that is

$$\int_K |z|^2 \mathrm{d}\nu_n(z) \leq 2np + 2(1-p) \ln{(1-p)},$$

where $p = \nu_n(K)$ is the measure of K. The application of Lemma 1.9 for the function $g(z) = 1 - \mathbf{1}_K(z)$ yields

$$-(1-p)\ln(1-p) \leq -\int_K \frac{|z|^2}{2} \mathrm{d}\nu_n(z) + np,$$

which is what we want. □

1.2.4 Corollaries

Theorem 1.4 immediately implies that the Cartesian products of cylinders support the $S\mathbb{C}$-inequality. As a consequence, the $S\mathbb{C}$-inequality possesses a tensorization property.

1.10 Corollary. *Let sets $K_1 \subset \mathbb{C}^{n_1}, \ldots, K_\ell \subset \mathbb{C}^{n_\ell}$ support the $S\mathbb{C}$-inequality. Then the set $K_1 \times \ldots \times K_\ell$ also supports the $S\mathbb{C}$-inequality.*

Proof. Choose a cylinder C with the same measure as the Cartesian product $K_1 \times \ldots \times K_\ell = \prod K_i$ and choose cylinders C_i with the same measure as the sets K_i respectively. Then we have that C and $\prod C_i$ have the same measure,

$$\nu_n(C) = \nu_n\left(\prod K_i\right) = \prod \nu_{n_i}(K_i) = \prod \nu_{n_i}(C_i) = \nu_n\left(\prod C_i\right).$$

Since we assume that each K_i supports the $S\mathbb{C}$-inequality and, as we said, so does $\prod C_i$, we get that for $t \geq 1$,

$$\nu_n\left(t \prod K_i\right) = \prod \nu_{n_i}(tK_i) \geq \prod \nu_{n_i}(tC_i) = \nu_n\left(t \prod C_i\right) \geq \nu_n(tC),$$

hence $\prod K_i$ supports the $S\mathbb{C}$-inequality as well. □

Another consequence of the main theorem is related to the standard symmetric exponential measure λ_n on $\mathbb{R}^n$, i.e.

$$\mathrm{d}\lambda_n(x) = \frac{1}{2^n} e^{-|x|_1} \mathrm{d}x, \qquad x \in \mathbb{R}^n,$$

where we denote $|(x_1, \ldots, x_n)|_1 = \sum_{i=1}^n |x_i|$. It turns out that certain subsets of $\mathbb{R}^n$ support the S-inequality for λ_n with strips as the optimal sets. To state the result, we need a few definitions. We say that a set $K \subset [0, \infty)^n$ is a *down set* if for every point $x \in K$, the set K contains the cube $[0, x_1] \times \ldots \times [0, x_n]$. A set $K \subset \mathbb{R}^n$ is called *unconditional* if $(\epsilon_1 x_1, \ldots, \epsilon_n x_n) \in K$ whenever $(x_1, \ldots, x_n) \in K$ and $\epsilon_1, \ldots, \epsilon_n \in \{-1, 1\}$. By an *unconditional down set* K in $\mathbb{R}^n$ we mean the unconditional set K such that the set $K \cap [0, \infty)^n$ is a down set. For instance, any unconditional convex set is also an unconditional down set.

1.11 Theorem. *For any closed unconditional down set $K \subset \mathbb{R}^n$ and for any strip $P = \{x \in \mathbb{R}^n,\ |x_1| \le p\}$, $p \ge 0$, we have*

$$\lambda_n(K) = \lambda_n(P) \quad \Longrightarrow \quad \forall t \ge 1\ \lambda_n(tK) \ge \lambda_n(tP), \tag{1.9}$$

and, equivalently,

$$\lambda_n(K) = \lambda_n(P) \quad \Longrightarrow \quad \forall t \le 1\ \lambda_n(tK) \le \lambda_n(tP). \tag{1.10}$$

Proof. The equivalence of (1.9) and (1.10) is straightforward. For instance, assume the latter does not hold. Then, there is $t_0 < 1$ such that $\lambda_n(t_0K) > \lambda_n(t_0P)$. We can find $s_0 < 1$ for which $\lambda_n(s_0t_0K) = \lambda_n(t_0P)$. Using (1.9) we get a contradiction

$$\lambda_n(K) > \lambda_n(s_0K) = \lambda_n\left(\frac{1}{t_0}(s_0t_0K)\right) \ge \lambda_n\left(\frac{1}{t_0}(t_0P)\right) = \lambda_n(P) = \lambda_n(K).$$

Consider the mapping $F\colon \mathbb{C}^n \longrightarrow [0,\infty)^n$ given by the formula

$$F(z_1,\ldots,z_n) = (|z_1|,\ldots,|z_n|).$$

Observe that for a down set $A \subset [0,\infty)^n$, the set $F^{-1}(A)$ is Reinhardt complete and integrating using the polar coordinates we find that

$$\nu_n\left(F^{-1}(A)\right) = \int_A \prod_{i=1}^n r_i e^{-r_i^2/2}\mathrm{d}r_1\ldots\mathrm{d}r_n.$$

Now, let us change the variables according to the mapping $G\colon [0,\infty)^n \longrightarrow [0,\infty)^n$,

$$G(x_1,\ldots,x_n) = \frac{1}{2}(x_1^2,\ldots,x_n^2).$$

We obtain

$$\nu_n\left(F^{-1}(A)\right) = \int_{G(A)} e^{-\sum_{i=1}^n x_i}\mathrm{d}x.$$

Since $G(A)$ is a down set if and only if A is a down set, we infer that for any unconditional down set $K \subset \mathbb{R}^n$

$$\lambda_n(K) = \nu_n(\widetilde{K}), \qquad \text{where} \qquad \widetilde{K} := G^{-1}F^{-1}\left(K \cap [0,\infty)^n\right).$$

Moreover, for a strip $P = \{x \in \mathbb{R}^n,\ |x_1| \le p\}$, the set $\widetilde{P} \subset \mathbb{C}^n$ is a cylinder. Note also that $\widetilde{tK} = \sqrt{t}\widetilde{K}$. These observations combined with Theorem 1.4 yield the assertion. □

In the next section, we will see a generalization of this theorem. Now, we finish our discussion with a result concerning the optimal comparison of moments of unconditional norms for the exponential measure.

1.12 Corollary. *Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ which is unconditional, i.e.*

$$\|(\epsilon_1 x_1, \dots, \epsilon_n x_n)\| = \|(x_1, \dots, x_n)\|,$$

for any $x_j \in \mathbb{R}$ and $\epsilon_j \in \{-1, 1\}$. Then for $p \geq q > 0$

$$\left(\int_{\mathbb{R}^n} \|x\|^p \mathrm{d}\lambda_n(x)\right)^{1/p} \leq C_{p,q} \left(\int_{\mathbb{R}^n} \|x\|^q \mathrm{d}\lambda_n(x)\right)^{1/q}, \tag{1.11}$$

where the constant

$$C_{p,q} = \frac{\left(\int_{\mathbb{R}} |x|^p \mathrm{d}\lambda_1(x)\right)^{1/p}}{\left(\int_{\mathbb{R}} |x|^q \mathrm{d}\lambda_1(x)\right)^{1/q}} = \frac{(\Gamma(p+1))^{1/p}}{(\Gamma(q+1))^{1/q}}$$

is the best possible.

Proof. It is enough to repeat an argument credited to S. Szarek presented in detail in the proof of Corollary 3 in [LO99]. We can write

$$\int_{\mathbb{R}^n} \|x\|^p \mathrm{d}\lambda_n(x) = \int_0^\infty p t^{p-1} \lambda_n\big(K_t^c\big) \mathrm{d}t,$$

where $K_t = \{x \in \mathbb{R}^n, \ \|x\| \leq t\}$ is a closed convex unconditional set, hence an unconditional down set and Theorem 1.11 provides optimal bounds for its measure. Specifically, we can compare it with the measure of a strip $S_t = \{x \in \mathbb{R}^n, \ |x_1| \leq t\}$. Moreover, note that we have $\lambda_n(S_t^c) = \lambda_1\{x \in \mathbb{R}, \ |x| > t\}$.

The argument starts with choosing a parameter $\alpha > 0$ so that

$$\int_{\mathbb{R}^n} \|x\|^p \mathrm{d}\lambda_n(x) = \int_{\mathbb{R}} |x/\alpha|^p \mathrm{d}\lambda_1(x).$$

Rewriting yields

$$\int_0^\infty t^{p-1} \lambda_n(K_t^c) \mathrm{d}t = \int_0^\infty t^{p-1} \lambda_n(S_{\alpha t}^c) \mathrm{d}t,$$

so there is $t_0 > 0$ such that K_{t_0} and $S_{\alpha t_0}$ have the same measure λ_n. Hence we have $\lambda_n(K_t^c) \geq \lambda_n(S_{\alpha t}^c)$ for $t \leq t_0$, and $\lambda_n(K_t^c) \leq \lambda_n(S_{\alpha t}^c)$ for $t \geq t_0$. It follows that for $t > 0$ and $p \geq q > 0$

$$\left(\frac{t}{t_0}\right)^{p-1} \big(\lambda_n(K_t^c) - \lambda_n(S_{\alpha t}^c)\big) \leq \left(\frac{t}{t_0}\right)^{q-1} \big(\lambda_n(K_t^c) - \lambda_n(S_{\alpha t}^c)\big).$$

Integrating gives

$$\int_0^\infty t^{q-1}\lambda_n(K_t^c)\mathrm{d}t \geq \int_0^\infty t^{q-1}\lambda_n(S_{\alpha t}^c)\mathrm{d}t,$$

or

$$\int_{\mathbb{R}^n} \|x\|^q \mathrm{d}\lambda_n(x) \geq \int_{\mathbb{R}} |x/\alpha|^q \mathrm{d}\lambda_1(x),$$

thus, given the choice of α, we obtain (1.11). $\square$

1.3 The Gamma and Weibull distributions

1.3.1 Preliminaries

The aim of this section is to extend Theorem 1.11 to the measures η_p^n on $\mathbb{R}^n$ with densities

$$\mathrm{d}\eta_p^n(x) = (c_p/2)^n e^{-|x|_p^p}\mathrm{d}x, \qquad x \in \mathbb{R}^n, \tag{1.12}$$

where we denote $|(x_1,\ldots,x_n)|_p = (\sum |x_i|^p)^{1/p}$ and $c_p = 1/\Gamma(1+1/p)$ is the appropriate normalization constant.

We begin with a few definitions. In this section, for a Borel measure μ on $\mathbb{R}$ its product measure $\mu \otimes \ldots \otimes \mu = \mu^{\otimes n}$ is denoted by μ^n. Recall that such a product measure μ^n on $\mathbb{R}^n$ is said to *support the S-inequality for a Borel set* $L \subset \mathbb{R}^n$ if for every $s > 0$ the dilation $K = sL$ and every strip $P = \{x \in \mathbb{R}^n,\ |x_1| \leq p\}$ satisfy

$$\mu^n(K) = \mu^n(P) \quad \Longrightarrow \quad \mu^n(tK) \geq \mu^n(tP), \qquad \text{for } t \geq 1. \tag{1.13}$$

If we assume that the function $\Psi(x) = \mu([-x,x])$ is invertible for $x \geq 0$, we can write (1.13) as

$$\mu^n(tK) \geq \Psi\Big[t\Psi^{-1}\big(\mu(K)\big)\Big], \qquad \text{for } t \geq 1. \tag{1.14}$$

A set $K \subset \mathbb{R}^n$ is called a *down set* if for every point $x \in K$, the set K contains the cube $[-|x_1|,|x_1|] \times \ldots \times [-|x_n|,|x_n|]$.

1.3.2 Results

Now we can state the main result.

1.13 Theorem. *Let $p \in (0,1]$. Then the measure η_p^n defined in (1.12) supports the S-inequality for all down sets in $\mathbb{R}^n$.*

As in Theorem 1.11, thanks to a simple coordinate-wise transport of measure argument, we establish the following corollary.

1.14 Corollary. *For $p \in (0,1]$ and $\alpha > 0$ introduce the measure $\mu_{p,\alpha}$ on $\mathbb{R}$ with density*

$$\mathrm{d}\mu_{p,\alpha}(x) = \alpha c_p |x|^{\alpha-1} e^{-|x|^{\alpha p}} \mathrm{d}x. \tag{1.15}$$

Then the product measures $\mu_{p,\alpha}^n$ supports the S-inequality for all down sets in $\mathbb{R}^n$. In particular, defining for $\alpha > 0$ and $q \geq 1$ on $\mathbb{R}$ the symmetric Weibull measure ω_α with the parameter α and the symmetric Gamma measure λ_q with the parameter q, given by

$$\mathrm{d}\omega_\alpha(x) = \frac{1}{2}\alpha |x|^{\alpha-1} e^{-|x|^{\alpha}} \mathrm{d}x, \tag{1.16}$$

$$\mathrm{d}\lambda_q(x) = \frac{1}{2\Gamma(q)} q |x|^{q-1} e^{-|x|} \mathrm{d}x. \tag{1.17}$$

we obtain that the product measures ω_α^n and λ_q^n support the S-inequality for all down sets in $\mathbb{R}^n$.

Recall that Corollary 1.12 says that the S-inequality for the symmetric exponential measure, (1.9), yields the optimal comparison of moments of unconditional norms. The same argument shows that the same holds true for any product measure.

1.15 Corollary. *Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ which is unconditional, i.e.*

$$\|(\epsilon_1 x_1, \ldots, \epsilon_n x_n)\| = \|(x_1, \ldots, x_n)\|$$

for any $x_j \in \mathbb{R}$ and $\epsilon_j \in \{-1, 1\}$. Suppose that a product Borel probability measure $\mu^n = \mu^{\otimes n}$ supports the S-inequality for all down sets in $\mathbb{R}^n$. Then for $p \geq q > 0$

$$\left(\int_{\mathbb{R}^n} \|x\|^p \mathrm{d}\mu^n(x) \right)^{1/p} \leq C_{p,q} \left(\int_{\mathbb{R}^n} \|x\|^q \mathrm{d}\mu^n(x) \right)^{1/q}, \tag{1.18}$$

where the constant

$$C_{p,q} = \frac{\left(\int_{\mathbb{R}} |x|^p \mathrm{d}\mu(x)\right)^{1/p}}{\left(\int_{\mathbb{R}} |x|^q \mathrm{d}\mu(x)\right)^{1/q}}$$

is the best possible. In particular, we might take $\mu = \eta_p, \omega_\alpha, \lambda_q$*, for* $p \in (0,1]$*,* $\alpha > 0$*,* $q \geq 1$ *(see* (1.12), (1.16), (1.17)*).*

In the next subsection we present the proof of the main result. The proofs of Corollaries 1.14 and 1.15 are essentially identical to those of Theorem 1.11 and Corollary 1.12 respectively. We omit the latter, but because of some subtleties we still discuss the proof of the former.

1.3.3 Proofs

Proof of Theorem 1.13

The theorem is trivial in one dimension. For higher dimensions the strategy of the proof is to reduce the problem to the two dimensional case where everything can be computed. This is done in the following proposition.

1.16 Proposition. *Let* μ *be a Borel probability measure on* $\mathbb{R}$*. Let* $\mu^n = \mu^{\otimes n}$ *be its product measure on* $\mathbb{R}^n$*. If* μ^2 *supports the S-inequality for all down sets on* $\mathbb{R}^2$ *then for any* $n \geq 2$ *the measure* μ^n *supports the S-inequality for all down sets on* $\mathbb{R}^n$*.*

Proof. We proceed by induction on n. Let us fix $n \geq 2$ and assume that μ^n supports the S-inequality for all down sets in $\mathbb{R}^n$. We would like to show that μ^{n+1} supports the S-inequality for all down sets in $\mathbb{R}^{n+1}$. To this end consider a down set $K \subset \mathbb{R}^{n+1}$ and fix $t \geq 1$. Thanks to Fubini's theorem

$$\mu^{n+1}(tK) = \int_{\mathbb{R}} \mu^n((tK)_x) \mathrm{d}\mu(x) = \int_{\mathbb{R}} \mu^n(tK_{x/t}) \mathrm{d}\mu(x),$$

where $A_x = \{y \in \mathbb{R}^n,\ (y,x) \in A\}$ is a section of a set $A \subset \mathbb{R}^{n+1}$ at a level $x \in \mathbb{R}$. For a set A let P_A denote a strip with a width w_A such that $\mu^n(A) = \mu^n(P_A)$. Since the section $K_{x/t}$ is a down set in $\mathbb{R}^n$, by the induction hypothesis we obtain

$$\mu^{n+1}(tK) \geq \int_{\mathbb{R}} \mu^n\left(tP_{K_{x/t}}\right) \mathrm{d}\mu(x) = \int_{\mathbb{R}} \mu\left([-tw_{K_{x/t}}, tw_{K_{x/t}}]\right) \mathrm{d}\mu(x).$$

For simplicity denote the function $x \mapsto w_{K_x}$ by f. If we put $G_f \subset \mathbb{R}^2$ to be a down set *generated* by f, i.e. $G_f = \{(x,y) \in \mathbb{R}^2,\ |y| \le f(x), x \in \mathbb{R}\}$, then its dilation tG_f is generated by the function $x \mapsto tf(x/t)$. Therefore

$$\int_{\mathbb{R}} \mu\left([-tw_{K_{x/t}}, tw_{K_{x/t}}]\right) \mathrm{d}\mu(x) = \mu^2(tG_f).$$

However, $\mu^2(G_f) = \mu^{n+1}(K)$, so taking the strip $P = [-w,w] \times \mathbb{R}^n$ with the same measure as K we see that the strip $[-w,w] \times \mathbb{R}$ has the same measure as G_f. Now the fact that μ^2 supports the S-inequality implies $\mu^2(tG_f) \ge \mu^2(t([-w,w]\times\mathbb{R})) = \mu^{n+1}(tP)$. Thus we have shown that $\mu^{n+1}(tK) \ge \mu^{n+1}(tP)$ and this completes the proof. $\square$

Thus it suffices to show the theorem when $n = 2$. Notice that any down set $K \subset \mathbb{R}^2$ can be described by a nonincreasing function $f : [0,\infty) \to [0,\infty)$, namely

$$K = \left\{(x,y) \in \mathbb{R}^2,\ |y| \le f(|x|)\right\}.$$

Fix such a function and take a strip $P = \{|x_1| \le w\}$ such that $\eta_p^2(K) = \eta_p^2(P)$. To prove that η_p^2 supports the S-inequality for the down set K it is enough to show that (see Proposition 1.5)

$$\frac{\mathrm{d}}{\mathrm{d}t}\eta_p^2(tK)\Big|_{t=1} \ge \frac{\mathrm{d}}{\mathrm{d}t}\eta_p^2(tP)\Big|_{t=1}.$$

Let

$$M_p(K) = \int_K (|x|^p + |y|^p)\, \mathrm{d}\eta_p^2(x,y).$$

We have

$$\eta_p^2(tK) = \frac{c_p^2}{4}\int_{tK} e^{-(|x|^p+|y|^p)}\, \mathrm{d}x\mathrm{d}y = \frac{c_p^2}{4}\int_K t^2 e^{-t^p(|x|^p+|y|^p)}\, \mathrm{d}x\mathrm{d}y,$$

hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\eta_p^2(tK)\Big|_{t=1} = 2\eta_p^2(K) - pM_p(K).$$

Therefore we are to prove that $M_p(K) \le M_p(P)$. Define the functions $T : [0,\infty) \to [0,1]$, $S : [0,\infty) \to [0,1]$

$$T(u) = c_p \int_u^\infty e^{-x^p}\, \mathrm{d}x, \qquad S(u) = c_p \int_0^u x^p e^{-x^p}\, \mathrm{d}x$$

and let μ_+ be the probability measure with density $c_p e^{-x^p}$ on $[0,\infty)$. Note that

$$S(u) = c_p \frac{1}{p}\int_0^u x(-e^{-x^p})'\,\mathrm{d}x = -\frac{c_p}{p}ue^{-u^p} + \frac{1}{p}(1-T(u)).$$

Thus $S(\infty) = 1/p$. We have

$$\begin{aligned} M_p(K) &= c_p^2 \int_0^\infty \int_0^{f(x)} (x^p + y^p)e^{-x^p-y^p}\,\mathrm{d}y\,\mathrm{d}x \\ &= c_p \int_0^\infty x^p e^{-x^p}(1-T(f(x)))\,\mathrm{d}x + c_p\int_0^\infty S(f(x))e^{-x^p}\,\mathrm{d}x \\ &= \frac{1}{p} - \int_0^\infty x^p T(f(x))\,\mathrm{d}\mu_+(x) + \int_0^\infty S(f(x))\,\mathrm{d}\mu_+(x). \end{aligned}$$

To calculate $M_p(P)$, it is enough to take $f(x) = \infty$ for $x < w$ and $f(x) = 0$ for $x \geq w$ in the above computations, so we obtain

$$\begin{aligned} \int_P (|x|^p + |y|^p)\,\mathrm{d}\eta_p^2(x,y) &= \frac{1}{p} - \left(\frac{1}{p} - S(w)\right) + \frac{1}{p}(1-T(w)) \\ &= \frac{1}{p} + S(w) - \frac{1}{p}T(w). \end{aligned}$$

Let $\Phi : [0,1] \to \mathbb{R}$, $\Phi = S\circ T^{-1}$ and $g : [0,\infty) \to [0,1]$, $g = T\circ f$. We would like to prove

$$\int \Phi(g)\,\mathrm{d}\mu_+ - \int_0^\infty x^p g(x)\,\mathrm{d}\mu_+(x) \leq S(w) - \frac{1}{p}T(w).$$

Observe that

$$\begin{aligned} \eta_p^2(K) &= c_p^2 \int_0^\infty \int_0^{f(x)} e^{-y^p-x^p}\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_0^\infty (1-T(f(x)))\,\mathrm{d}\mu_+(x) = 1 - \int g\,\mathrm{d}\mu_+. \end{aligned}$$

Our assumption $\eta_p^2(K) = \eta_p^2(P)$ yields $\int g\,\mathrm{d}\mu_+ = T(w)$. Moreover,

$$S(w) = \Phi(T(w)) = \Phi\left(\int g\,\mathrm{d}\mu_+\right).$$

Therefore our inequality can be expressed in the following form

$$\int \Phi(g)\,\mathrm{d}\mu_+ - \Phi\left(\int g\,\mathrm{d}\mu_+\right) \leq \int_0^\infty g(x)\left(x^p - \frac{1}{p}\right)\mathrm{d}\mu_+(x).$$

Note that $g : [0,\infty) \to [0,1]$ is nondecreasing. Summing up, to establish Theorem 1.13 it suffices to prove the following lemma.

1.17 Lemma. *Let $p \in (0,1]$ and let μ_+ be a measure with density $c_p e^{-x^p}$ supported on $[0,\infty)$. Then for all nondecreasing functions $g : [0,\infty) \to [0,1]$ we have*

$$\int \Phi(g)\, d\mu_+ - \Phi\left(\int g\, d\mu_+\right) \le \int_0^\infty g(x)\left(x^p - \frac{1}{p}\right) d\mu_+(x). \tag{1.19}$$

In order to prove Lemma 1.17 we shall need a lemma of R. Latała and K. Oleszkiewicz (see [LO00, Lemma 4] or [Wol07, Theorem 1]). For convenience let us recall their result.

1.18 Lemma (Latała–Oleszkiewicz)**.** *Let (Ω, ν) be a probability space and suppose that $\Phi : [0,1] \to \mathbb{R}$ has strictly positive second derivative and $1/\Phi''$ is concave. For a nonnegative function $g : \Omega \to [0,1]$ define a functional*

$$\Psi_\Phi(g) = \int_\Omega \Phi(g)\, d\nu - \Phi\left(\int_\Omega g\, d\nu\right). \tag{1.20}$$

Then Ψ_Φ is convex, namely

$$\Psi_\Phi(\lambda f + (1-\lambda)g) \le \lambda \Psi_\Phi(f) + (1-\lambda)\Psi_\Phi(g).$$

Now we show that our function $\Phi = S \circ T^{-1}$ satisfies the assumptions of Lemma 1.18.

1.19 Lemma. *The function $\Phi = S \circ T^{-1} : [0,1] \to \mathbb{R}$ satisfies $\Phi'' > 0$ and $(1/\Phi'')'' \le 0$.*

Proof. Let $T^{-1} = F$. Note that $F' = \frac{1}{T'(F)} = -\frac{1}{c_p} e^{F^p}$. We have

$$\Phi' = S'(F)F' = c_p F^p e^{-F^p}\left(-\frac{1}{c_p} e^{F^p}\right) = -F^p$$

and

$$\Phi'' = -pF^{p-1}F' = \frac{p}{c_p} F^{p-1} e^{F^p} > 0.$$

Moreover,

$$\begin{aligned}(1/\Phi'')' &= \frac{c_p}{p}\left(F^{1-p} e^{-F^p}\right)' \\ &= \frac{c_p}{p}\left((1-p)F^{-p} - pF^{1-p}F^{p-1}\right) e^{-F^p} F' = 1 - \frac{1-p}{p} F^{-p}\end{aligned}$$

and

$$(1/\Phi'')'' = (1-p)F^{-p-1}F' = -\frac{1-p}{c_p}F^{-p-1}e^{F^p} \leq 0.$$

□

1.20 Remark. The reader might want to notice that the last inequality is the place where the proof of the theorem does not work for other values of p.

We are ready to give the proof of Lemma 1.17.

Proof of Lemma 1.17. Combining Lemmas 1.18 and 1.19 we see that the left hand side of (1.19) is a convex functional of g. The right hand side is linear in g and therefore we see that $\lambda g_1 + (1-\lambda)g_2$ satisfies (1.19) for every $\lambda \in [0,1]$ whenever g_1, g_2 satisfy (1.19). By an approximation argument it suffices to prove our inequality for nondecreasing right-continuous piecewise constant functions that take finitely many values. Every such a function is a convex combination of a finite collection of functions of the form $g_a(x) = \mathbf{1}_{[a,\infty)}(x)$, where $a \in [0,\infty]$. Therefore it suffices to check (1.19) for the functions g_a. Since $\Phi(0) = S(\infty) = 1/p$ and $\Phi(1) = 0$, we have

$$\int \Phi(g_a)\,\mathrm{d}\mu_+ - \Phi\left(\int g_a\,\mathrm{d}\mu_+\right) = \frac{1}{p}\left(1 - T(a)\right) - S(a)$$

and

$$\int_0^\infty g_a(x)\left(x^p - \frac{1}{p}\right)\mathrm{d}\mu_+(x) = \frac{1}{p} - S(a) - \frac{1}{p}T(a),$$

thus we have equality in (1.19). □

The proof of Theorem 1.13 is now complete.

Proof of Corollary 1.14

Recall that the idea is that once a measure supports the S-inequality for all down sets then so does its image under a properly chosen transformation (cf. the proof of Theorem 1.11). Fix $p \in (0,1]$ and $\alpha > 0$. Consider the mapping $F\colon [0,\infty)^n \longrightarrow [0,\infty)^n$ given by the formula

$$F(x_1,\ldots,x_n) = (x_1^\alpha,\ldots,x_n^\alpha).$$

We will use it to change the variables. So, take a down set $K \subset \mathbb{R}^n$, the strip $P \subset \mathbb{R}^n$ such that $\eta_p^n(K) = \eta_p^n(P)$, and compute the measure of the dilation tK for some $t \geq 1$

$$\eta_p^n(tK) = \left(\frac{c_p}{2}\right)^n \int_{tK} e^{-|x|_p^p} \mathrm{d}x = c_p^n \int_{tK\cap[0,\infty)^n} e^{-\sum x_i^p} \mathrm{d}x$$
$$= (\alpha c_p)^n \int_{F^{-1}(tK\cap[0,\infty)^n)} \prod y_i^{\alpha-1} e^{-y_i^{\alpha p}} \mathrm{d}y.$$

In the first equality we have used the symmetries of down sets, while in the last one we have changed the variables putting $x = F(y)$. Introducing the measure $\mu_{p,\alpha}$ on $\mathbb{R}$ with density (1.15) we thus have seen that

$$\eta_p^n(tK) = \mu_{p,\alpha}(\widetilde{tK}),$$

where for a down set A in $\mathbb{R}^n$ the set $\widetilde{A}$ denotes a down set such that $\widetilde{A} \cap [0,\infty)^n = F^{-1}(A \cap [0,\infty)^n)$ (note that it makes sense as F is monotone with respect to each coordinate). The point is that due to the homogeneity of F we have $\widetilde{tK} = t^{1/\alpha}\widetilde{K}$. Moreover, strips are mapped onto strips. Therefore

$$\mu_{p,\alpha}(t^{1/\alpha}\widetilde{K}) = \eta_p^n(tK) \geq \eta_p^n(tP) = \mu_{p,\alpha}(t^{1/\alpha}\widetilde{P}),$$

which means that $\mu_{p,\alpha}$ supports the S-inequality for the down set $\widetilde{K}$. Since the down set K is arbitrary, we conclude that $\mu_{p,\alpha}$ supports the S-inequality for all down sets. To finish the proof notice that we recover the Weibull distribution putting $p = 1$, namely $\omega_\alpha = \mu_{1,\alpha}$. To obtain the Gamma distribution set $\alpha = 1/p = q$, as then $\lambda_q = \mu_{1/q,q}$.

1.21 Remark. We might use more general change of variables $y_i = V(x_i)$ for some increasing function $V\colon [0,\infty) \longrightarrow [0,\infty)$, $V(0) = 0$ and ask whether we will derive the S-inequality for other measures than $\mu_{p,\alpha}$ exploiting the above technique. Since we would like to have $\widetilde{tK} = u(t)\widetilde{K}$ for a monotone function u, we check it would imply that $V(st) = CV(s)V(t)$, and C is a constant. So V should be a power function but this case has been studied in the above proof.

1.4 Notes and comments

The results of Section 1.2 come from the publication [NT13]. The change of variables used in the proof of Theorem 1.11, which establishes the S-inequality for the symmetric exponential measure, was pointed out by B. Maurey after a seminar talk by the author. R. Adamczak's remark regarding Lemma 1.7, the one-dimensional entropy inequality, led to the general formulation presented here.

Section 1.3 is based on the authors' further work on the S-inequality, the article [NT14b].

P. Nayar and T. Tkocz are both including the results of the two aforementioned papers into their PhD theses. They worked together and contributed equally to the results obtained.

Chapter 2

Tensor products of random unitary matrices

2.1 Introduction

In quantum mechanics, the time evolution of two noninteracting subsystems can be described by an operator $e^{itH} \otimes e^{itH'}$, where H and H' are Hamiltonians of the subsystems (see e.g. chapters 2.2 and 3.1 in [BP02]). In applications, the unitary operator e^{itH}, which is *a priori* complicated, is replaced by a random unitary matrix, to make a model tractable. This powerful idea goes back to E. Wigner (see e.g. his seminal paper [Wig55]). Here by an $n \times n$ random unitary matrix we mean a matrix chosen according to the Haar measure on the unitary group $U(n)$. From this point of view it seems natural to study asymptotic local properties of spectra of the tensor product $U_m \otimes V_n$ of two independent $m \times m$ and $n \times n$ random unitary matrices.

More generally, consider a quantum system consisting of M noninteracting subsystems. For simplicity we can assume that each of them is represented in an n dimensional Hilbert space, so that any local unitary dynamics can be written as $U_1 \otimes \ldots \otimes U_M$, where the U_j are $n \times n$ unitary matrices. If the unitary dynamics of each subsystem is generic, the matrices U_j can be represented by independent random unitary matrices of size n.

2.2 Background and notation

Since we shall view a collection of eigenvalues of a random matrix as a *point process*, for clarity we need to start off by recalling some definitions and known facts. A user-friendly and brief introduction to the theory of point processes can be found in [HKPV09]. The monograph [AGZ10] is a good reference for some background knowledge on random matrix theory.

A *point process* Γ on $\mathbb{R}$ is a random integer-valued positive and σ-finite Borel measure on $\mathbb{R}$. In other words, $\Gamma\colon \Omega \longrightarrow \mathcal{M}(\mathbb{R})$ is a random variable taking values in the subset of integer-valued positive measures of the set $\mathcal{M}(\mathbb{R})$ of all σ-finite Borel measures on $\mathbb{R}$. If for every $x \in \mathbb{R}$, $\Gamma(\{x\}) \leq 1$ a.s, then Γ is called *simple*. It is not hard to see that simple point processes on $\mathbb{R}$ correspond to random discrete subsets of $\mathbb{R}$. The latter point of view is particularly useful for us. Indeed, the eigenvalues $\lambda_1, \ldots, \lambda_n$ of, say an $n \times n$ Hermitian random matrix constitute a.s. a discrete random subset $\{\lambda_1, \ldots, \lambda_n\}$ of $\mathbb{R}$, which defines a random counting measure

$$\Gamma(D) = \sum_{i=1}^{n} \delta_{\{\lambda_i\}}(D).$$

The set $\mathcal{M}(\mathbb{R})$ is a metric space and the σ-algebra of Borel subsets is generated by the cylinders which are the subsets of the form $C^{(I_i)}_{(D_i), i\leq k} = \{\mu \in \mathcal{M}(\mathbb{R});\ \forall i \leq k\ \mu(D_i) \in I_i\}$, given Borel subsets $D_1, \ldots, D_k \subset \mathbb{R}$ and intervals $I_1, \ldots, I_k$. Thus, to determine the distribution of a point process, it is enough to specify the probabilities $\mathbb{P}\left(\Gamma \in C^{(I_i)}_{(D_i), i\leq k}\right)$. For instance, a Poisson point process Γ on $\mathbb{R}$ with intensity $\lambda > 0$ is characterized by setting

$$\mathbb{P}\left(\Gamma(D_i) = k_i, i \leq n\right) = \prod_{i=1}^{n} e^{-\lambda|D_i|} \frac{(\lambda|D_i|)^{k_i}}{k_i!},$$

for every $n \geq 1$, integers $k_1, \ldots, k_n \geq 0$, and mutually disjoint bounded Borel subsets D_i of $\mathbb{R}$, that is $\Gamma(D_i)$ are independent Poisson random variables with parameters $|D_i|$ (Lebesgue measure of D_i). The existence of such process is guaranteed by the Kolmogorov consistency theorem. When $\lambda = 1$ we call it the standard Poisson point process and denote it by Π.

Sometimes in practice it is more convenient to describe distributions of point processes by so-called *intensity functions*, or in the language of physicists, *correlation functions*. Given a simple point process Γ, its k-th correlation function (if exists) $\rho_\Gamma^{(k)} \colon \mathbb{R}^k \longrightarrow [0, \infty)$ is defined by the fact that for every mutually disjoint Borel subsets $D_1, \ldots, D_k$ of $\mathbb{R}$ there holds

$$\mathbb{E} \prod_{i=1}^{k} \Gamma(D_i) = \int_{D_1 \times \ldots \times D_k} \rho_\Gamma^{(k)}(x_1, \ldots, x_k) \mathrm{d}x_1 \ldots \mathrm{d}x_k.$$

For example, for the standard Poisson process we have

$$\rho_\Pi^{(k)} \equiv 1$$

because $\mathbb{E} \prod_{i=1}^{k} \Gamma(D_i)$ is the product of Lebesgue measures of the D_i. Under some mild technical assumptions on Γ, the correlation functions are akin to densities

$$\rho_\Gamma^{(k)}(x_1, \ldots, x_k) = \lim_{\epsilon \to 0} \frac{\mathbb{P}\left(\Gamma \text{ has a point in } (x_i - \epsilon, x_i + \epsilon) \text{ for each } i \leq k\ \right)}{(2\epsilon)^k},$$

for every sequence of pairwise distinct points $x_1, \ldots, x_k$.

An important class of point processes are *determinantal point processes*: processes for which the correlation functions can be put in the form

$$\det[K(x_i, x_j)]_{i,j=1}^{k}$$

for some *kernel* function $K \colon \mathbb{R}^2 \longrightarrow \mathbb{R}$. For instance, the *sine process* Σ on $\mathbb{R}$ is defined by setting its correlation functions $\rho_\Sigma^{(k)}$ to be

$$\rho_\Gamma^{(k)}(x_1, \ldots, x_k) = \det[Q(x_i, x_j)]_{i,j=1}^{k},$$

with the *sine kernel* $Q(x, y) = q(x - y)$, where q is given by

$$q(u) = \frac{\sin(\pi u)}{\pi u}.$$

Obviously, this is a determinantal point process. Let us see why it is important.

Given an $n \times n$ random unitary matrix with eigenvalues $e^{i\xi_1}, \ldots, e^{i\xi_n}$, where $\xi_i \in [0, 2\pi)$ are *eigenphases*, we define the point process $\Xi_n = \{\xi_1, \ldots, \xi_n\}$ (here and throughout we identify a simple point process with a random discrete subset of its atoms). It is well known that this process is determinantal with the kernel $S_n(x,y) = s_n(x-y)$, where

$$s_n(u) = \frac{1}{2\pi}\frac{\sin\left(\frac{nu}{2}\right)}{\sin\left(\frac{u}{2}\right)}. \tag{2.1}$$

Since $\frac{2\pi}{n}s_n\left(\frac{2\pi}{n}u\right) \xrightarrow[n\to\infty]{} q(u)$, when n becomes large, the process $\frac{n}{2\pi}(\Xi_n - \pi)$ of the rescaled eigenphases of the $n\times n$ random unitary matrix locally behaves as the sine process Σ.

By the *superposition* of two simple point processes $\Psi = \{\psi_1,\ldots,\psi_M\}$, $\Phi = \{\phi_1,\ldots,\phi_N\}$, $M,N \le \infty$, we mean the union

$$\Psi \cup \Phi = \{\psi_1,\ldots,\psi_M,\phi_1,\ldots,\phi_N\}.$$

2.3 Results

2.3.1 Two matrices of large sizes

Given two independent $m\times m$ and $n\times n$ random unitary matrices U and U' we get two independent point processes of their eigenphases $\Xi_m = \{\xi_1,\ldots,\xi_m\}$ and $\Xi'_n = \{\xi'_1,\ldots,\xi'_n\}$ respectively. We define the point process $\Xi_m \otimes \Xi'_n$ of the eigenphases of the matrix $U\otimes U'$ as

$$\Xi_m \otimes \Xi'_n = \{\xi_i + \xi'_j \bmod 2\pi,\ i=1,\ldots,m, j=1,\ldots,n\}.$$

Our first main result reads as follows.

2.1 Theorem. *Let Ξ_m and Ξ'_n be point processes of eigenphases of two independent $m\times m$ and $n\times n$ random unitary matrices. Let $\Sigma_1,\ldots,\Sigma_m$ be independent sine processes and let Π be a Poisson process on $\mathbb{R}$. Then for each $k\le n$ the k-th correlation function of the process $\Xi_m\otimes\Xi'_n$ exists and*

(a) $\rho^{(k)}_{\frac{mn}{2\pi}(\Xi_m\otimes\Xi'_n-\pi)} \xrightarrow[n\to\infty]{} \rho^{(k)}_{m\Sigma_1\cup\ldots\cup m\Sigma_m}$,

(b) $\rho^{(k)}_{\frac{mn}{2\pi}(\Xi_m\otimes\Xi'_n-\pi)} \xrightarrow[m,n\to\infty]{} \rho^{(k)}_{\Pi}$,

uniformly on all compact sets in $\mathbb{R}^k$.

2.2 Remark (Weak convergence). We say that a sequence of point processes (τ_n) converges *in distribution* to a point process τ if the law ν_n of τ_n converges weakly to that of τ, say ν, in the space $\mathcal{M}_1(\mathcal{M}(\mathbb{R}))$ of probability measures on $\mathcal{M}(R)$, i.e. $\int f \mathrm{d}\nu_n \to \int f \mathrm{d}\nu$ for any bounded continuous function on $\mathcal{M}(\mathbb{R})$. Clearly, these integrals can be expressed using correlation functions, so the theorem implies the convergence in distribution of the point processes in question.

2.3 Remark (Heuristic behind (a)). In the simplest case $m = 2$ we have

$$\begin{aligned}\Xi_2 \otimes \Xi'_n =&\{\xi_1 + \xi'_1 \mod 2\pi, \dots, \xi_1 + \xi'_n \mod 2\pi\}\\ &\cup \{\xi_2 + \xi'_1 \mod 2\pi, \dots, \xi_2 + \xi'_n \mod 2\pi\}.\end{aligned}$$

After shifting and rescaling we end up with two families of rescaled eigenphases of an $n \times n$ random unitary matrix which differ roughly by a large shift $\frac{n}{2\pi}(\xi_1 - \xi_2)$ which is independent of the matrix. That makes the families independent and in the limit, according to $\rho^{(k)}_{\frac{n}{2\pi}(\Xi_n - \pi)} \xrightarrow[n\to\infty]{} \rho^{(k)}_{\Sigma}$, they look like sine processes.

2.4 Remark (Superposition of many sine processes becomes a Poisson point process). Notice that for m independent copies $\Phi_1, \dots, \Phi_m$ of a point process Φ we have

$$\rho^{(k)}_{\Phi_1 \cup \ldots \cup \Phi_m}(x_1, \dots, x_k) = \sum_{p=1}^{m \wedge k} \sum_{\pi \in \mathfrak{S}(k,p)} \frac{m!}{(m-p)!} \prod_{j=1}^{p} \rho^{(\sharp \pi_j)}_{\Phi}((x_i)_{i \in \pi_j}),$$

where $\mathfrak{S}(k,p)$ is the collection of all partitions of the set $\{1, \dots, k\}$ into p nonempty pairwise disjoint subsets. By this we mean that if π is such a partition then $\pi = \{\pi_1, \dots, \pi_p\}$, where $\pi_q = \{\pi(q,1), \dots, \pi(q, \sharp\pi_q)\}$ is the q-th block of the partition π.

Along with the fact that if we rescale, $\rho^{(k)}_{\lambda\Phi}(x)$ becomes $\frac{1}{\lambda^k}\rho^{(k)}_{\Phi}\left(\frac{1}{\lambda}x\right)$, the previous observation yields

$$\rho^{(k)}_{m\Sigma_1 \cup \ldots \cup m\Sigma_m}(x) = \sum_{p=1}^{m \wedge k} \sum_{\pi \in \mathfrak{S}(k,p)} \frac{1}{m^k} \frac{m!}{(m-p)!} \prod_{j=1}^{p} \rho^{(\sharp \pi_j)}_{\Sigma}\left(\frac{1}{m}(x_i)_{i \in \pi_j}\right). \tag{2.2}$$

When m goes to infinity we thus get

$$\lim_{m\to\infty} \rho^{(k)}_{m\Sigma_1 \cup \ldots \cup m\Sigma_m}(x) = \lim_{m\to\infty} \prod_{j=1}^{k} \rho^{(1)}_{\Sigma}\left(\frac{1}{m}(x_i)_{i \in \pi_j}\right) = 1 = \rho^{(k)}_{\Pi}.$$

This retrieves a special case of a high dimensional phenomenon presented in [CD10]. Namely, the authors say "[...] a Poisson process can be viewed as an infinite superposition of determinantal or permanental point processes" (see Theorem 4 therein and the two preceding paragraphs). In view of Theorem (a) this implies that

$$\lim_{m\to\infty}\lim_{n\to\infty}\rho^{(k)}_{\frac{mn}{2\pi}(\Xi_m\otimes\Xi'_n-\pi)}=1.$$

Note that in the second part of the theorem we establish a stronger statement: that letting the dimensions of two independent random unitary matrices go to infinity eliminates all the correlations in their tensor product.

2.3.2 The tensor product of a large number of 2×2 matrices

We next consider M independent 2×2 random unitary matrices $U_1,\ldots,U_M$ and study the asymptotic properties of the phase-spectrum of the matrix $U_1\otimes\ldots\otimes U_M$. Our main result is as follows.

2.5 Theorem. *Let θ_j^1,θ_j^2, $j=1,\ldots,M$ be the eigenphases of independent 2×2 random unitary matrices $U_1,\ldots,U_M$. Define the point process τ_M of the rescaled eigenphases of the matrix $U_1\otimes\ldots\otimes U_M$ as*

$$\tau_M(D)=\sum_{\epsilon=(\epsilon_1,\ldots,\epsilon_M)\in\{1,2\}^M}\boldsymbol{1}_{\left\{\frac{2^M}{2\pi}\left(\theta_1^{\epsilon_1}+\ldots+\theta_M^{\epsilon_M}\mod 2\pi\right)\in D\right\}},\tag{2.3}$$

for any compact set $D\subset[0,\infty)$. Then, for each k there exists a continuous function $\delta_k:[0,\infty)\to[0,\infty)$ with $\delta_k(0)=0$ so that for any mutually disjoint intervals $I_1,\ldots,I_k\subset[0,\infty)$

$$\limsup_{M\to\infty}\frac{\mathbb{P}\left(\tau_M(I_1)>0,\ldots,\tau_M(I_k)>0\right)}{|I_1|\cdot\ldots\cdot|I_k|}\le\left(1+\delta_k(\max_j|I_j|)\right),$$
$$\liminf_{M\to\infty}\frac{\mathbb{P}\left(\tau_M(I_1)>0,\ldots,\tau_M(I_k)>0\right)}{|I_1|\cdot\ldots\cdot|I_k|}\ge\left(1-\delta_k(\max_j|I_j|)\right).$$

Note that the statement of Theorem 2.5 is weaker than that of Theorem 2.1. This is due to the fact that stronger correlations exist in the point process τ_M, which prevent us from demonstrating the convergence of its intensities to those of a Poisson process. The mode of convergence is however strong enough to deduce interesting information, including the weak convergence of the processes (cf. the remark following Theorem 2.1).

2.4 Proofs

2.4.1 Proof of Theorem 2.1

For the sake of convenience, let us recall a few basic facts which will be used frequently in the proof.

Note the following easy estimate (for the definition see (2.1))

$$\sup_{x\in\mathbb{R}}\left|\frac{2\pi}{n}s_n(x)\right| = 1. \tag{2.4}$$

Combined with Hadamard's inequality (see e.g. (3.4.6) in [AGZ10]), this allows us to bound the correlation functions,

$$\sup_{x\in\mathbb{R}^k}\rho^{(k)}_{\Xi_n}(x) \le k^{k/2}\|s_n\|^k_\infty = \frac{k^{k/2}}{(2\pi)^k}n^k. \tag{2.5}$$

Proof of Theorem 2.1 (a). Let $\Theta_{m,n} = \frac{mn}{2\pi}(\Xi_m\otimes\Xi'_n - \pi)$. Fix a natural number k. Since we will let n go to infinity, we may assume that $k \le n$. First we show that there exists functions $\rho^{(k)}_{\Theta_{m,n}}\colon \mathbb{R}^k \longrightarrow [0,\infty)$ so that for any bounded, measurable function $f\colon \mathbb{R}^k \longrightarrow \mathbb{R}$ we have

$$\mathbb{E}\sum f(\theta_1,\ldots,\theta_k) = \int_{\mathbb{R}^k} f(x)\rho^{(k)}_{\Theta_{m,n}}(x)\mathrm{d}x,$$

where the summation is over all ordered k-tuples $(\theta_1,\ldots,\theta_k)$ of distinct points of $\Theta_{m,n}$. This will prove that $\rho^{(k)}_{\Theta_{m,n}}$ are the correlation functions of $\Theta_{m,n}$. Then we will deal with the limit when $n\to\infty$.

Fix f. Since for each $s = 1,\ldots,k$, $\theta_s = \frac{mn}{2\pi}(\xi_{i_s} + \xi'_{j_s} \mod 2\pi - \pi)$ for some $i_s \in \{1,\ldots,m\}$, $j_s \in \{1,\ldots,n\}$ we can write

$$\mathbb{E}\sum f(\theta_1,\ldots,\theta_k) = \mathbb{E}\sum_{\substack{i\in\{1,\ldots,m\}^k\\ j\in\{1,\ldots,n\}^k}} f\left(\left(\frac{mn}{2\pi}(\xi_{i_s} + \xi'_{j_s} \mod 2\pi - \pi)\right)_{s=1}^k\right),$$

where the second sum is over k-tuples i, j such that the pairs $(i_1, j_1), \dots,$ (i_k, j_k) are pairwise distinct. This certainly happens when all the j_s are distinct. Call these choices of i and j *good* and the rest *bad*. So

$$\mathbb{E}\sum_{i,j} f = \mathbb{E}\sum_{\text{good } i,j} f + \mathbb{E}\sum_{\text{bad } i,j} f.$$

First we handle the *good* sum. Some of the i_s may overlap and we will control them using partitions of the set $\{1, \dots, k\}$ into $p \le k \wedge m$ nonempty pairwise disjoint subsets (see Remark 2.4 for the notation) so that $i_s = i_t$ whenever s and t belong to the same block of a partition. We have

$$\mathbb{E}\sum_{\text{good } i,j} f = \sum_{p=1}^{k\wedge m}\sum_{\pi\in\mathfrak{S}(k,p)}\mathbb{E}\sum_{\substack{\text{distinct}\\ i_{\pi(1,1)},\dots,i_{\pi(p,1)}}}\ \sum_{\substack{\text{distinct}\\ j_1,\dots,j_k}} f.$$

The sums over i and j have been separated. Therefore taking advantage of independence as well as recalling the definitions of the p-th and k-th correlation functions of Ξ_m and Ξ'_n we find

$$\begin{aligned}\mathbb{E}\sum_{\text{good } i,j} f = \sum_{p,\pi}\int_{[0,2\pi]^p}\int_{[0,2\pi]^k} f&\left(\left(\frac{mn}{2\pi}(x_{\pi(s)} + y_s \mod 2\pi - \pi)\right)_{s=1}^k\right)\\ &\cdot\rho^{(p)}_{\Xi_m}(x_1,\dots,x_p)\rho^{(k)}_{\Xi'_n}(y_1,\dots,y_k)\\ &\mathrm{d}x_1\dots\mathrm{d}x_p\mathrm{d}y_1\dots\mathrm{d}y_k,\end{aligned}$$

where we note $\pi(s) = q \iff s \in \pi_q$. Finally, we need to address the technicality concerning the addition $\mod 2\pi$. Keeping in mind that we integrate over $[0,2\pi]^p$ and $[0,2\pi]^k$ we consider for $\eta \in \{0,1\}^k$ the set

$$\begin{aligned}U_\eta = \Big\{x \in [0,2\pi]^p, y\in[0,2\pi]^k;\ \forall s\le k\ x_{\pi(s)} + y_s < 2\pi \text{ if } \eta_s = 0, \text{ and}\\ x_{\pi(s)} + y_s \ge 2\pi \text{ if } \eta_s = 1\Big\}.\end{aligned}$$

Then on U_η we have $x_{\pi(s)} + y_s \mod 2\pi = x_{\pi(s)} + y_s - 2\pi\eta_s$, thus changing the variables on U_η so that $z_s = \frac{mn}{2\pi}(x_{\pi(s)} + y_s - 2\pi\eta_s - \pi)$ we get

$$\begin{aligned}\mathbb{E}\sum_{\text{good } i,j} f = \int_{\mathbb{R}^k} f(z)\Bigg(\sum_{p,\pi,\eta}\mathbf{1}_{W_\eta}(z)\int_{[0,2\pi]^p}\mathbf{1}_{V_\eta}(x)\rho^{(p)}_{\Xi_m}(x)\left(\frac{2\pi}{mn}\right)^k\\ \cdot\rho^{(k)}_{\Xi'_n}(y(z,x))\mathrm{d}x\Bigg)\mathrm{d}z,\end{aligned}$$

where $y_s(z,x) = \frac{2\pi}{mn} z_s - x_{\pi(s)} + 2\pi\eta_s + \pi$,

$$V_\eta = \left\{ x \in \mathbb{R}^p;\ \forall s \le k\ \frac{2\pi}{mn} z_s + 2\pi\eta_s - \pi \le x_{\pi(s)} \le \frac{2\pi}{mn} z_s + 2\pi\eta_s + \pi \right\},$$

and

$$W_\eta = \left\{ z \in \mathbb{R}^k;\ \forall s \le k\ z_s \le mn/2 \text{ if } \eta_s = 0, \text{ and } z_s \ge -mn/2 \text{ if } \eta_s = 1 \right\}.$$

Summarizing, we have just seen that the correlation function $\rho^{(k)}_{\Theta_{m,n}}(z)$ takes the form

$$\begin{aligned}\rho^{(k)}_{\Theta_{m,n}}(z) = \sum_{p,\pi,\eta} \mathbf{1}_{W_\eta}(z) \int_{[0,2\pi]^p} \mathbf{1}_{V_\eta}(x) \rho^{(p)}_{\Xi_m}(x) \left(\frac{2\pi}{mn}\right)^k \rho^{(k)}_{\Xi'_n}(y(z,x)) \mathrm{d}x \\ + B_{m,n}(z),\end{aligned} \tag{2.6}$$

where the term $B_{m,n}$ corresponds to the sum over bad indices $\mathbb{E} \sum_{\text{bad } i,j} f$. By the same kind of reasoning we can show that

$$B_{m,n}(z) = \sum_{p=1}^{k} \sum_{q=1}^{k-1} \sum_{\substack{\pi \in \mathfrak{S}(k,p) \\ \tau \in \mathfrak{S}(k,q)}} \sum_{\eta} \mathbf{1}_{\tilde{W}_\eta}(z) \left(\frac{2\pi}{mn}\right)^k \int_{[0,2\pi]^{p+q-k}} \mathbf{1}_{\tilde{V}_\eta}(x) \rho^{(p)}_{\Xi_m}(\tilde{x}(z,x)) \rho^{(q)}_{\Xi'_n}(\tilde{y}(z,x)) \mathrm{d}x,$$

where the sums are over appropriate partitions and $\tilde{W}_\eta$, $\tilde{V}_\eta$ are suitable sets which appear after changing the variables. Now, by (2.5),

$$\|\rho^{(p)}_{\Xi_m} \cdot \rho^{(q)}_{\Xi'_n}\|_\infty \le \frac{p^{p/2} q^{q/2}}{(2\pi)^{p+q}} m^p n^q, \tag{2.7}$$

so

$$B_{m,n}(z) \le C_k \frac{1}{n},$$

where the constant C_k depends only on k (roughly, it equals the number of summands times k^k). Hence, when taking $n \to \infty$ we will not have to worry about $B_{m,n}$.

Let us look at (2.6) and compute the limit of the first term when $n \to \infty$. We observe that $\mathbf{1}_{W_\eta} \to 1$ pointwise on $\mathbb{R}^k$. Moreover, $\sum_\eta \mathbf{1}_{V_\eta} \to \mathbf{1}_{[0,2\pi)^p}$, and $\mathbf{1}_{V_\eta} \to 0$ for η such that $\eta_s \ne \eta_t$ but $\pi(s) = \pi(t)$ for some $s \ne t$. Thus we consider only η such that $\eta_s = \eta_t$ whenever $\pi(s) = \pi(t)$ and then the following simple observation

$$\frac{2\pi}{mn} s_n \left(\frac{2\pi}{mn} u + v \right) \xrightarrow[n\to\infty]{} \begin{cases} 0, & v \neq 0 \\ \frac{1}{m} q\left(\frac{u}{m}\right), & v = 0 \end{cases} \tag{2.8}$$

yields that for all these η,

$$\left(\frac{2\pi}{mn}\right)^k \rho^{(k)}_{\Xi'_n}(y) = \det\left[\frac{2\pi}{mn} s_n \left(\frac{2\pi}{mn}(z_s - z_t) + 2\pi(\eta_s - \eta_t) + x_{\pi(t)} - x_{\pi(s)}\right)\right]_{s,t=1}^{k}$$
$$\xrightarrow[n\to\infty]{} \prod_{j=1}^{p} \det\left[\frac{1}{m} q\left(\frac{z_s - z_t}{m}\right)\right]_{s,t\in\pi_j}$$
$$= \frac{1}{m^k} \prod_{j=1}^{p} \rho^{(\sharp\pi_j)}_{\Sigma}\left(\frac{1}{m}(z_i)_{i\in\pi_j}\right).$$

By estimate (2.5), $\left(\frac{2\pi}{mn}\right)^k \rho^{(k)}_{\Xi'_n}(y)$ is bounded by $k^{k/2}/m^k$, so the integrand in (2.6) can be simply bounded. Thus by Lebesgue's dominated convergence theorem

$$\rho^{(k)}_{\Theta_{m,n}}(z) \xrightarrow[n\to\infty]{} \sum_{p,\pi} \frac{1}{m^k} \prod_{j=1}^{p} \rho^{(\sharp\pi_j)}_{\Sigma}\left(\frac{1}{m}(z_i)_{i\in\pi_j}\right) \cdot \int_{[0,2\pi]^p} \rho^{(p)}_{\Xi_m}(x)\mathrm{d}x.$$

For any $p \le m$ the integral $\int_{[0,2\pi)^p} \rho^{(p)}_{\Xi_m}(x)\mathrm{d}x$ just equals $m!/(m-p)!$. Consequently, we finally obtain

$$\rho^{(k)}_{\Theta_{m,n}}(z_1, \ldots, z_k) \xrightarrow[n\to\infty]{} \sum_{p,\pi} \frac{1}{m^k} \frac{m!}{(m-p)!} \prod_{j=1}^{p} \rho^{(\sharp\pi_j)}_{\Sigma}\left(\frac{1}{m}(z_i)_{i\in\pi_j}\right).$$

In view of (2.2) this completes the proof. $\square$

Proof of Theorem 2.1 (b). Fix a point $z = (z_1, \ldots, z_k) \in \mathbb{R}^k$. We let m and n tend to infinity and want to prove that $\rho^{(k)}_{\Theta_{mn}}(z)$ tends to 1. Recall (2.6) and notice that due to estimate (2.7) all the terms with $p \le k-1$ are bounded above by C_k/m, so we can write

$$\rho^{(k)}_{\Theta_{m,n}}(z) = O\left(\frac{1}{m} + \frac{1}{n}\right) + \sum_{\eta} \mathbf{1}_{W_\eta}(z) \int_{[0,2\pi]^k} \mathbf{1}_{V_\eta}(x) \left(\frac{2\pi}{mn}\right)^k \rho^{(k)}_{\Xi_m}(x)$$
$$\rho^{(k)}_{\Xi'_n}\left(y(z,x)\right)\mathrm{d}x.$$

Using the formulas for the correlation functions and the permutational definition of the determinant, we can put the integrand in the following form

$$\frac{\mathbf{1}_{V_\eta}(x)}{(2\pi)^k} \cdot \det\left[\frac{2\pi}{m} s_m(x_s - x_t)\right]_{s,t=1}^{k} \cdot \det\left[\frac{2\pi}{n} s_n\left(y_s - y_t\right)\right]_{s,t=1}^{k}$$
$$= \frac{\mathbf{1}_{V_\eta}(x)}{(2\pi)^k}\left(1 + \sum_{\sigma\neq \mathrm{id} \text{ or } \tau\neq\mathrm{id}} \operatorname{sgn}\sigma \operatorname{sgn}\tau \prod_{i=1}^{k} \frac{2\pi}{m} s_m(x_i - x_{\sigma(i)}) \cdot \frac{2\pi}{n} s_n(y_i - y_{\tau(i)})\right),$$

where the second summation runs through permutations σ and τ of k indices. The point is that each term in this sum tends to zero with m and n going to infinity as we have $\frac{2\pi}{m} s_m(x_i - x_{\sigma(i)}) \xrightarrow[m\to\infty]{\text{a.e.}} 0$ for i such that $i \neq \sigma(i)$, and $\frac{2\pi}{n} s_n(y_i - y_{\tau(i)}) \xrightarrow[n\to\infty]{\text{a.e.}} 0$ if $i \neq \tau(i)$ (see (2.8) and bear in mind the fact that actually y depends on m and n). Recall also that $\mathbf{1}_{W_\eta} \to 1$ and $\sum_\eta \mathbf{1}_{V_\eta} \to \mathbf{1}_{[0,2\pi)^k}$. Moreover, (2.4) yields that the whole sum is bounded by $(k!)^2/(2\pi)^k$. Therefore by Lebesgue's dominated convergence theorem we conclude that

$$\rho^{(k)}_{\Theta_{m,n}}(z) \xrightarrow[m,n\to\infty]{} \int \mathbf{1}_{[0,2\pi)^k}(x) \frac{1}{(2\pi)^k} \mathrm{d}x = 1,$$

which finishes the proof. $\square$

2.4.2 Proof of Theorem 2.5

In the course of the proof we will need three lemmas. Let us start with them.

2.6 Lemma. *Fix a positive integer s and a number $\gamma \in (0, 1/s)$. For a positive integer n define the set $\mathcal{L}_n = \{\ell = (\ell_1, \ldots, \ell_s),\ \mathbb{Z} \ni \ell_j \geq 0,\ \sum_{j=1}^{s} \ell_j = n\}$. Then*

$$\sum_{\ell\in\mathcal{L}_n, \exists j\ \ell_j/n \leq \gamma} \frac{1}{s^n}\frac{n!}{\ell!} = 1 - \sum_{\ell\in\mathcal{L}_n, \forall j\ \ell_j/n > \gamma} \frac{1}{s^n}\frac{n!}{\ell!} \xrightarrow[n\to\infty]{} 0. \tag{2.9}$$

Here we adopt the convention that $\ell! = \ell_1! \cdot \ldots \cdot \ell_s!$.

Proof. The first estimate we make is simply the union bound

$$\sum_{\ell, \exists j\ \ell_j/n\leq\gamma} \frac{1}{s^n}\frac{n!}{\ell!} \leq s \sum_{\ell, \ell_1/n \leq \gamma} \frac{1}{s^n}\frac{n!}{\ell!}.$$

Now we rearrange this sum

$$
\begin{aligned}
\sum_{\ell,\ell_1/n\le\gamma} \frac{1}{s^n}\frac{n!}{\ell!} &= \sum_{\ell_1=0}^{\lfloor\gamma n\rfloor} \frac{1}{s^n}\frac{n!}{\ell_1!(n-\ell_1)!} \sum_{\ell_2+\ldots+\ell_s\le n-\ell_1} \frac{(n-\ell_1)!}{\ell_2!\cdot\ldots\cdot\ell_s!} \\
&= \sum_{\ell_1=0}^{\lfloor\gamma n\rfloor} \frac{1}{s^n}\binom{n}{n-\ell_1}(s-1)^{n-\ell_1} \\
&= \sum_{k=n-\lfloor\gamma n\rfloor}^{n} \binom{n}{k}\left(1-\frac{1}{s}\right)^k\left(\frac{1}{s}\right)^{n-k}.
\end{aligned}
$$

Let $X_1, X_2, \ldots$ be i.i.d. Bernoulli random variables such that $\mathbb{P}(X_1 = 0) = 1/s = 1 - \mathbb{P}(X_1 = 1)$. Denote $S_n = X_1 + \ldots + X_n$. Then the last expression equals $\mathbb{P}(S_n \ge n - \lfloor\gamma n\rfloor)$. The second estimate we make is the following probabilistic bound

$$
\begin{aligned}
\mathbb{P}(S_n \ge n - \gamma n) &= \mathbb{P}\left(\frac{S_n - \mathbb{E}S_n}{n} \ge \frac{1}{s} - \gamma\right) \\
&\le \exp\big(-2n(1/s-\gamma)^2\big) \xrightarrow[n\to\infty]{} 0,
\end{aligned}
$$

where the inequality follows for instance from Hoeffding's inequality. $\square$

2.7 Lemma. *Let X be a random vector in $\mathbb{R}^n$ with a bounded density. Let $A\colon \mathbb{R}^n \longrightarrow \mathbb{R}^k$ be a linear mapping of rank r. Then there exists a constant C such that for any intervals $I_1, \ldots, I_k \subset \mathbb{R}$ of finite length we have*

$$
\mathbb{P}(AX \in I_1 \times \ldots I_k) \le C|I_{i_1}|\cdot\ldots\cdot|I_{i_r}|,
$$

where $1 \le i_1 < \ldots < i_r \le k$ are indices of those rows of the matrix A which are linearly independent.

Proof. Let $a_1, \ldots, a_k \in \mathbb{R}^n$ be rows of the matrix A. We know there are r of them, say $a_1, \ldots, a_r$, which are linearly independent. Thus there exists an invertible $r \times r$ matrix U such that

$$
U\begin{bmatrix} a_1 \\ \vdots \\ a_r \end{bmatrix} = \begin{bmatrix} e_1 \\ \vdots \\ e_r \end{bmatrix} =: E,
$$

where $e_i \in \mathbb{R}^n$ is the i-th vector of the standard basis of $\mathbb{R}^n$. Notice that

$$\begin{aligned}\mathbb{P}\left(AX \in I_1 \times \ldots \times I_k\right) &\leq \mathbb{P}\left(U^{-1}EX \in I_1 \times \ldots \times I_r\right) \\ &= \mathbb{P}\left((X_1, \ldots, X_r) \in U(I_1 \times \ldots I_r)\right) \\ &\leq C|U(I_1 \times \ldots \times I_k)| = C|\det U| \cdot |I_1| \cdot \ldots \cdot |I_r|,\end{aligned}$$

for the vector $(X_1, \ldots, X_r)$ also has a bounded density on $\mathbb{R}^r$. This finishes the proof. $\square$

2.8 Lemma. *Let A be a matrix of dimension $k \times j$, with entries in $\{0, 1\}$, and satisfying the following conditions*

(i) no two columns are equal.

(ii) no two rows are equal.

(iii) no zero row

Then, the rank of A is at least $\min(k, \lfloor \log_2 j \rfloor + 1)$.

Proof. (Due to Dima Gourevitch) Denote $r = \mathrm{rank} A$. The assertion of the lemma is equivalent to the statement that $2^r \geq j$ and if $2^r = j$ then $r = k$.

We may assume without loss of generality that the first r rows of A are linearly independent and the others are their linear combinations. Under this assumption, if two columns are identical in the first r coordinates then they are identical in all coordinates. By condition *(i)*, such columns do not exist. Therefore the $r \times j$ submatrix B which consists of the first r rows has distinct columns. As a result $j \leq 2^r$.

Now suppose $j = 2^r$. If $k > r$, consider the $r + 1$ row of A. It is a linear combination of the first r rows. Since the columns of B include the column $e_i = (0, .., 0, 1, 0, .., 0)$ for all $i = 1, \ldots, r$, the coefficient of each row is either 0 or 1. B includes also a column of all 1s, thus there is at most one nonzero coefficient (if there were more than one, a certain entry would be greater than 1). Consequently, the coefficient of exactly one row is 1, and all other coefficients vanish, because if all coefficients were zero, the $r + 1$ row would be zero which contradicts *(iii)*. Thus, the $r+1$-th row is identical to one of the first r rows - in contradiction to condition *(ii)*. $\square$

Proof of Theorem 2.5. Fix an integer $k \geq 1$ and finite intervals $I_1, \ldots, I_k \subset [0, \infty)$ which are mutually disjoint. We need to compute the probability of the event $\{\tau_M(I_j) > 0, j = 1, \ldots, k\}$ which means that in each interval I_j there is a rescaled eigenphase. Each such eigenphase is of the form $\frac{2^M}{2\pi}(\theta_1^{\epsilon_1} + \ldots + \theta_M^{\epsilon_M} \bmod 2\pi)$ for some $\epsilon = (\epsilon_1, \ldots, \epsilon_M) \in \{1, 2\}^M$. Therefore

$$\{\tau_M(I_j) > 0,\ j = 1, \ldots, k\} = \bigcup_{\epsilon} A_\epsilon,$$

where

$$A_\epsilon = \left\{ \sum_{i=1}^{M} \theta_i^{\epsilon_i^j} \bmod 2\pi \in \underbrace{\frac{2\pi}{2^M} I_j}_{J_j}, j = 1, \ldots, k \right\}, \tag{2.10}$$

and ϵ runs over the set

$$\mathcal{E} = \left\{ [\epsilon_i^j]_{i=1,\ldots,M}^{j=1,\ldots,k},\ \epsilon_i^j \in \{1,2\},\ \epsilon^u \neq \epsilon^v, \text{for } u \neq v, u, v = 1, \ldots, k \right\} \tag{2.11}$$

of all $k \times M$ matrices with entries $1, 2$ which have pairwise distinct rows $\epsilon^j = (\epsilon_1^j, \ldots, \epsilon_M^j) \in \{1, 2\}^M$, $j = 1, \ldots, k$ (j-th row ϵ^j describes the j-th eigenphase and since intervals are disjoint we assume the rows are distinct). Column vectors are denoted by $\epsilon_i = [\epsilon_i^1, \ldots, \epsilon_i^k]^T$, $i = 1, \ldots, M$.

We say that ϵ is *bad* if the collection of its column vectors $\{\epsilon_i, i \leq M\}$ has cardinality less than 2^k. Otherwise ϵ is called *good*. Obviously,

$$\mathbb{P}\left(\bigcup_{\text{good } \epsilon} A_\epsilon\right) \leq \mathbb{P}\left(\bigcup_{\epsilon} A_\epsilon\right) \leq \mathbb{P}\left(\bigcup_{\text{good } \epsilon} A_\epsilon\right) + \mathbb{P}\left(\bigcup_{\text{bad } \epsilon} A_\epsilon\right).$$

The strategy is to show that the contribution of bad ϵ vanishes for large M while good ϵ essentially provide the desired result $\prod_j |I_j|$ when M goes to infinity. So the proof will be divided into several parts.

Good ϵ.

The goal here is to prove

$$\lim_{\max_j |I_j| \to 0} \lim_{M \to \infty} \frac{1}{|I_1| \cdot \ldots \cdot |I_k|} \mathbb{P}\left(\bigcup_{\text{good } \epsilon} A_\epsilon\right) = 1, \tag{2.12}$$

with the required uniformity in the choice of the disjoint intervals I_j. By virtue of the fact that

$$\sum_{\text{good }\epsilon} \mathbb{P}\left(A_\epsilon\right) - \sum_{\substack{\text{good }\epsilon,\widetilde{\epsilon}\\ \epsilon\neq\widetilde{\epsilon}}} \mathbb{P}\left(A_\epsilon \cap A_{\widetilde{\epsilon}}\right) \leq \mathbb{P}\left(\bigcup_{\text{good }\epsilon} A_\epsilon\right) \leq \sum_{\text{good }\epsilon} \mathbb{P}\left(A_\epsilon\right)$$

it suffices to prove that

$$\lim_{M\to\infty} \sum_{\text{good }\epsilon} \mathbb{P}\left(A_\epsilon\right) = \prod |I_j| \tag{2.13}$$

uniformly, and that the correlations between two different good ϵ do not matter

$$\limsup_{\max_j |I_j|\to 0} \limsup_{M\to\infty} \frac{1}{\prod |I_j|} \sum_{\substack{\text{good }\epsilon,\widetilde{\epsilon}\\ \epsilon\neq\widetilde{\epsilon}}} \mathbb{P}\left(A_\epsilon \cap A_{\widetilde{\epsilon}}\right) = 0. \tag{2.14}$$

Let us now prove (2.13). The proof of (2.14) is deferred to the very end as we will need the ideas developed here as well as in the part devoted to bad ϵ.

Given $\epsilon \in \mathcal{E}$ and a vector $\alpha = [\alpha_1 \dots \alpha_k]^T \in \{1,2\}^k$ we count how many column vectors of ϵ equals α and call this number ℓ_α. Then $\sum_\alpha \ell_\alpha = M$. Note that ϵ is good iff all ℓ_α are nonzero. The crucial observation is that the probability of the event A_ϵ does depend only on the vector $\ell = (\ell_\alpha)_{\alpha\in\{1,2\}^k}$ associated with ϵ as described before. Indeed, the sum $\sum_{i=1}^M [\theta_i^{\epsilon_i^1} \dots \theta_i^{\epsilon_i^k}]^T \bmod 2\pi$ is identically distributed as the random vector $\sum_\alpha \psi(\alpha, \ell_\alpha) \bmod 2\pi$, where

$$\psi(\alpha,\ell_\alpha) = \begin{bmatrix} \psi_1(\alpha,\ell_\alpha) \\ \vdots \\ \psi_k(\alpha,\ell_\alpha) \end{bmatrix} = \begin{bmatrix} \theta_{i_1}^{\alpha_1} \\ \vdots \\ \theta_{i_1}^{\alpha_k} \end{bmatrix} + \dots + \begin{bmatrix} \theta_{i_{\ell_\alpha}}^{\alpha_1} \\ \vdots \\ \theta_{i_{\ell_\alpha}}^{\alpha_k} \end{bmatrix} \bmod 2\pi \tag{2.15}$$

is a sum modulo 2π of i.i.d. vectors. Note that the distribution of $\psi(\alpha, \ell_\alpha)$ does not depend on the choice of indices $i_1, \dots, i_{\ell_\alpha}$ but only on α and ℓ_α. Consequently, denoting by $\mathcal{E}_\ell$ the set of all ϵ such that there are exactly ℓ_α indices $1 \leq i_1 < \dots < i_{\ell_\alpha} \leq M$ for which $\epsilon_{i_1} = \dots = \epsilon_{i_{\ell_\alpha}} = \alpha$, we have that the value of $\mathbb{P}\left(A_\epsilon\right)$ is the same for all $\epsilon \in \mathcal{E}_\ell$. Clearly $\sharp\mathcal{E}_\ell = \frac{M!}{\ell!}$, whence

$$\sum_{\text{good }\epsilon} \mathbb{P}\left(A_\epsilon\right) = \sum_{\text{good }\ell} \frac{M!}{\ell!} \mathbb{P}\left(\sum_{\alpha\in\{1,2\}^k} \psi(\alpha,\ell_\alpha) \bmod 2\pi \in J_1 \times \dots \times J_k\right). \tag{2.16}$$

The idea is to identify those terms which will sum up to $\prod |I_i|$ and the rest which will be neglected in the limit of large M. To do this, set a positive parameter $\gamma < 1/2^k$ and let us call a good ℓ *very good* (*v.g.* for short) if $\ell_\alpha > \gamma M$ for every α and *quite good* (*q.g.* for short) otherwise. We claim that

$$\mathbb{P}\left(\sum \psi(\alpha, \ell_\alpha) \bmod 2\pi \in J_1 \times \ldots \times J_k\right) \leq C \prod |J_j|, \quad \text{for a good } \ell \tag{C1}$$

and

$$\mathbb{P}\left(\sum \psi(\alpha, \ell_\alpha) \bmod 2\pi \in J_1 \times \ldots \times J_k\right) = \frac{\prod |J_j|}{(2\pi)^k}\left(1 + \frac{r_\ell}{\sqrt{M}}\right), \quad |r_\ell| \leq C, \quad \text{for a very good } \ell, \tag{C2}$$

where C is a constant (from now on in this proof we adopt the convention that C is a constant depending only on k which may differ from line to line).

Let us postpone the proofs and see how to conclude (2.13). Notice that $\frac{\prod |J_j|}{(2\pi)^k} = \frac{1}{2^{kM}} \prod |I_j|$. Thus applying (C1) we obtain

$$\sum_{\text{q.g. } \ell} \mathbb{P}\left(\sum \psi(\alpha, \ell_\alpha) \bmod 2\pi \in J_1 \times \ldots \times J_k\right) \leq \prod |I_j| \cdot C \sum_{\text{q.g. } \ell} \frac{1}{2^{kM}} \frac{M!}{\ell!}.$$

By Lemma 2.6 this vanishes when $M \to \infty$. Now we deal with very good ℓ writing with the aid of (C2) that

$$\begin{aligned}\sum_{\text{v.g. } \ell} &\mathbb{P}\left(\sum \psi(\alpha, \ell_\alpha) \bmod 2\pi \in J_1 \times \ldots \times J_k\right) \\ &= \prod |I_j| \left(\sum_{\text{v.g. } \ell} \frac{1}{2^{kM}} \frac{M!}{\ell!} + \sum_{\text{v.g. } \ell} \frac{1}{2^{kM}} \frac{M!}{\ell!} \frac{r_\ell}{\sqrt{M}}\right).\end{aligned}$$

The first term in the bracket approaches 1 in the limit $M \to \infty$ due to Lemma 2.6, while the second one approaches 0 as it is bounded above by $C\frac{1}{\sqrt{M}}$.

Proof of (C1). Let us define the vectors

$$e_j = (\underbrace{2, \ldots, 2}_{j-1}, 1, \underbrace{2, \ldots, 2}_{k-j}) \in \{1, 2\}^k, \qquad j = 1, \ldots, k.$$

Since ℓ is good, in particular we have that $\ell_{e_j} > 0$, so denoting the random vector $\psi(e_j, \ell_{e_j})$ by Ψ^j we have

$$\sum_\alpha \psi(\alpha, \ell_\alpha) = (\Psi^1 + \ldots + \Psi^k) + \sum_{\alpha \notin \{e_1, \ldots, e_k\}} \psi(\alpha, \ell_\alpha).$$

By independence it is enough to show that the random vector $\Psi = \Psi^1 + \ldots + \Psi^k \mod 2\pi$ has a bounded density on $[0, 2\pi)^k$. Equation (2.15) yields that

$$\Psi^j = (\underbrace{Y_j, \ldots, Y_j}_{j-1}, X_j, \underbrace{Y_j, \ldots, Y_j}_{k-j}),$$

where (X_j, Y_j) are independent random vectors on $[0, 2\pi)^2$ with the same distributions as the vectors $(\theta_1^1 + \ldots + \theta_{\ell_{e_j}}^1 \mod 2\pi, \theta_1^2 + \ldots + \theta_{\ell_{e_j}}^2 \mod 2\pi)$ respectively. Clearly, the vector (X_j, Y_j) has a bounded density on $[0, 2\pi)^2$ because the vector (θ_1^1, θ_1^2) has a bounded density. Therefore the vector $(X_1, Y_1, \ldots, X_k, Y_k)$ has a bounded density on $[0, 2\pi)^{2k}$. A certain linear transformation with determinant 1 maps this vector to $(\Psi^1 + \ldots + \Psi^k, Y_1, \ldots, Y_k)$ which consequently also has a bounded density. We project it to the first k coordinates and then take care of addition modulo 2π obtaining that Ψ has a bounded density, which finishes the proof. □

Proof of (C2). Given a vector $\alpha \in \{1, 2\}^k$ let Θ^α denote the random vector in $[0, 2\pi)^k$ identically distributed as the vector $(\theta_1^{\alpha_1}, \ldots, \theta_1^{\alpha_k})$. Take independent copies $\Theta_1^\alpha, \Theta_2^\alpha, \ldots$ of Θ^α such that the family $\{\Theta_1^\alpha, \Theta_2^\alpha, \ldots\}_{\alpha \in \{1,2\}^k}$ also consists of independent random vectors. Then $\mathbb{E}\Theta^\alpha = [\pi, \ldots, \pi]^T$, and

$$\begin{aligned}
p_{\ell,M} &= \mathbb{P}\left(\sum_\alpha \psi(\alpha, \ell_\alpha) \mod 2\pi \in J_1 \times \ldots \times J_k\right) \\
&= \mathbb{P}\left(\sum_\alpha \sum_{l=1}^{\ell_\alpha} \Theta_l^\alpha \mod 2\pi \in J_1 \times \ldots \times J_k\right) \\
&= \sum_{i_1, \ldots, i_k = 0}^{M-1} \mathbb{P}\left(\sum_\alpha \sum_{l=1}^{\ell_\alpha} \Theta_l^\alpha \in (J_1 + 2\pi i_1) \times \ldots \times (J_k + 2\pi i_k)\right) \\
&= \sum_i \mathbb{P}\Bigg(\sum_\alpha \sum_{l=1}^{\ell_\alpha} \frac{\Theta_l^\alpha - \mathbb{E}\Theta_l^\alpha}{\sqrt{M}} \in \frac{1}{\sqrt{M}}(J_1 + 2\pi(i_1 - M/2)) \times \ldots \\
&\qquad\qquad \times \frac{1}{\sqrt{M}}(J_k + 2\pi(i_k - M/2))\Bigg).
\end{aligned}$$

To ease the notation we introduce new indices

$$j = \left(i_1 - \frac{M}{2}, \ldots, i_k - \frac{M}{2}\right) \in \left\{-\frac{M}{2}, -\frac{M}{2} + 1, \ldots, \frac{M}{2} - 1\right\}^k,$$

sets

$$K_{j,M} = \frac{1}{\sqrt{M}}(J_1 + 2\pi j_1) \times \ldots \times \frac{1}{\sqrt{M}}(J_k + 2\pi j_k),$$

and the vector

$$S_M = \sum_\alpha \sum_{l=1}^{\ell_\alpha} \frac{\Theta_l^\alpha - \mathbb{E}\Theta_l^\alpha}{\sqrt{M}}.$$

Now we intend to use the local Central Limit Theorem of [BRR86]. Indeed, due to independence such a theorem should hopefully yield that S_M has a normal distribution for large M. To be more precise, let us consider the matrix $\operatorname{Cov} S_M = \sum_\alpha \frac{\ell_\alpha}{M} \operatorname{Cov} \Theta^\alpha$ and its eigenvalues. Since for any $x \in \mathbb{R}^k$

$$x^T(\operatorname{Cov} S_M)x = \sum_\alpha \frac{\ell_\alpha}{M} x^T(\operatorname{Cov} \Theta^\alpha)x \leq \underbrace{\max_\alpha \| \operatorname{Cov} \Theta^\alpha \|}_{C} |x|^2,$$

it is clear that the largest eigenvalues are uniformly (with respect to M) bounded by C, which depends only on k. To provide a uniform bound for the smallest eigenvalues let us observe that (recall that $e_i = (2, \ldots, 2, 1, 2, \ldots, 2)$)

$$x^T(\operatorname{Cov} S_M)x \geq \sum_{j=1}^k \frac{\ell_{e_j}}{M} x^T(\operatorname{Cov} \Theta^{e_j})x > \gamma x^T\Big(\sum_{j=1}^k \operatorname{Cov} \Theta^{e_j}\Big)x \geq \gamma \cdot \frac{\pi^2}{3}|x|^2,$$

where the second inequality holds because ℓ is very good.

It is a matter of a direct computation to see the last inequality since for $k \geq 2$ we have $\sum_{j=1}^k \operatorname{Cov} \Theta^{e_j} = \big((k-2)\pi^2/3 - 2\big)[1 \ldots 1]^T[1 \ldots 1] + \operatorname{diag}(2 + 2\pi^2/3, \ldots, 2 + 2\pi^2/3)$ and for $k = 1$ the sum equals $\pi^2/3$. Therefore, with the matrix B_M given by

$$B_M^2 = (\operatorname{Cov} S_M)^{-1}$$

we get that

$$\frac{1}{C}|x| \leq |B_M x| \leq C|x|.$$

Therefore the assumptions of [BRR86, Corollary 19.4] are satisfied (for the family of independent random vectors $\{\Theta_1^\alpha, \Theta_2^\alpha, \ldots\}_{\alpha \in \{1,2\}^k}$), so the vector $B_M S_M$ possesses a density q_M and

$$\sup_{x \in \mathbb{R}^k} \big(1 + |x|^{k+2}\big) \left(q_M(x) - \phi(x) - \frac{1}{\sqrt{M}} P_M(x)\phi(x)\right) = O(M^{-k/2}),$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}^k} e^{-|x|^2/2}$ is the density of the standard normal distribution in $\mathbb{R}^k$ and P_M is a polynomial of degree $k-1$ whose coefficients depends on the cumulants of the vectors $B_M \Theta^\alpha$. We may put it differently, i.e.

$$q_M(x) = \phi(x) + \frac{1}{\sqrt{M}} \Bigg(\underbrace{P_M(x)\phi(x) + \frac{f_M(x)}{1+|x|^{k+2}}}_{h_M(x)} \Bigg),$$

for some functions f_M uniformly bounded $\sup_M \sup_{x\in\mathbb{R}^k} |f_M(x)| = C < \infty$. Therefore, denoting $L_{j,M} = B_M K_{j,M}$,

$$\begin{aligned} p_{\ell,M} &= \sum_j \mathbb{P}\left(S_M \in K_{j,M}\right) = \sum_j \mathbb{P}\left(B_M S_M \in B_M K_{j,M}\right) \\ &= \sum_j \int_{L_{j,M}} q_M = \sum_j \int_{L_{j,n}} \phi + \frac{1}{\sqrt{M}} \sum_j \int_{L_{j,n}} h_M \qquad (2.17) \\ &= a_M + \frac{1}{\sqrt{M}} b_M. \end{aligned}$$

Let us first deal with the error term b_M. Denoting

$$\kappa = \frac{|J_1| \cdot \ldots \cdot |J_k|}{(2\pi)^k}$$

we are to show that

$$|b_M| \le C\kappa. \qquad (2.18)$$

To do this we estimate the integrated function

$$|h_M(x)| \le |P_M(x)|\phi(x) + \frac{C}{1+|x|^{k+2}}.$$

Therefore we define

$$h(x) = |P_M(x)|\phi(x) + \frac{C}{1+|x|^{k+2}}$$

and then $|b_M| \le \sum_j \int_{L_{j,M}} h$. Introduce the boxes

$$F_{j,M} = B_M \left(\frac{1}{\sqrt{M}}([0,2\pi) + 2\pi j_1) \times \ldots \times \frac{1}{\sqrt{M}}([0,2\pi) + 2\pi j_k) \right)$$

and observe that

$$\int_{L_{j,M}} h = \frac{|L_{j,M}|}{|F_{j,M}|} |F_{j,M}| \frac{1}{|L_{j,n}|} \int_{L_{j,M}} h \le \kappa |F_{j,M}| \sup_{L_{j,M}} h \le \kappa |F_{j,M}| \sup_{F_{j,M}} h.$$

Since $\mathrm{diam} F_{j,M} \leq C\frac{2\pi\sqrt{k}}{\sqrt{M}} \xrightarrow[M\longrightarrow\infty]{} 0$, the sets $F_{j,M}$ are pairwise disjoint and sum up to $B_M[-\pi\sqrt{M}, \pi\sqrt{M})^k$, we can infer that the sum $\sum_j |F_{j,M}| \sup_{F_{j,M}} h$ converges to $\int_{\mathbb{R}^k} h = C < \infty$. Hence, this sum is bounded by C and we get (2.18).

Now we handle the main term a_M. We prove it equals κ up to another error $\kappa\frac{C}{\sqrt{M}}$. Let $A_{j,M} : \mathbb{R}^k \longrightarrow \mathbb{R}^k$ be the linear isomorphism mapping $F_{j,M}$ onto $L_{j,M}$. It equals $B_M \tilde{A}_{j,M} B_M^{-1}$, where $\tilde{A}_{j,M}$ is the linear mapping transforming the box $B_M^{-1}F_{j,M}$ onto the box $B_M^{-1}L_{j,M}$, whence $|\det A_{j,M}| = \kappa$. Thus, changing the variable we obtain

$$\int_{L_{j,M}} \phi(x)\mathrm{d}x = \kappa \int_{F_{j,M}} \phi(A_{j,M}x)\mathrm{d}x.$$

Notice that $A_{j,M}x$ is close to x, whenever $x \in F_{j,M}$, for

$$|A_{j,n}x - x| \leq \mathrm{diam} F_{j,M}, \qquad x \in F_{j,M}.$$

Consequently, on $F_{j,M}$, $\phi(A_{j,M}x)$ is close to $\phi(x)$. Strictly, we use the mean value theorem and get

$$\int_{L_{j,M}} \phi(x)\mathrm{d}x = \kappa \int_{F_{j,M}} \phi(x)\mathrm{d}x + \kappa \int_{F_{j,M}} \nabla\phi_V(\eta_x) \cdot (A_{j,M}x - x)\mathrm{d}x,$$

for some mean points $\eta_x \in [x, A_{j,M}x]$. This results in

$$\begin{aligned} a_M &= \sum_j \int_{L_{j,M}} \phi(x)\mathrm{d}x = \kappa \sum_j \int_{F_{j,M}} \phi + \kappa \sum_j \int_{F_{j,M}} \nabla\phi(\eta_x) \cdot (A_{j,M}x - x)\mathrm{d}x \\ &= \kappa\Bigg(1 - \underbrace{\int_{\mathbb{R}^k \setminus B_M[-\pi\sqrt{M}, \pi\sqrt{M})^k} \phi}_{c_M} + \underbrace{\sum_j \int_{F_{j,M}} \nabla\phi_V(\eta_x) \cdot (A_{j,M}x - x)\mathrm{d}x}_{d_M}\Bigg). \end{aligned}$$

We are almost done. Clearly c_M converges to 0 faster that $1/\sqrt{M}$, so $|c_M| \leq C/\sqrt{M}$. For d_M we use the Cauchy-Schwarz inequality and integrability of $|\nabla\phi(\eta_x)|$

$$\begin{aligned} |d_M| &\leq \sum_j \int_{F_{j,M}} |\nabla\phi(\eta_x)||A_{j,M}x - x|\mathrm{d}x \\ &\leq \mathrm{diam} F_{j,M} \int_{\bigcup F_{j,M}} |\nabla\phi(\eta_x)|\mathrm{d}x \leq \frac{C}{\sqrt{M}}. \end{aligned}$$

This completes the proof of (C2). $\square$

We have proved claims (C1) and (C2), so the proof of the part concerning good ϵ is now complete. Let us proceed to tackle bad ϵ.

Bad ϵ.

The goal here is to show that

$$\lim_{M\to\infty} \mathbb{P}\left(\bigcup_{\text{bad } \epsilon} A_\epsilon\right) = 0, \tag{2.19}$$

again, with the required uniformity. Obviously it suffices to show that

$$\sum_{\text{bad } \epsilon} \mathbb{P}(A_\epsilon) \xrightarrow[M\to\infty]{} 0.$$

Let $\mathcal{F}_j$ be the set of those bad ϵ for which the cardinality of the set $\{\epsilon_i, i \leq M\}$ equals j. Observe that $\sharp\mathcal{F}_j \leq j^M$. With the aid of Lemma 2.8 we will show that

$$\forall \epsilon \in \mathcal{F}_j \ \mathbb{P}(A_\epsilon) \leq C \cdot 2^{-M(1+\lfloor \log_2 j \rfloor)} \cdot O\big((\max_j |I_j|)^{1+\lfloor \log_2 j \rfloor}\big), \tag{2.20}$$

when $\max_j |I_j| \longrightarrow 0$. This will finish the proof, for

$$\begin{aligned}
\sum_{\text{bad } \epsilon} \mathbb{P}(A_\epsilon) &\leq C \cdot O(\max_j |I_j|) \sum_{j=1}^{2^k-1} j^M \cdot 2^{-M(1+\lfloor \log_2 j \rfloor)} \\
&= C \cdot O(\max_j |I_j|) \sum_{j=1}^{2^k-1} 2^{-M(1+\lfloor \log_2 j \rfloor - \log_2 j)} \xrightarrow[M\to\infty]{} 0.
\end{aligned} \tag{2.21}$$

For the proof of (2.20) fix $\epsilon \in \mathcal{F}_j$. We have seen that

$$\mathbb{P}(A_\epsilon) = \mathbb{P}\left(\sum \psi(\alpha, \ell_\alpha) \bmod 2\pi \in J_1 \times \ldots \times J_k\right)$$

and we know that there are exactly j numbers ℓ_α which are nonzero, say those which correspond to vectors $\alpha^1, \ldots, \alpha^j \in \{1,2\}^k$. Denote $\Psi^i = \psi(\alpha^i, \ell_{\alpha^i})$, $i = 1, \ldots, j$ and consider the random vector $S_j = \Psi^1 + \ldots + \Psi^j$ in $\mathbb{R}^k$. As in the proof of Claim (C1) we observe that S_j is a linear image of the vector $(X_1, Y_1, \ldots, X_j, Y_j)$. This mapping is given by the matrix $A = [a_{vi}]$ where

$$a_{v,2i-1} = \begin{cases} 1, & \alpha_v^i = 1 \\ 0, & \alpha_v^i = 2 \end{cases}, \qquad a_{v,2i} = \begin{cases} 0, & \alpha_v^i = 1 \\ 1, & \alpha_v^i = 2 \end{cases}.$$

By Lemma 2.7 we obtain

$$
\begin{aligned}
\mathbb{P}\left(S_j \bmod 2\pi \in J_1 \times \ldots \times J_k\right) &\le C \max\left(|J_{i_1}| \cdot \ldots \cdot |J_{i_r}|\right) \\
&= C \cdot O\big((\max_j |I_j|)^r\big) \cdot 2^{-Mr},
\end{aligned}
\tag{2.22}
$$

where $r = \operatorname{rank} A$. The number r does not change if we replace the $2i$-th column of A with the vector e with 1 at each its entry, as the sum of $2i-1$-th and $2i$-th columns is just e. Now taking only the columns $1, 2, 3, 5, \ldots, 2j-1$ we get the matrix B which has the same rank as A. It has $j+1$ columns and fulfils the assumptions of Lemma 2.8 (it has no zero row as the second column consists of all 1s). Thus $r \ge \min(1 + \lfloor \log_2(1+j) \rfloor, k)$ and when $j < 2^k - 1$ this minimum equals $1 + \lfloor \log_2(1+j) \rfloor \ge 1 + \lfloor \log_2 j \rfloor$. If $j = 2^k - 1$ in the matrix A there must be two identical columns, one with even, say $2u$, and one with odd, say $2v-1$ index, which means that the u-th and the v-th column of B add up to e, so the v-th column may be erased and the rank of B does not change. Therefore we apply the lemma to the matrix B with erased the v-th column which is of size $k \times j$ and get again $r \ge \min(1 + \lfloor \log_2 j \rfloor, k) = 1 + \lfloor \log_2 j \rfloor$. This completes the proof of (2.20).

Pairs of good ϵ, i.e. the proof of (2.14).

We denote by $\Theta_i(\epsilon)$ the random vector $(\theta_i^{\epsilon_i^1}, \ldots, \theta_i^{\epsilon_i^k})$. By the definition of A_ϵ we may write

$$
A_\epsilon \cap A_{\widetilde{\epsilon}} = \left\{ \sum_{i=1}^{M} \begin{bmatrix} \Theta_i(\epsilon) \\ \Theta_i(\widetilde{\epsilon}) \end{bmatrix} \bmod 2\pi \in \begin{matrix} J_1 \times \ldots \times J_k \\ J_1 \times \ldots \times J_k \end{matrix} \right\}. \tag{2.23}
$$

Since the intervals J_u and J_v are disjoint for $u \ne v$, we may restrict ourselves to those ϵ and $\widetilde{\epsilon}$ for which $\epsilon^u \ne \widetilde{\epsilon}^v$ whenever $u \ne v$, $u, v = 1, \ldots, k$ as otherwise the event $A_\epsilon \cap A_{\widetilde{\epsilon}}$ is impossible. However it might happen that $\epsilon^u = \widetilde{\epsilon}^u$. Let us count for how many u it takes place, i.e. given $s \in \{1, \ldots, k\}$ let $\mathcal{P}_s$ be the set of all considered unordered pairs $\{\epsilon, \widetilde{\epsilon}\}$ for which there are exactly $k-s$ indices $1 \le u_1 < \ldots < u_{k-s} \le k$ such that $\epsilon^{u_j} = \widetilde{\epsilon}^{u_j}$, $j = 1, \ldots, k-s$. The value $s=0$ is excluded as $\epsilon \neq \widetilde{\epsilon}$. We have

$$\sum_{\epsilon\neq\widetilde{\epsilon}} \mathbb{P}\left(A_\epsilon \cap A_{\widetilde{\epsilon}}\right) = \sum_{s=1}^{k} \sum_{\{\epsilon,\widetilde{\epsilon}\}\in\mathcal{P}_s} \mathbb{P}\left(A_\epsilon \cap A_{\widetilde{\epsilon}}\right).$$

Thus we fix s and prove that

$$\limsup_{\max_j |I_j|\to 0} \limsup_{M\to\infty} \frac{1}{\prod |I_j|} \sum_{\{\epsilon,\widetilde{\epsilon}\}\in\mathcal{P}_s} \mathbb{P}\left(A_\epsilon \cap A_{\widetilde{\epsilon}}\right) = 0.$$

There are two cases. A pair $\{\epsilon,\widetilde{\epsilon}\}\in\mathcal{P}_s$ can be *good* which means

$$\sharp\left\{ \left[\begin{smallmatrix}\epsilon_i\\ \widetilde{\epsilon}_i\end{smallmatrix}\right],\ i=1,\dots,M\right\} \geq 2^{k+s},$$

or, otherwise we call it *bad*. We obtain a decomposition $\mathcal{P}_s = \mathcal{P}_s^{\mathrm{good}} \cup \mathcal{P}_s^{\mathrm{bad}}$. Now for a good pair, applying the reasoning already used for bad ϵ, i.e. combining lemmas 2.7 and 2.8, we get the estimate

$$\begin{aligned}\mathbb{P}\left(A_\epsilon \cap A_{\widetilde{\epsilon}}\right) &\leq C|J_1|\cdot\ldots\cdot|J_k|\big(\max_{j=1,\dots,k}|J_j|\big)^s\\ &= \frac{C}{2^{(k+s)M}}\left(\prod |I_j|\right)\big(\max_j |I_j|\big)^s.\end{aligned}$$

But $\sharp\mathcal{P}_s^{\mathrm{good}} \leq \sharp\mathcal{P}_s \leq \binom{k}{s}\cdot 2^{(k+s)M}$, so

$$\limsup_{\max_j |I_j|\to 0} \limsup_{M\to\infty} \frac{1}{\prod |I_j|} \sum_{\{\epsilon,\widetilde{\epsilon}\}\in\mathcal{P}_s^{\mathrm{good}}} \mathbb{P}\left(A_\epsilon \cap A_{\widetilde{\epsilon}}\right) = 0.$$

For a bad pair $\{\epsilon,\widetilde{\epsilon}\}$ we know that there are $k+s$ different rows and at most $2^{k+s}-1$ different columns in the matrix $\left[\begin{smallmatrix}\epsilon\\ \widetilde{\epsilon}\end{smallmatrix}\right]$. Hence we repeat the argument of the part concerning bad ϵ. As in that part we use Lemma 2.8 in order to establish an appropriate inequality in the spirit of (2.20). Then we follow the estimate of (2.21) and conclude that

$$\lim_{M\to\infty} \sum_{\{\epsilon,\widetilde{\epsilon}\}\in\mathcal{P}_s^{\mathrm{bad}}} \mathbb{P}\left(A_\epsilon \cap A_{\widetilde{\epsilon}}\right) = 0.$$

This finishes the proof of Theorem 2.5. □

2.5 Notes and comments

The idea of studying tensor products of random unitary matrices stemmed from the author's discussions with the physicists M. Kuś and K. Życzkowski.

The author discussed the related mathematical problems with O. Zeitouni. Under his supervision, the author proved first Theorem 2.1(b) in the particular case of $m = n$. After some time, guided by some key ideas produced by O. Zeitouni, the author proved Theorem 2.5. K. Życzkowski's PhD student at the time, M. Smaczyński, provided some numerical data which, along with the main results, Theorem 2.1(b) ($m = n$) and Theorem 2.5, were published in article [TSK+12]. Later, working under the supervision of N. O'Connell, the author proved Theorem 2.1 and published it in [Tko13]. The asymptotics of the extreme statistics of spectra of tensor products of random unitary matrices have recently been investigated numerically (see [STKZ13]). For a single random unitary matrix, such statistics are well understood (see [BAB13]).

Sections 2.3.1 and 2.4.1 are based on [Tko13], whereas sections 2.3.2 and 2.4.2 are based on [TSK+12].

We would like to end this discussion with a conjecture that generalises our main results. The conjecture appeared in [TSK+12].

2.9 Conjecture. *Let $\theta_j^1, \ldots, \theta_j^N$, $j = 1, \ldots, M$ be the eigenphases of independent random unitary matrices $U_1, \ldots, U_M$ of size N. Define the point process $\tau_{M,N}$ of the rescaled eigenphases of the matrix $U_1 \otimes \ldots \otimes U_M$ as*

$$\tau_{M,N}(D) := \sum_{\epsilon=(\epsilon_1,\ldots,\epsilon_M)\in\{1,\ldots,N\}^M} \boldsymbol{1}_{\left\{\frac{N^M}{2\pi}\left(\theta_1^{\epsilon_1}+\ldots+\theta_M^{\epsilon_M} \mod 2\pi\right)\in D\right\}}, \tag{2.24}$$

for any compact set $D \subset [0, \infty)$. Then, for each k there exists a continuous function $\delta_k : [0, \infty) \to [0, \infty)$ with $\delta_k(0) = 0$ so that for any mutually disjoint intervals $I_1, \ldots, I_k \subset [0, \infty)$

$$\limsup \frac{\mathbb{P}\left(\tau_{M,N}(I_1) > 0, \ldots, \tau_{M,N}(I_k) > 0\right)}{|I_1| \cdot \ldots \cdot |I_k|} \le (1 + \delta_k(\max_j |I_j|)),$$
$$\liminf \frac{\mathbb{P}\left(\tau_{M,N}(I_1) > 0, \ldots, \tau_{M,N}(I_k) > 0\right)}{|I_1| \cdot \ldots \cdot |I_k|} \ge (1 - \delta_k(\max_j |I_j|))$$

with fixed $N > 2$ and $M \to \infty$, or $N \to \infty$ and fixed $M > 2$.

Chapter 3

Invertibility of L_1 operators.

3.1 Introduction

In this chapter we will frequently work with the Lebesgue space $L_p[0,1]$ for $p \geq 1$ of measurable functions $f\colon [0,1] \longrightarrow \mathbb{R}$ with finite p-norm which will be denoted by $\|f\|_{L_p} = \|f\|_p = \left(\int_0^1 |f(x)|^p \mathrm{d}x\right)^{1/p}$. We will be interested in linear operators, which when restricted to functions with small support, are invertible. By the *support* of a function f, $\mathrm{supp}(f)$ we mean, as usual, the set of points where the function is *not* zero-valued. For the present purpose, whether or not the support is small will be decided by its Lebesgue measure which we denote $|\cdot|$.

Given two positive constants ϵ and c, let us define the class $\mathcal{T}_{\epsilon,c}(L_p)$ of linear operators T acting on $L_p[0,1]$, with values in the same space, with norm at most one, and having the property that restricted to $L_p(A)$ for any measurable subset A of $[0,1]$ of Lebesgue measure c, they are invertible with $\|(T|_{L_p(A))})^{-1}\| \geq \epsilon^{-1}$; in other words, satisfying

$$\forall f \in L_p[0,1] \;\; |\mathrm{supp}(f)| \leq c \implies \|Tf\|_{L_p} \geq \epsilon \|f\|_{L_p}.$$

G. Schechtman posed the following question (personal communication). In an equivalent form it was asked by B. Johnson in connection with A. Nasseri's question on Mathoverflow [Nas12]).

3.1 Problem (Schechtman)**.** Fix $\epsilon, c > 0$. Is it true that there exists a natural number $k = k(\epsilon, c)$ depending only on ϵ and c such that for any operator $T \in \mathcal{T}_{\epsilon,c}(L_1)$ we can find $\delta > 0$ and k functions $g_1, \ldots, g_k \in L_\infty[0,1]$ such that

$$\|Tf\|_{L_1} \geq \delta \|f\|_{L_1}$$

for every function f from the subspace $\{f \in L_1[0,1],\ \int_0^1 fg_j = 0, j \leq k\}$?

Loosely speaking, the question asks whether local invertibility of an operator on L_1 implies its global invertibility on a subspace of fixed codimension. The aim of the first part of this chapter is to explain why the answer to this question is negative. Our approach will go through the study of the finite dimensional analogue of Schechtman's question. We will collect some recent results from combinatorics discovered in the context of sparse signal recovery and, based on that, we will build a counter-example in L_1. In the second part of this chapter, we will establish some Sobolev-type inequalities and find a certain large class of convolution operators which are nicely invertible. This will hopefully emphasise even more how careful we have to be in choosing the right operator to answer Schechtman's question in the negative, since taking *decent* convolution operators will not work.

To get some intuition about the main issue, let us look at an example of a locally invertible operator which is also globally invertible. Consider $T\colon L_1([0,1]) \longrightarrow L_1([0,1]^2)$ given by the formula

$$(Tf)(x,y) = \frac{1}{2}\big(f(x) - f(y)\big), \qquad x, y \in [0,1].$$

(Since $L_1([0,1]^2)$ is isometrically isomorphic to $L_1([0,1])$, this T can be used to define an operator from $L_1([0,1])$ to itself, but for simplicity we will work with T.) Clearly, $\|Tf\| = \frac{1}{2}\iint |f(x) - f(y)| \leq \int |f| = \|f\|$, and hence $\|T\| \leq 1$. (In fact, $\|T\| = 1$). If $|\mathrm{supp}(f)| \leq c$, we get

$$\|Tf\| \geq \int_{\mathrm{supp}(f)} \int_{(\mathrm{supp}(f))^c} |f(x) - f(y)| = (1-c)\|f\|,$$

so that T is in the class $\mathcal{T}_{\epsilon,c}(L_1)$ with $\epsilon = 1 - c$. However, if f is in the subspace of functions with mean 0, $f \in V = \{f \in L_1[0,1],\ \int_0^1 f = 0\}$, then

$$\|Tf\| \geq \frac{1}{2}\int_0^1 \left| \int_0^1 \big(f(x) - f(y)\big)\mathrm{d}y \right| \mathrm{d}x = \frac{1}{2}\|f\|.$$

Therefore operator T, begin in the class $\mathcal{T}_{\epsilon,c}(L_1)$, is moreover globally invertible on a 1-codimensional subspace.

To finish this introduction, we remark that Schechtman's question really touches upon some geometric subtleties of L_1. We can ask a similar question about the L_p spaces for $p > 1$, but the answer is much simpler.

3.2 Problem. Let $p > 1$ and fix $\epsilon, c > 0$. Is it true that there exists a natural number $k = k(\epsilon, c, p)$ such that for any operator $T \in \mathcal{T}_{\epsilon,c}(L_p)$ we can find $\delta > 0$ and k functions $g_1, \ldots, g_k \in L_q[0,1]$, $1/p + 1/q = 1$, such that

$$\|Tf\|_{L_p} \geq \delta \|f\|_{L_p}$$

for every function f from the subspace $\{f \in L_p[0,1];\ \int_0^1 f g_j = 0, j \leq k\}$?

The answer is negative. To see this, fix $1 < p < \infty$ and let $g_1, g_2, \ldots$ be i.i.d. standard Gaussian random variables (mean 0, variance 1) on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let G be the closed span of $\{g_i,\ i \geq 1\}$ in $L_p((\Omega, \mathcal{F}, \mathbb{P}))$. It is well known that G is *complemented* in L_p, $1 < p < \infty$ (see first few pages of Chapter 2 in [Pis89]). So there is a linear projection operator Q from L_p onto G which is continuous. Consider $T = I - Q$.

First, there is no $\delta > 0$ and a subspace V of the form $\{X \in L_p,\ \mathbb{E}XY_j = 0,\ j = 1, \ldots, k\}$ for some $Y_1, \ldots, Y_k \in L_q$, $1/p + 1/q = 1$, such that $\|TX\|_p \geq \delta \|X\|_p$ for all $X \in V$. Indeed, considering $X = \sum_{i=1}^n \alpha_i g_i$ we have $TX = 0$, but for any $n > k$ it is possible to find a nonzero sequence $(\alpha_i)_{i=1}^n$ such that $X \in V$.

Second, we shall show that $T/\|T\|$ is in the class $\mathcal{T}_\epsilon(L_p)$ for some positive constant $\epsilon = \epsilon(p)$. To this end, fix a random variable $X \in L_2$ with the L_p norm 1 and let $A = \{X \neq 0\}$ be the support of X. We assume that $\mathbb{P}(A) \leq 1/2$ and we want to show that $\|TX\|_p \geq \epsilon$ as then $\left\| \frac{T}{\|T\|} X \right\| \geq \frac{\epsilon}{\|T\|} \geq \frac{\epsilon}{1 + \|Q\|}$. Notice that

$$\|TX\|_p \geq \|(TX) \cdot \mathbf{1}_{A^c}\|_p = \|(X - QX) \cdot \mathbf{1}_{A^c}\|_p = \|(QX) \cdot \mathbf{1}_{A^c}\|_p.$$

Because Q is a projection on G, the last expression becomes

$$\frac{\|QX\|_p}{\|g_1\|_p}\|g_1\mathbf{1}_{A^c}\|_p$$

(we can think of QX as a sum $\sum \alpha_i g_i$ which has the same distribution as $(\sum \alpha_i^2)^{1/2} g_1)$. As this is bounded below by $\eta_p\|QX\|_p$ with η_p being a positive constant,

$$\eta_p = \frac{\inf_A\{\|g_1\mathbf{1}_{A^c}\|_p,\ \mathbb{P}(A) = 1/2\}}{\|g_1\|_p},$$

we have

$$\|TX\|_p \geq \eta_p\|QX\|_p,$$

so in view of $1 = \|X\|_p = \|(Q+T)X\|_p \leq \|QX\|_p + \|TX\|_p$, we obtain

$$\|TX\|_p \geq \frac{\eta_p}{1+\eta_p}.$$

3.2 A finite dimensional analogue

Here we will work with the space ℓ_1^n ($\mathbb{R}^n$ equipped with the ℓ_1-norm $\|x\|_{\ell_1^n} = |x_1| + \ldots + |x_n|$). Informally we will distinguish *short* vectors $x \in \mathbb{R}^n$ meaning that their support, $\operatorname{supp}(x)$, which is the set of indices of nonzero coordinates of x, is small. By $\#A$, we denote the cardinality of A.

By analogy with the L_1 case, let us consider the class $\mathcal{T}_{\epsilon,c}^n$ of linear operators $T\colon \ell_1^n \longrightarrow \ell_1^n$ acting on ℓ_1^n, with norm at most one, and having the property that T is nicely invertible on the set of short vectors,

$$\forall x \in \mathbb{R}^n \quad \#\operatorname{supp}(x) \leq cn \implies \|Tx\|_{\ell_1^n} \geq \epsilon\|x\|_{\ell_1^n}.$$

Instead of asking about rather subtle invertibility properties of such operators, to begin with, let us observe that if $T \in \mathcal{T}_{\epsilon,c}^n$ for some $\epsilon > 0$ and $c \in (0,1)$ then

$$\dim\ker T < (1-c)n + 1 \tag{3.1}$$

as for any subspace V in $\mathbb{R}^n$ of dimension k there is a choice of k coordinates such that every vector from V is determined by these coordinates, hence such a subspace contains a nonzero vector with at most $n-k+1$ nonzero coordinates.

What is the sharp bound for $\dim\ker T$? If the answer to Schechtman's problem was positive, we would expect that there would be an integer $k = k(\epsilon)$ such that $\dim\ker T \leq k(\epsilon)$, for every $T \in \mathcal{T}_\epsilon^n$. As we shall see in the next section, the correct estimate is essentially given by the above simple bound.

3.3 Expanders

We say that a bipartite simple graph $G = (A, B, E)$ is an (r, θ, ϵ) *regular unbalanced expander* if it is left r-regular (every vertex from A has degree r) and for every subset X in A with $\#X \leq \theta \cdot \#A$ the set $\mathrm{Neighb}(X)$ of its neighbours is large, $\#\mathrm{Neighb}(X) \geq (1-\epsilon)r\#X$.

Of course, if we took $\#B = r \cdot \#A$ with all vertices in B of degree 1, then we would get an $(r, 1, 0)$ expander. The goal is to make $\#B, r, \epsilon$ as small as possible with θ being a positive constant (ideally close to 1). In particular, we are interested in $\#B$ of the same order as $\#A$, say $\#B \geq \frac{1}{2}\#A$ (the size of B cannot be too small as $\#B \geq (1-\epsilon)r\theta\#A$).

It was first observed by M. Pinsker that expanders exist (see [Pin73]). Probabilistic constructions of such expanders are mathematical folklore (see, e.g., [Lub94, Chapter 1.2], or [Vad12, Theorem 4.4] with a slightly different counting argument and [HLW06, Lemma 1.9] where certain expanders are called *magical*). We present one of them, following closely the proof of [HLW06, Lemma 1.9].

3.3 Proposition ([HLW06])**.** *Let $r \geq 3$, $\epsilon = 2/r$ and $\theta \in (0,1)$, $n \geq 2$ such that $10^r\theta < 1$, $\theta n \geq 1$. Then the probability that a uniformly chosen random left r-regular bipartite graph $G(A, B, E)$ with the left set size $\#A = n$ and the right set size $\#B = \lfloor n/2 \rfloor$ is* not *an (r, θ, ϵ) expander is less than $(10^{-r}\theta^{-1} - 1)^{-1}$.*

Proof. To build G, for each vertex from the left set A, independently, we assign r vertices from the right set B to be its neighbours.

Fix a nonempty subset S in A of cardinality $s \leq \theta n$ and then a subset T

in B of cardinality $t \leq (1-\epsilon)rs$. Define the random variable

$$X_{S,T} = \begin{cases} 1, & \text{if all the edges of } G \text{ from } S \text{ go to } T, \\ 0, & \text{otherwise.} \end{cases}$$

Notice that if $\sum X_{S,T} = 0$, where the sum is over all subsets S, T as above, then G is an (r, θ, ϵ) expander. Therefore,

$$p = \mathbb{P}\left(G \text{ is } \textit{not} \text{ an } (r,\theta,\epsilon) \text{ expander}\right) \leq \mathbb{P}\left(\sum X_{S,T} > 0\right).$$

Moreover, it is clear that $\mathbb{P}\left(X_{S,T} = 1\right) = \left[\binom{t}{r}/\binom{\#B}{r}\right]^s \leq [t/\#B]^{rs}$. By the union bound and then the standard estimate $\binom{n}{k} \leq (en/k)^k$ we obtain

$$p \leq \sum \mathbb{P}\left(X_{S,T} = 1\right) < \sum_{1 \leq s \leq \theta n} e^{s+(1-\epsilon)rs} r^{\epsilon rs} \left(\frac{n}{s}\right)^s \left(\frac{s}{\#B}\right)^{\epsilon rs}.$$

Since $\#B = \lfloor n/2 \rfloor > n/e$ for $n \geq 2$, we get

$$p < \sum_{1 \leq s \leq \theta n} e^{(1+r)s} r^{\epsilon rs} \left(\frac{s}{n}\right)^{(\epsilon r - 1)s}.$$

The choice $\epsilon r = 2$, along with the simple estimate $e^{1+r} r^2 < 10^r$, $r \geq 3$, gives

$$p < \sum_{1 \leq s \leq \theta n} [10^r \theta]^s < (10^{-r}\theta^{-1} - 1)^{-1}.$$

□

It was discovered by R. Berinde, R. Gilbert, P. Indyk, H. Karloff and M. Strauss that the adjacency matrices of bipartite unbalanced expanders (completed with zeros) are in the class $\mathcal{T}^n_{\epsilon,c}$. However, clearly they have kernels of dimension proportional to n because the size of the right vertex set B is a fraction of the size of the left vertex set A (see [BGI+08]). That is why, estimate (3.1) is of the the right order and, as we will see in the next section, the answer to Schechtman's question is negative.

Fix n, r, θ, ϵ as in Proposition 3.3 so that the probability appearing there is less than 1. Take, say $r = 8$, $\epsilon = 1/4$, $\theta = 10^{-9}$, $n \geq 10^9$. Let Φ be the adjacency matrix of an (r, θ, ϵ) expander $G(A, B, E)$ provided by that proposition (Φ is a $\#B \times \#A$ matrix with 1 at the entry (i, j) if $(i, j) \in E$, 0 otherwise).

In [BGI$^+$08], it is proved that the expanding property of G guarantees that Φ has good local invertibility properties in ℓ_1 (in fact the authors prove more, that the converse is true as well). For the reader's convenience, we now sketch the proof of their key observation.

3.4 Lemma ([BGI$^+$08])**.** *For every $x \in \mathbb{R}^n$ with $\#\mathrm{supp}(x) \leq \theta n$ we have*

$$\|\Phi x\| \geq (1-2\epsilon)r\|x\|,$$

where $\|\cdot\|$ denotes the ℓ_1 norm.

Proof. (Sketch) Without loss of generality let us assume that $|x_1| \geq \ldots \geq |x_n|$ and say $|x_{k+1}| = \ldots = |x_n| = 0$, $k = \#\mathrm{supp}(x)$. We order the edges of G by going over Φ *column by column, top to bottom* and setting $e_t = (i_t, j_t)$ if and only if $\Phi_{j_t,i_t} = 1$, $t = 1,\ldots,rn$ (so $e = (i,j)$ implicitly means that e is an edge from $i \in A$ to $j \in B$ and (j_t, i_t) are the entries in Φ having ones). We say that $e_t = (i_t, j_t)$ causes a *collision* if there is $s < t$ and an edge $e_s = (i_s, j_s)$ with $j_s = j_t$. Let E' be the set of edges which do *not* cause collisions and let $E'' = E \setminus E'$. The key observation is that

$$\sum_{(i,j)\in E''} |x_i| \leq \epsilon r\|x\|. \tag{3.2}$$

To see it, denote by l_i be the number of edges among $e_{(i-1)r+1}, e_{(i-1)r+2}, \ldots, e_{ir}$ which cause collisions, $i = 1,\ldots,n$. By the expanding property of G we know that

$$l_1 + \ldots l_i \leq \epsilon \cdot ir, \qquad i = 1,\ldots,k.$$

Applying summation by parts to $\sum_{(i,j)\in E''} |x_i| = l_1|x_1| + \ldots + l_k|x_k|$ proves (3.2).

To bound $\|\Phi x\|$ we notice that

$$\|\Phi x\| = \sum_j \left|\sum_i \Phi_{j,i} x_i\right| = \sum_{j\in B} \left|\sum_{i\in A:(i,j)\in E} x_i\right|,$$

break the sum $\sum_{i:(i,j)\in E} x_i$ into two bits $\sum_{i:(i,j)\in E'} x_i + \sum_{i:(i,j)\in E''} x_i$, and, after using the triangle inequality, we will get that the first bit gives the major contribution whereas the second one, by (3.2), is small. This will finish the proof of the lemma. $\square$

In order to construct a matrix invertible for short vector but with the kernel of dimension proportional to n, that is to see why the bound (3.1) is essentially sharp, we put

$$A_n = \left[\begin{array}{c} \frac{1}{r}\Phi \\ \hline 0 \end{array} \right]$$

which is an $n \times n$ block matrix — the $\lfloor n/2 \rfloor \times n$ matrix $\frac{1}{r}\Phi$ completed with $n - \lfloor n/2 \rfloor$ zero rows. Clearly, $\dim\ker A_n \geq n - \lfloor n/2 \rfloor \geq n/2$. Moreover, $\|A_n\|_{\ell_1^n \to \ell_1^n}$ is equal to the maximum of the ℓ_1 norms of the columns of A_n which is 1 (every column of Φ has exactly r nonzero entries which are 1). Finally, for a vector $x \in \mathbb{R}^n$ with $\#\mathrm{supp}(x) \leq \theta n$ we have $\|A_n x\| = \|\frac{1}{r}\Phi x\| \geq (1 - 2\epsilon)\|x\| = \frac{1}{2}\|x\|$ by Lemma 3.4. This shows the following result.

3.5 Theorem ([BGI$^+$08])**.** *For every $n \geq n_0 = 10^9$ there is an $n \times n$ matrix A_n with the properties*

1) $\|A_n\|_{\ell_1^n \to \ell_1^n} = 1$.

2) $\|A_n x\|_{\ell_1^n} \geq \frac{1}{2}\|x\|_{\ell_1^n}$ *for every vector $x \in \mathbb{R}^n$ with $\#\mathrm{supp}(x) \leq \theta n$, $\theta = 10^{-9}$.*

3) $\dim\ker A_n \geq \frac{n}{2}$.

3.4 Operators on L_1

Now we are ready to show how to construct examples of operators acting on L_1 and prove that the answer to Schechtman's question is negative.

3.6 Theorem. *Let θ, n_0 and A_n for $n \geq n_0$ be provided by Theorem 3.5. Let the space $\big(L_1([0,1]^n)\big)^n$ of the Cartesian product of n copies of $L_1([0,1]^n)$ be equipped with the norm*

$$\|(f_1, \ldots, f_n)\| = \frac{1}{n}\sum_{i=1}^{n} \|f_i\|_{L_1([0,1]^n)}.$$

Define an operator

$$T\colon L_1([0,1]) \longrightarrow \big(L_1([0,1]^n)\big)^n$$

by

$$(Tf)_i(x_1,\dots,x_n)=\sum_{j=1}^{n}A_n(i,j)f_j(x_j),\qquad i=1,\dots,n,$$

where $A_n(i,j)$ is the (i,j)-entry of A_n and for $j=1,\dots,n$

$$f_j(t)=f\big((j-1)/n+t/n\big),\qquad t\in[0,1].$$

Then for $n\geq\max\{n_0,3/\theta\}$ operator T possesses the following properties

(i) $\|T\|\leq 1$.

(ii) $\|Tf\|\geq\frac{1}{4}\|f\|$ *for every function $f\in L_1([0,1])$ with $|\mathrm{supp}(f)|\leq\theta/3$.*

(iii) If $n>2k$, then for every functions $g_1,\dots,g_k\in L_\infty([0,1])$ there is a nonzero function $f\in L_1([0,1])$ such that $\int_0^1 f\cdot g_j=0$ for $j=1,\dots,k$, but $Tf=0$.

3.7 Remark. The space $\big(L_1([0,1]^n)\big)^n$ is isometrically isomorphic to $L_1[0,1]$, so the above construction also yields an operator acting from $L_1[0,1]$ to $L_1[0,1]$ with the same properties (i) — (iii).

Proof. Fix a function $f\in L_1([0,1])$. For $x=(x_1,\dots,x_n)\in[0,1]^n$ we shall denote the vector $(f_1(x_1),\dots,f_n(x_n))$ by $F(x)$. Note that

$$\|f\|_{L_1([0,1])}=\frac{1}{n}\sum_{i=1}^{n}\|f_i\|_{L_1([0,1])}=\frac{1}{n}\int_{[0,1]^n}\|F(x)\|_{\ell_1^n}\mathrm{d}x.$$

Moreover,

$$\begin{aligned}\|Tf\|&=\frac{1}{n}\sum_{i=1}^{n}\|(Tf)_i\|_{L_1([0,1]^n)}=\frac{1}{n}\sum_{i=1}^{n}\int_{[0,1]^n}\left|\sum_{j=1}^{n}A_n(i,j)f_j(x_j)\right|\mathrm{d}x\\&=\frac{1}{n}\int_{[0,1]^n}\|A_nF(x)\|_{\ell_1^n}\mathrm{d}x.\end{aligned}$$

Property (i). Using $\|A_nF(x)\|_{\ell_1^n}\leq\|F(x)\|_{\ell_1^n}$ we get from the above formulae that $\|Tf\|\leq\|f\|$.

Property (ii). Let $S=\mathrm{supp}(f)$. We can assume that $|S|=\theta/3$ (if that is not the case, put ϵ in the places where f is zero). Observe that

$$\begin{aligned}S_i=\mathrm{supp}(f_i)&=\{t\in[0,1],\ (i-1)/n+t/n\in S\}=(nS-(i-1))\cap[0,1]\\&=\big(nS\cap[i-1,i]\big)-(i-1)\end{aligned}$$

and

$$\sum_{i=1}^{n} |S_i| = \sum_{i=1}^{n} \big|nS \cap [i-1,i]\big| = n|S| = n\theta/3.$$

Given a subset $I \subset \{1,\ldots,n\} = [n]$ by S^I we denote the set $B_1 \times \ldots \times B_n$ with $B_i = S_i$ for $i \in I$ and $B_i = S_i^c$ for $i \in [n] \setminus I$. Clearly, for $x \in S^I$ the support of the vector $F(x)$ has cardinality $\#I$, thus the idea is that by restricting the integration to the union of S^I with $\#I \leq \theta n$ we will be able to use the invertibility of A_n. We obtain

$$\begin{aligned}
\|Tf\| &\geq \frac{1}{n} \sum_{1\leq k\leq \theta n} \sum_{\#I=k} \int_{S^I} \|A_n F(x)\|_{\ell_1^n} \mathrm{d}x \\
&\geq \frac{1}{2n} \sum_{1\leq k\leq \theta n} \sum_{\#I=k} \int_{S^I} \|F(x)\|_{\ell_1^n} \mathrm{d}x \\
&= \frac{1}{2n} \sum_{i=1}^{n} \sum_{1\leq k\leq \theta n} \sum_{\#I=k} \int_{S^I} |f_i(x_i)| \mathrm{d}x.
\end{aligned}$$

For a fixed i the integral $\int_{S^I} |f_i(x_i)|\mathrm{d}x$ is 0 if $i \notin I$, otherwise it equals

$$\|f_i\| \left(\prod_{j\in I\setminus\{j\}} |S_j| \right) \left(\prod_{j\in[n]\setminus I} |S_j^c| \right).$$

Suppose for a moment that $i = n$. Then we can write

$$\begin{aligned}
\sum_{1\leq k\leq \theta n} \sum_{\#I=k} \int_{S^I} |f_i(x_i)| \mathrm{d}x &= \sum_{0\leq k\leq \theta n-1} \sum_{\substack{I\subset[n-1]\\ \#I=k}} \|f_n\| \cdot \prod_{j\in I} |S_j| \prod_{j\in[n-1]\setminus I} |S_j^c| \\
&= \|f_n\| \cdot \mathbb{P}\left(X_1 + \ldots + X_{n-1} \leq \theta n - 1\right),
\end{aligned}$$

where X_i are independent random variables with distribution $\mathbb{P}(X_i = 1) = |S_i| = 1 - \mathbb{P}(X_i = 0)$. Since for $n \geq 3/\theta$

$$2\mathbb{E}(X_1 + \ldots + X_{n-1}) \leq 2(|S_1| + \ldots + |S_n|) = 2n\theta/3 \leq n\theta - 1,$$

by Chebyshev's inequality we get

$$\begin{aligned}
\mathbb{P}\left(X_1 + \ldots + X_{n-1} > \theta n - 1\right) &\leq \mathbb{P}\left(X_1 + \ldots + X_{n-1} > 2\mathbb{E}(X_1 + \ldots + X_{n-1})\right) \\
&\leq 1/2,
\end{aligned}$$

hence

$$\sum_{1\le k\le \theta n}\sum_{\#I=k}\int_{S^I}\|f_n(x_n)\|_{\ell_1^n}\mathrm{d}x \ge \frac{1}{2}\|f_n\|.$$

Dealing with $i<n$ similarly we finally get

$$\|Tf\| \ge \frac{1}{2n}\sum_{i=1}^{n}\frac{1}{2}\|f_i\| = \frac{1}{4}\|f\|.$$

Property (iii). Let $V\subset\mathbb{R}^n$ be the space of the solutions of the system of equations

$$\left\{\sum_{i=1}^{n}\left(\int_{(i-1)/n}^{i/n} g_j(t)\mathrm{d}t\right)x_i = 0, \qquad j=1,\dots,k.\right.$$

There are n variables and k equations, thus $\dim V \ge n-k$. Therefore

$$\dim(V\cap\ker A_n) \ge n-k+n/2-n = n/2-k.$$

As a result, if $n>2k$, there is a nonzero vector $x\in V\cap\ker A_n$. Take

$$f(t)=\sum_{i=1}^{n}x_i\mathbf{1}_{[(i-1)/n,i/n)}(t), \qquad t\in[0,1].$$

□

3.5 Convolution operators

Let $\mathbb{T}=\mathbb{R}/\mathbb{Z}$ be the one dimensional torus viewed as a compact group with addition modulo 1 denoted by $x\oplus y$, equipped with the Haar measure (inherited from Lebesgue measure). To begin with, fix $1\le p\le\infty$ and consider the averaging operator U_t acting on $L_p(\mathbb{T})$ (with the usual norm $\|f\| = \left(\int_{\mathbb{T}}|f|^p\right)^{1/p}$ for $p<\infty$, and $\|f\| = \operatorname{ess\,sup}_{\mathbb{T}}|f|$ for $p=\infty$)

$$(U_tf)(x) = \frac{1}{2t}\int_{-t}^{+t} f(x\oplus s)\,\mathrm{d}s, \quad t\in(0,1). \tag{3.3}$$

If t is small, is the operator $I-U_t$ invertible, or, in other words, how much does U_tf differ from f? Of course, averaging a constant function does not change it, but excluding such a trivial case, we get a quantitative answer.

3.8 Theorem. *Let $t \in (0,1)$. There exists a positive constant c such that for every $1 \le p \le \infty$ and every $f \in L_p(\mathbb{T})$ with $\int_{\mathbb{T}} f = 0$ we have*

$$\|f - U_t f\| \ge ct^2 \|f\|, \tag{3.4}$$

where $\|\cdot\|$ denotes the L_p norm.

Note that if p was equal to 2, then, with the aid of Fourier analysis, the above estimate would be trivial. However to get an estimate for L_p that is independent of p is more subtle.

When $p = 1$, if we further estimate the left hand side of (3.4) using the Sobolev inequality, see [GT01], we obtain the following corollary.

3.9 Corollary. *There is a positive constant c such that for every function f from the Sobolev space $W^{1,1}(\mathbb{T})$ with $\int_{\mathbb{T}} f = 0$ and every $t \in (0,1)$ we have*

$$\int_{\mathbb{T}} \left| f'(x) - \frac{f(x \oplus t) - f(x \oplus -t)}{2t} \right| dx \ge ct^2 \int_{\mathbb{T}} |f(x)|\, dx, \tag{3.5}$$

3.10 Remark. Setting $t = 1/2$, inequality (3.5) becomes the usual Sobolev inequality, so (3.5) can be viewed as a certain generalization of the Sobolev inequality.

3.11 Remark. Set $f(x) = \cos(2\pi x)$. Then $\|f - U_t f\| = \|f\| \left(1 - \frac{1}{2\pi t}\sin(2\pi t)\right) \approx t^2 \|f\|$, for small t. Therefore, the inequality in Theorem 3.8 is sharp up to an absolute constant.

In this section we give a proof of a generalization of Theorem 3.8. We say that a $\mathbb{T}$-valued random variable Z is *c-good* with some positive constant c if $\mathbb{P}(Z \in A) \ge c|A|$ for all measurable $A \subset \mathbb{T}$. Equivalently, by Lebesgue's decomposition theorem it means that the absolutely continuous part of Z (with respect to the Lebesgue measure) has a density bounded below by a positive constant. We say that a real random variable Y is *ℓ-decent* if $Y_1 + \ldots + Y_\ell$ has a nontrivial absolutely continuous part, where $Y_1, Y_2, \ldots$ are i.i.d. copies of Y. Our main result reads

3.12 Theorem. *Given $t \in (0,1)$ and an ℓ-decent real random variable Y, consider the operator A_t given by*

$$(A_t f)(x) = \mathbb{E} f(x \oplus tY). \tag{3.6}$$

Then there exists a positive constant c which depends only on the distribution of the random variable Y such that for every $1 \le p \le \infty$ and every $f \in L_p(\mathbb{T})$ with $\int_{\mathbb{T}} f = 0$ we have

$$\|f - A_t f\| \ge ct^2 \|f\|,$$

where $\|\cdot\|$ denotes the L_p norm.

3.13 Remark. We cannot hope to prove a statement similar to Theorem 3.12 for purely atomic measures. Indeed, just consider the case $p = 1$ and let Y be distributed according to the law $\mu_Y = \sum_{i=1}^{\infty} p_i \delta_{x_i}$. Then for every $\epsilon > 0$ and every $t \in (0,1)$ there exists $f \in L_1(\mathbb{T})$ such that $\|f - A_t(f)\| < \epsilon$ and $\|f\| = 1$. To see this take N such that $\sum_{i=N+1}^{\infty} p_i < \epsilon/4$ and let $f_n(x) = \frac{\pi}{2}\sin(2\pi n x)$. Then $\|f_n\| = 1$. Let $n_0 \ge 8\pi/\epsilon$. Consider a sequence $\big((\pi n t x_1 \mod 2\pi, \dots, \pi n t x_N \mod 2\pi)\big)_n$ for $n = 0, 1, 2, \dots, n_0^N$ and observe that by the pigeonhole principle there exist $0 \le n_1 < n_2 \le n_0^N$ such that for all $1 \le i \le N$ we have $\operatorname{dist}(\pi t x_i(n_1 - n_2), 2\pi\mathbb{Z}) \le \frac{2\pi}{n_0}$. Taking $n = n_2 - n_1$ we obtain

$$\begin{aligned}
\|f_n - A_t(f_n)\| &\le \frac{\pi}{2}\sum_{i=1}^{N} p_i \|\sin(2\pi n x) - \sin(2\pi n(x + t x_i))\| + \frac{\epsilon}{2} \\
&= \pi \sum_{i=1}^{N} p_i |\sin(\pi n t x_i)| \cdot \|\cos(2\pi n x \oplus \pi n t x_i)\| + \frac{\epsilon}{2} \\
&\le 2\sum_{i=1}^{N} p_i |\sin(\pi n t x_i)| + \frac{\epsilon}{2} \le \frac{4\pi}{n_0}\sum_{i=1}^{N} p_i + \frac{\epsilon}{2} \le \epsilon.
\end{aligned}$$

For the proof of Theorem 3.12 we will need two lemmas. The first one shows why we bother about ℓ-decent and c-good random variables (the point being, of course, that we can apply a local version of the central limit theorem for ℓ-decent random variables). The second lemma explains why convolving with good random variables gives operators which are strong contractions.

3.14 Lemma. *Suppose Y is an ℓ-decent random variable. Let $Y_1, Y_2, \ldots$ be independent copies of Y. Then there exist a positive integer $N = N(Y)$ and numbers $c = c(Y) > 0$, $C_0 = C_0(Y) \geq 1$ such that for all $C \geq C_0$ and $n \geq N$ the random variable*

$$X_n^{(C)} = \left(C \cdot \frac{Y_1 + \ldots + Y_n}{\sqrt{n}} \right) \mod 1 \tag{3.7}$$

is c-good.

Proof. We prove the lemma in a few steps considering more and more general assumptions about Y.

Step I. Suppose that the characteristic function of Y belongs to $L_p(\mathbb{R})$ for some $p \geq 1$. In this case, by a certain version of the Local Central Limit Theorem, e.g. Theorem 19.1 in [BRR86], p. 189, we know that the density q_n of $(Y_1 + \ldots + Y_n - n\mathbb{E}Y)/\sqrt{n}$ exists for sufficiently large n, and satisfies

$$\sup_{x\in\mathbb{R}} \left| q_n(x) - \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \right| \xrightarrow[n\to\infty]{} 0, \tag{3.8}$$

where $\sigma^2 = \mathrm{Var}(Y)$. Observe that the density $g_n^{(C)}$ of $X_n^{(C)}$ equals

$$g_n^{(C)}(x) = \sum_{k\in\mathbb{Z}} \frac{1}{C} q_n\left(\frac{1}{C}(x+k) - \sqrt{n}\mathbb{E}Y \right), \qquad x \in [0,1].$$

Using (3.8), for $\delta = \frac{e^{-2/\sigma^2}}{\sqrt{2\pi}\sigma}$ we can find $N = N(Y)$ such that

$$q_n(x) > \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} - \delta/8, \qquad x \in \mathbb{R},\ n \geq N.$$

Therefore, to be close to the maximum of the Gaussian density we sum over only those k for which $x + k \in (-2C, 2C) + C\sqrt{n}\mathbb{E}Y$ for all $x \in [0,1]$. Since there are at least C and at most $4C$ such k, we get that

$$g_n^{(C)}(x) > \frac{1}{C}\frac{1}{\sqrt{2\pi}\sigma} e^{-2/\sigma^2} \cdot C - \frac{1}{C}\frac{\delta}{8} \cdot 4C = \frac{1}{2\sqrt{2\pi}\sigma} e^{-2/\sigma^2}.$$

In particular, this implies that $X_n^{(C)}$ is c-good with $c = \frac{1}{2\sqrt{2\pi}\sigma} e^{-2/\sigma^2}$. Thus, in this case, it suffices to set $C_0 = 1$.

Step II. Suppose that the law of Y is of the form $q\mu + (1-q)\nu$ for some $q \in (0,1]$ and some Borel probability measures μ, ν on $\mathbb{R}$ such that the characteristic function of μ belongs to $L_p(\mathbb{R})$ for some $p \geq 1$. Notice that

$$\mu_{Y_1+\ldots+Y_N} = \mu_Y^{\star N} = (q\mu + (1-q)\nu)^{\star N} = \sum_{k=0}^{N} \binom{N}{k} q^k (1-q)^{N-k} \mu^{\star k} \star \nu^{\star(N-k)}$$
$$\geq \sum_{k=N_0}^{N} \binom{N}{k} q^k (1-q)^{N-k} \mu^{\star k} \star \nu^{\star(N-k)} = c_{N,N_0} \left(\mu^{\star N_0} \star \rho_{N,N_0}\right),$$

where

$$\rho_{N,N_0} = \frac{1}{c_{N,N_0}} \sum_{k=N_0}^{N} \binom{N}{k} q^k (1-q)^{N-k} \mu^{\star k - N_0} \star \nu^{\star(N-k)}$$

is a probability measure, and

$$c_{N,N_0} = \sum_{k=N_0}^{N} \binom{N}{k} q^k (1-q)^{N-k}$$

is a normalisation constant. Choosing $N_0 = \lfloor qN - C_1\sqrt{q(1-q)N} \rfloor$ we can guarantee that $c_{N,N_0} \geq 1/2$ eventually, say for $N \geq \tilde{N}$. Denoting by $\bar{Y}$, Z the random variables with the law μ, ρ_{N,N_0} respectively and by $\bar{Y}_i$ i.i.d. copies of $\bar{Y}$, we get

$$\mathbb{P}\left(X_N^{(C)} \in A\right) \geq c_{N,N_0} \mathbb{P}\left(\left(C\frac{\bar{Y}_1 + \ldots + \bar{Y}_{N_0}}{\sqrt{N}} + C\frac{Z_{N,N_0}}{\sqrt{N}}\right) \mod 1 \in A\right).$$

By Step I, the first bit $C(\bar{Y}_1 + \ldots + \bar{Y}_{N_0})/\sqrt{N}$ is c-good for some $c > 0$ and $C \geq C_0^{(II)} = \sup_{N \geq \tilde{N}} \sqrt{N/N_0}$. Moreover, note that if U is a c-good $\mathbb{T}$-valued r.v., then so is $U \oplus V$ for every $\mathbb{T}$-valued r.v. V which is independent of U. As a result, $X_N^{(C)}$ is $c/2$-good.

Step III. Now we consider the general case, i.e. Y is ℓ-decent for some $\ell \geq 1$. For $n \geq \ell$ we can write

$$C \cdot \frac{Y_1 + \ldots + Y_n}{\sqrt{n}} = C\sqrt{\frac{\lfloor n/\ell \rfloor}{n}} \cdot \frac{\tilde{Y}_1 + \ldots + \tilde{Y}_{\lfloor n/\ell \rfloor}}{\sqrt{\lfloor n/\ell \rfloor}} + C\frac{\tilde{R}}{\sqrt{n}}$$

with $\tilde{Y}_j = Y_{(j-1)\ell+1} + \ldots + Y_{j\ell}$ for $j = 1, \ldots, \lfloor n/\ell \rfloor$, and $\tilde{R} = Y_{\lfloor n/\ell \rfloor \ell+1} + \ldots + Y_n$. Since the absolutely continuous part of the law μ of $\tilde{Y}_j$ is nontrivial, then μ is of the form $q\nu_1 + (1-q)\nu_2$ with $q \in (0,1]$ and the characteristic function of ν_1 belonging to some L_p. Indeed, μ has a bit which is a uniform distribution on some measurable set whose characteristic function is in L_2. Therefore, applying Step II for $\tilde{Y}_j$ we get that $X_n^{(C)}$ is c-good when $C\sqrt{\frac{\lfloor n/\ell \rfloor}{n}} \geq C_0^{(II)}$. So we can set $C_0 = C_0^{(II)}\sqrt{2\ell}$. $\square$

3.15 Lemma. *Suppose Z is a $\mathbb{T}$-valued c-good random variable and B_Z is the operator defined by $(B_Z f)(x) = \mathbb{E} f(x \oplus Z)$. Then for every $1 \leq p \leq \infty$ and every $f \in L_p(\mathbb{T})$ with $\int_{\mathbb{T}} f = 0$ we have $\|B_Z f\| \leq (1-c)\|f\|$, where $\|\cdot\|$ is the L_p norm.*

Proof. Fix $1 \leq p < \infty$. Let μ be the law of Z. Define the measure $\nu(A) = (\mu(A) - c|A|)/(1-c)$ for measurable $A \subset \mathbb{T}$. Since μ is c-good, ν is a Borel probability measure on $\mathbb{T}$. Take $f \in L_p(\mathbb{T})$ with mean zero. Then by Jensen's inequality we have

$$\begin{aligned}
\|B_Z f\|^p &= \int_0^1 \left| \int_0^1 f(x \oplus s)\, \mathrm{d}\mu(s) \right|^p \mathrm{d}x \\
&= (1-c)^p \int_0^1 \left| \int_0^1 f(x \oplus s)\, \mathrm{d}\nu(s) \right|^p \mathrm{d}x \\
&\leq (1-c)^p \int_0^1 \int_0^1 |f(x \oplus s)|^p\, \mathrm{d}\nu(s)\, \mathrm{d}x \\
&= (1-c)^p \|f\|^p \int_0^1 \mathrm{d}\nu(s) = (1-c)^p \|f\|^p .
\end{aligned}$$

Since c does not depend on p we get the same inequality for $p = \infty$ by passing to the limit. $\square$

Now we are ready to give the proof of Theorem 3.12.

Proof of Theorem 3.12. Fix $1 \leq p \leq \infty$. Let $Y_1, Y_2, \ldots$ be independent copies of Y. Observe that

$$\begin{aligned}
(A_t^n f)(x) &= \mathbb{E} f\left(x \oplus tY_1 \oplus \ldots \oplus tY_n\right) \\
&= \mathbb{E} f\left(x \oplus \left(t\sqrt{n}\left(\frac{Y_1 + \ldots + Y_n}{\sqrt{n}}\right) \bmod 1\right)\right).
\end{aligned}$$

Take $n(t) = C_0^2 \lceil 1/t^2 \rceil N$, where C_0 and N are the numbers given by Lemma 3.14. Therefore, with $X_{n(t)}^{(C)}$ defined by (3.7), we can write

$$(A_t^{n(t)} f)(x) = \mathbb{E} f\left(x \oplus X_{n(t)}^{(C)}\right),$$

where $C = t\sqrt{n(t)} = tC_0\sqrt{\lceil 1/t^2 \rceil N} \geq C_0\sqrt{N} \geq C_0$. Thus $X^{(C)}_{n(t)}$ is $c(Y)$-good with some constant $c(Y) \in (0,1)$. From Lemma 3.15 we have

$$\left\| A_t^{n(t)} f \right\| \leq (1 - c(Y)) \|f\|$$

for all f satisfying $\int_{\mathbb{T}} f = 0$.

The operator A_t is a contraction, namely $\|A_t f\| \leq \|f\|$ for all $f \in L_1(\mathbb{T})$. Using this observation and the triangle inequality we obtain

$$\begin{aligned} \|f - A_t f\| &\geq \frac{1}{n} \left(\|f - A_t f\| + \left\| A_t f - A_t^2 f \right\| + \ldots + \left\| A_t^{n-1} f - A_t^n f \right\| \right) \\ &\geq \frac{1}{n} \|f - A_t^n f\| . \end{aligned}$$

Taking $n = n(t)$ we arrive at

$$\frac{1}{n(t)} \left\| f - A_t^{n(t)} f \right\| \geq \frac{1}{t^{-2}+1} \cdot \frac{1}{C_0^2 \cdot N} \left(\|f\| - \left\| A_t^{n(t)} f \right\| \right) \geq \frac{c(Y)}{2C_0^2 \cdot N} t^2 \|f\| .$$

To finish the proof, it suffices to take $c = c(Y)/(2C_0^2 \cdot N)$. □

3.16 Remark. Consider an ℓ-decent random variable Y. As was noted in the proof of Lemma 3.14 (Step III), the law $Y_1 + \ldots + Y_\ell$ has a bit whose characteristic function is in L_2. Conversely, if the law of $S_m = Y_1 + \ldots + Y_m$ has the form $q\mu + (1-q)\nu$ with $q \in (0,1]$ and the characteristic function of μ belongs to L_p for some $p \geq 1$, then the characteristic function of the bit $\mu^{\star \lceil p/2 \rceil}$ of the sum of $\lceil p/2 \rceil$ i.i.d. copies of S_m is in L_2. In particular, that bit has a density function in $L_1 \cap L_2$. Thus Y is $(m \lceil p/2 \rceil)$-decent.

3.17 Remark. The idea to study the operators A_t (see (3.6)) stemmed from Schechtman's question, Problem 3.1, presented and discussed earlier in the chapter. Our hope was that an operator $T = I - A_t$, for some Y, would provide a negative answer to Schechtman's question. However, Theorem 3.12 says that if Y is an ℓ-decent random variable, then T is nicely invertible on the subspace of functions $f \in L_1$ such that $\int f \cdot 1 = 0$.

3.6 Notes and comments

The question mentioned at the start of the chapter asked by A. Nasseri was this (see [Nas12]): *does there exist a nonsurjective bounded linear operator on*

ℓ_∞ *with dense range?* Its main difficulty is captured in Schechtman's question, Problem 3.1.

Using more technical arguments involving tools such as ultraproducts, Theorem 3.6 can be *souped up* to provide a positive answer to Nasseri's question as well as some other related questions. This is done by the author's collaborators in the paper [JNST14]. What is presented in Sections 3.1-3.4 reflects the author's contribution to [JNST14].

Section 3.5 is based on the publication [NT14a] joint with P. Nayar. He and the author equally contributed to the results obtained therein as they worked together. It should be remarked that the present general statement of Theorem 3.12 was obtained thanks to useful comments of K. Oleszkiewicz to whom we are indebted. The remarks about purely atomic measures and the sharpness of the constant in (3.4) were pointed out by S. Kwapień.

Bibliography

[AGZ10] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[BAB13] G. Ben Arous and P. Bourgade. Extreme gaps between eigenvalues of random matrices. *Ann. Probab.*, 41(4):2648–2681, 2013.

[BGI+08] R. Berinde, A.C. Gilbert, P. Indyk, H. Karloff, and M.J. Strauss. Combining geometry and combinatorics: A unified approach to sparse signal recovery. In *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on Communication, Control, and Computing*, pages 798–805, Sept 2008.

[BP02] H.-P. Breuer and F. Petruccione. *The theory of open quantum systems*. Oxford University Press, New York, 2002.

[BRR86] R. N. Bhattacharya and R. Ranga Rao. *Normal approximation and asymptotic expansions*. Robert E. Krieger Publishing Co., Inc., Melbourne, FL, 1986. Reprint of the 1976 original.

[CD10] I. Camilier and L. Decreusefond. Quasi-invariance and integration by parts for determinantal and permanental processes. *J. Funct. Anal.*, 259(1):268–300, 2010.

[GT01] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[HKPV09] J. B. Hough, M. Krishnapur, Y. Peres, and B. Virág. *Zeros of Gaussian analytic functions and determinantal point processes*, volume 51 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2009.

[HLW06] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S.)*, 43(4):439–561 (electronic), 2006.

[JNST14] W. Johnson, A. Nasseri, G. Schechtman, and T. Tkocz. Injective Tauberian operators on L_1 and operators with dense range on ℓ_∞. *Canad. Math. Bull.*, 2014. to appear.

[KS93] S. Kwapień and J. Sawa. On some conjecture concerning Gaussian measures of dilatations of convex symmetric sets. *Studia Math.*, 105(2):173–187, 1993.

[Lat02] R. Latała. On some inequalities for Gaussian measures. In *Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002)*, pages 813–822. Higher Ed. Press, Beijing, 2002.

[Led01] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

[LO99] R. Latała and K. Oleszkiewicz. Gaussian measures of dilatations of convex symmetric sets. *Ann. Probab.*, 27(4):1922–1938, 1999.

[LO00] R. Latała and K. Oleszkiewicz. Between Sobolev and Poincaré. In *Geometric aspects of functional analysis*, volume 1745 of *Lecture Notes in Math.*, pages 147–168. Springer, Berlin, 2000.

[Lub94] A. Lubotzky. *Discrete groups, expanding graphs and invariant measures*, volume 125 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 1994. With an appendix by Jonathan D. Rogawski.

[Nas12] A. Nasseri. Surjectivity of operators on ℓ^∞, 2012. Mathoverflow. `http://mathoverflow.net/questions/101253`.

[NT13] P. Nayar and T. Tkocz. The unconditional case of the complex S-inequality. *Israel J. Math.*, 197(1):99–106, 2013.

[NT14a] P. Nayar and T. Tkocz. A note on certain convolution operators. In Bo'az Klartag and Emanuel Milman, editors, *Geometric Aspects of Functional Analysis*, volume 2116 of *Lecture Notes in Mathematics*, pages 405–412. Springer International Publishing, 2014.

[NT14b] P. Nayar and T. Tkocz. S-inequality for certain product measures. *Math. Nachr.*, 287(4):398–404, 2014.

[Pin73] M. S. Pinsker. On the complexity of a concentrator. In *7th Annual Teletraffic Conference*, pages 318/1–318/4, 1973.

[Pis89] G. Pisier. *The volume of convex bodies and Banach space geometry*, volume 94 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1989.

[Sha92] B. V. Shabat. *Introduction to complex analysis. Part II*, volume 110 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1992. Functions of several variables, Translated from the third (1985) Russian edition by J. S. Joel.

[STKZ13] M. Smaczyński, T. Tkocz, M. Kuś, and K. Życzkowski. Extremal spacings between eigenphases of random unitary matrices and their tensor products. *Phys. Rev. E*, 88:052902, Nov 2013.

[Sza91] S. J. Szarek. Condition numbers of random matrices. *J. Complexity*, 7(2):131–149, 1991.

[Tko11] T. Tkocz. Gaussian measures of dilations of convex rotationally symmetric sets in $\mathbb{C}^N$. *Electron. Commun. Probab.*, 16:38–49, 2011.

[Tko13] T. Tkocz. A note on the tensor product of two random unitary matrices. *Electron. Commun. Probab.*, 18:no. 16, 7, 2013.

[TSK+12] T. Tkocz, M. Smaczyński, M. Kuś, O. Zeitouni, and K. Życzkowski. Tensor products of random unitary matrices. *Random Matrices Theory Appl.*, 1(4):1250009, 26, 2012.

[Vad12] S. P. Vadhan. Pseudorandomness. *Foundations and Trends® in Theoretical Computer Science*, 7(1–3):1–336, 2012.

[Wig55] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math. (2)*, 62:548–564, 1955.

[Wol07] P. Wolff. Some remarks on functionals with the tensorization property. *Bull. Pol. Acad. Sci. Math.*, 55(3):279–291, 2007.

[ZS74] V. A. Zalgaller and V. N. Sudakov. Some problems on centrally symmetric convex bodies. *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 45:75–82, 119, 1974. Problems in global geometry.