SECOND ORDER CONCENTRATION ON THE SPHERE

S. G. BOBKOV[1,4], G. P. CHISTYAKOV[2], AND F. GÖTZE[3]

Abstract. Sharpened forms of the concentration of measure phenomenon for classes of functions on the sphere are developed in terms of Hessians of these functions.

1. Introduction

Let σ_{n-1} denote the normalized Lebesgue measure on the unit sphere

$$S^{n-1} = \{x \in \mathbf{R}^n : |x| = 1\}, \qquad n \geq 2,$$

in the Euclidean n-space which is equipped with the canonical inner product $\langle \cdot, \cdot \rangle$ and the norm $|\cdot|$. The spherical concentration phenomenon asserts in particular that mean zero smooth functions f on S^{n-1} are of order at most $\frac{1}{\sqrt{n}}$ on a large part of the sphere in the sense of σ_{n-1}. This follows already from the Poincaré inequality

$$\int f^2 \, d\sigma_{n-1} \leq \frac{1}{n-1} \int |\nabla_S f|^2 \, d\sigma_{n-1}, \tag{1.1}$$

where $\nabla_S f$ stands for the spherical gradient of f (cf. e.g. [L1]). Hence, if the integral on the right-hand side is of order 1, the L^2-norm of f will be of order at most $\frac{1}{\sqrt{n}}$. Moreover, in case $|\nabla_S f| \leq 1$, there is a considerably stronger property

$$\int e^{(n-1)\, f^2/c} \, d\sigma_{n-1} \leq 2$$

involving some absolute constant $c > 0$. Using a standard normal random variable Z, it may be stated informally as stochastic dominance

$$|f| \preceq c\, \frac{|Z|}{\sqrt{n}}, \tag{1.2}$$

which means a corresponding inequality for the measures/probabilities of the tail sets $|f| \geq r$ and $\frac{c}{\sqrt{n}}\, |Z| \geq r$ for all $r > 0$. This property was first emphasized in the early 70's by V. D. Milman in the context of the local theory of Banach spaces and led him to the

1991 *Mathematics Subject Classification.* Primary 60E.

Key words and phrases. Concentration of measure phenomenon, logarithmic Sobolev inequalities.

1) School of Mathematics, University of Minnesota, USA; Email: bobkov@math.umn.edu.

2) Faculty of Mathematics, University of Bielefeld, Germany; Email: chistyak@math.uni-bielefeld.de.

3) Faculty of Mathematics, University of Bielefeld, Germany; Email: goetze@math.uni-bielefeld.de.

4) Research partially supported by Alexander von Humboldt Foundation.

understanding of the concentration of measure phenomenon in a much broader sense; cf. V. D. Milman, G. Schechtman [M-S], subsequent works by M. Talagrand [T1-2], and the book by M. Ledoux [L2] for an account of basic ideas and results in this direction up to the end of 90's.

Returning to the sphere, in certain problems one deals however with smooth functions that turn out to be of a much smaller order than $\frac{1}{\sqrt{n}}$. This cannot be guaranteed just by the Lipschitz condition $|\nabla_S f| \leq 1$, even if f is orthogonal to linear functions in $L^2(S^{n-1}, \sigma_{n-1})$ (which play an extremal role in (1.1)). Hence, conditions on higher derivatives of f are required. The aim of this note is to study corresponding conditions in terms of the Hessian of f''_S of f by involving both the operator norms $\|f''_S(\theta)\|$ and the Hilbert-Schmidt norms $\|f''_S(\theta)\|_{\rm HS}$ of the matrices $f''_S(\theta)$ $(\theta \in S^{n-1})$.

Orthogonality of functions on the unit sphere will be understood as orthogonality in the Hilbert space $L^2(S^{n-1}, \sigma_{n-1})$. Restrictions of affine, linear and quadratic functions on $\mathbf{R}^n$ to the sphere S^{n-1} will be again called affine, linear and quadratic functions respectively on the sphere.

Theorem 1.1. *Assume that f is a C^2-smooth function on S^{n-1} which is orthogonal to all affine functions. If $\|f''_S\| \leq 1$ at all points on the sphere and $\int \|f''_S\|_{\rm HS}^2\, d\sigma_{n-1} \leq b^2$, then*

$$\int \exp\Big\{\frac{n-1}{2(1+b^2)}\,|f|\Big\}\, d\sigma_{n-1} \leq 2. \tag{1.3}$$

By Chebyshev's inequality, (1.3) provides bounds on tails, which may be written similarly to (1.2) as

$$|f| \,\preceq\, c_b\Big(\frac{Z}{\sqrt{n}}\Big)^2,$$

however – with the right-hand side behaving like $\frac{1}{n}$ with respect to the dimension (provided that b is of order 1).

We refer to Theorem 1.1 as (a variant of) the second order concentration on the sphere. It is consistent with a second order Poincaré-type inequality

$$\int f^2\, d\sigma_{n-1} \leq \frac{1}{2n(n+2)} \int \|f''_S\|_{\rm HS}^2\, d\sigma_{n-1},$$

valid for all smooth f on S^{n-1} that are orthogonal to affine functions (with equality attainable for all quadratic spherical harmonics). This inequality can be derived using the spectral decomposition of f in spherical harmonics by means of the identity

$$\int \|f''_S\|_{\rm HS}^2\, d\sigma_{n-1} = \int f\,\big(\Delta_S(\Delta_S f) + (n-2)\Delta_S f\big)\, d\sigma_{n-1}. \tag{1.4}$$

Here and in the sequel $\Delta_S = {\rm Tr}\, f''_S$ denotes the Laplacian operator on S^{n-1} which acts diagonally on all homogeneous spherical harmonics. Although typically $\Delta_S f$ behaves in a more "chaotic" (oscillatory) way than f, the average in (1.4) captures and cancels such potentially large oscillations.

The conditions on the spherical second derivative in Theorem 1.1 are fulfilled, for example, when $\|f_S''\|_{\mathrm{HS}} \leq b$ on S^{n-1}. However, in applications, one might prefer to deal with functions on the sphere induced by smooth functions in $\mathbf{R}^n$ or at least in a neighbourhood of the sphere via restriction and using the Euclidean derivatives of such functions, rather than intrinsic derivatives on S^{n-1}. Using this Euclidean setup, we may formulate a related statement as follows.

In the sequel we denote by $f''(x) = \big(\partial_{ij} f(x)\big)_{i,j=1}^n$ the matrix of partial derivatives of f of second order at the point x, and by I_n the identity $n \times n$ matrix.

Theorem 1.2. *Let f be defined and C^2-smooth in some open neighbourhood of S^{n-1}. Assume that it is orthogonal to all affine functions and satisfies $\|f'' - aI_n\| \leq 1$ on S^{n-1} together with*

$$\int \|f'' - aI_n\|_{\mathrm{HS}}^2 \, d\sigma_{n-1} \leq b^2, \tag{1.5}$$

for some $a \in \mathbf{R}$ and $b \geq 0$. Then

$$\int \exp\Big\{\frac{n-1}{2(1+4b^2)}\,|f|\Big\}\, d\sigma_{n-1} \leq 2. \tag{1.6}$$

In Theorems 1.1-1.2 one may also start with an arbitrary C^2-smooth function f, but apply the hypotheses and the conclusions (1.3)/(1.6) to the projection Tf of f onto the orthogonal complement of the space of all affine functions on the sphere in $L^2(S^{n-1}, \sigma_{n-1})$. The "affine" part of f may be described as $l(\theta) = m + \langle v, \theta\rangle$ with

$$m = \int f(x)\, d\sigma_{n-1}(x), \quad v = n \int x f(x)\, d\sigma_{n-1}(x),$$

so $Tf(\theta) = f(\theta) - l(\theta)$. For example, if f is even, i.e. $f(-\theta) = f(\theta)$ for all $\theta \in S^{n-1}$, then $Tf = f - m$.

In the setting of Theorem 1.2, the functions Tf and f have identical Euclidean second derivatives. Hence, if we want to obtain an inequality similar to (1.6) without the orthogonality assumption (still assuming conditions on the Euclidean second derivative), we need to verify that the affine part l is of order $\frac{1}{n}$. This may be achieved by estimating the L^2-norm of l and using the well-known fact that the linear functions on the sphere behave like Gaussian random variables. If, for definiteness, f has mean zero, then

$$\|l\|_{L^2}^2 = \frac{1}{n}\,|v|^2 = nI, \quad \text{where} \ \ I = \iint \langle x, y\rangle\, f(x) f(y)\, d\sigma_{n-1}(x) d\sigma_{n-1}(y).$$

Therefore, a natural requirement would be a bound $I \leq \frac{b_0}{n^3}$ with b_0 of order 1. This leads to a variant of Theorem 1.2 which is more flexible in applications.

Theorem 1.3. *Let f be defined and C^2-smooth in some open neighbourhood of S^{n-1}. Assume that it has mean zero and*

$$\int\!\!\int \langle x,y\rangle\ f(x)f(y)\,d\sigma_{n-1}(x)d\sigma_{n-1}(y) \leq \frac{b_0}{n^3}, \qquad b_0 \geq 0.$$

If $\|f'' - aI_n\| \leq 1$ holds on S^{n-1} together with (1.5), *then*

$$\int \exp\Big\{\frac{n-1}{4(1+b_0^2+4b^2)}\,|f|\Big\}\,d\sigma_{n-1} \leq 2.$$

We believe that the second order concentration on the sphere may indeed be useful in various applications. One motivating example has been the question of optimal rates of approximation in the central limit theorem for linear forms $X_\theta = \langle X,\theta\rangle$, where $X = (X_1,\dots,X_n)$ is a given random vector in $\mathbf{R}^n$ whose components are not necessarily independent. If the covariance matrix of X has a bounded spectral radius, a celebrated result of Sudakov [S] indicates that, for n large, the distributions F_θ of X_θ are concentrated for most of θ (in the sense of σ_{n-1}) around a certain typical measure F on the real line, which may or may not be Gaussian. Many authors studied various aspects of this interesting phenomenon, and we omit references. Let us mention only that one can study the deviations F_θ from F in terms of the Fourier-Stieltjes transforms

$$f_t(\theta) = \mathbf{E}\,e^{it\langle\theta,X\rangle} = \int_{-\infty}^{\infty} e^{it\langle\theta,x\rangle}\,dF_\theta(x) \qquad (t\in\mathbf{R},\ \theta\in\mathbf{R}^n),$$

which are naturally defined as smooth functions on the whole space $\mathbf{R}^n$. By the direct differentiation in θ,

$$\langle f_t''(\theta)v,w\rangle = -t^2\,\mathbf{E}\,\langle v,X\rangle\,\langle w,X\rangle\,e^{it\langle\theta,X\rangle}.$$

Here, condition (1.5) leads to a certain correlation-type condition for products X_jX_k, such that (1.6) will ensure $\frac{1}{n}$-bounds for typical deviations of F_θ from F (in contrast with $\frac{1}{\sqrt{n}}$-bounds in the classical Berry-Esseen theorem). Such improving effects have recently been shown in the work of B. Klartag and S. Sodin in case of independent summands ([K-S], cf. also [K]). As for the general setting, this concentration problem will be dealt with in a separate paper and hence will not be discuss it here further.

The proof of Theorems 1.1-1.2 is based on the application of the logarithmic Sobolev inequality on the sphere and requires derivation of bounds on the integrals

$$\int |\nabla_S f|^2\,d\sigma_{n-1}, \qquad \int |\nabla f|^2\,d\sigma_{n-1} \tag{1.7}$$

in terms of the second derivatives. Basic tools leading to exponential bounds under logarithmic Sobolev inequalities are rather universal and can be developed in the setting of abstract metric spaces, cf. Section 2. Then we turn to the case of the sphere and sharpen the Poincaré inequality by involving the norm $\|f_S''\|$ (Section 3). Sections 5-6 are devoted to the estimation of the integrals (1.7). As a preliminary step, the identity

(1.4) is derived separately in Section 4. The proofs of Theorem 1.1 and Theorems 1.2-1.3 are completed in Sections 5 and 7, respectively. After Section 7 we add an Appendix (Sections 8-14) providing for the readers convenience more details on the underlying computations in spherical calculus.

Acknowledgement. We would like to thank Michel Ledoux for valuable comments concerning the differential geometric motivation of Proposition 4.1.

2. **Logarithmic Sobolev Inequalities on Metric Spaces**

Assume that a metric space (M, ρ) is equipped with a Borel probability measure μ. The triple (M, ρ, μ) is said to satisfy a logarithmic Sobolev inequality with constant $\sigma^2 < \infty$, if

$$\mathrm{Ent}_\mu(f^2) \leq 2\sigma^2 \int |\nabla f|^2 \, d\mu \tag{2.1}$$

for any bounded function f on M with finite Lipschitz semi-norm $\|f\|_{\mathrm{Lip}}$. The optimal value of σ^2 is then called the logarithmic Sobolev constant.

Here

$$\mathrm{Ent}_\mu(u) = \int u \log u \, d\mu - \int u \, d\mu \log \int u \, d\mu \qquad (u \geq 0)$$

is the entropy functional defined for non-negative measurable functions on M. As for the modulus of the gradient in (2.1), it may be understood in the generalized sense as

$$|\nabla f(x)| = \limsup_{y \to x} \frac{|f(x) - f(y)|}{\rho(x, y)} \qquad (x \in M). \tag{2.2}$$

This function is always Borel measurable, whenever f is continuous. In this abstract setting, (2.1) actually extends to the larger class of all f that have a finite Lipschitz semi-norm on every ball in M; such functions will be called locally Lipschitz.

Now, define the function

$$|\nabla^2 f(x)| = |\nabla \, |\nabla f(x)| \, | = \limsup_{y \to x} \frac{| \, |\nabla f(x)| - |\nabla f(y)| \, |}{\rho(x, y)}, \tag{2.3}$$

which we call a second order modulus of the gradients of f.

The Lipschitz property $\|f\|_{\mathrm{Lip}} \leq 1$ implies that $|\nabla f(x)| \leq 1$ for all $x \in M$. The converse is also true, at least when M is a (connected) Riemannian manifold. In this case, the assumption $|\nabla^2 f(x)| \leq 1$ for every x in M means that the function $|\nabla f|$ is Lipschitz. If $|\nabla f|$ is locally Lipschitz, then f is of course locally Lipschitz as well.

The next statement indicates how the definition (2.3) could be used in applications.

Proposition 2.1. *Assume that a metric probability space (M, ρ, μ) satisfies a logarithmic Sobolev inequality with constant σ^2. Then, for any locally Lipschitz function*

f on M with μ-mean zero, such that $|\nabla f|$ is locally Lipschitz and $|\nabla^2 f| \leq 1$ on the support of μ, we have

$$\int \exp\Big\{\frac{1}{2\sigma^2}\, f\Big\}\, d\mu \leq \exp\Big\{\frac{1}{2\sigma^2} \int |\nabla f|^2\, d\mu\Big\}. \tag{2.4}$$

Proof. The argument is based on two general results that relate (2.1) to the exponential integrability of Lipschitz functions. Namely, for any locally Lipschitz μ-integrable function u on M,

$$\int e^{u - \int u\, d\mu}\, d\mu \leq \int e^{\sigma^2 |\nabla u|^2}\, d\mu. \tag{2.5}$$

In addition, if $|\nabla u| \leq 1$ on the support of μ, say M_1, then for all $0 \leq t < \frac{1}{2\sigma^2}$,

$$\int e^{tu^2}\, d\mu \leq \exp\Big\{\frac{t}{1 - 2\sigma^2 t} \int u^2\, d\mu\Big\}. \tag{2.6}$$

On the basis of (2.1), the inequality (2.5) was derived in [B-G], cf. also [L1-2]. The second inequality, (2.6), is a classical result of Aida, Masuda and Shigekawa [A-M-S]. We refer to [B-G] for a detailed discussion.

We apply (2.6) with $t = \sigma^2\lambda^2$ to the locally Lipschitz function $u = |\nabla f|$. Since the condition $|\nabla u| \leq 1$ is assumed to hold on M_1, we get that

$$\int e^{\sigma^2\lambda^2 |\nabla f|^2}\, d\mu \leq \exp\Big\{\frac{\sigma^2\lambda^2}{1 - 2\sigma^4\lambda^2} \int |\nabla f|^2\, d\mu\Big\}, \qquad \lambda^2 < \frac{1}{2\sigma^4}.$$

On the other hand, since f is locally Lipschitz and has μ-mean zero, one may apply (2.5), which gives

$$\int e^{\lambda f}\, d\mu \leq \int e^{\sigma^2\lambda^2 |\nabla f|^2}\, d\mu.$$

Hence, the combination of these two bounds yields

$$\int e^{\lambda f}\, d\mu \leq \exp\Big\{\frac{\sigma^2\lambda^2}{1 - 2\sigma^4\lambda^2} \int |\nabla f|^2\, d\mu\Big\}.$$

Here one may choose $\lambda = \frac{1}{2\sigma^2}$, and then we arrive at the required inequality (2.4). $\square$

When M is an open region in $\mathbf{R}^n$ (with the Euclidean distance), the definition (2.1) leads to the usual notion of a logarithmic Sobolev inequality, holding for all locally Lipschitz functions on M. To avoid possible confusion about being locally Lipschitz, let us emphasize that, when f is differentiable at a given point x, (2.2) does coincide with the modulus (the length) of the Euclidean gradient. The same remark applies to the sphere $M = S^{n-1}$ with the geodesic or induced Euclidean distances, in which case (2.2) defines $|\nabla_S f(x)|$, the length of the spherical gradient of f.

The second order modulus of the gradients may also be related to the usual (Euclidean) derivatives. Namely, if f is C^2-smooth in the open set M in $\mathbf{R}^n$, the function $|\nabla f|$ will be locally Lipschitz, and

$$|\nabla^2 f(x)| = |\nabla f(x)|^{-1} |f''(x) \nabla f(x)|, \qquad x \in M. \tag{2.7}$$

Here the ratio should be understood as $\|f''(x)\|$ in case $|\nabla f(x)| = 0$. In particular,

$$|\nabla^2 f(x)| \leq \|f''(x)\|. \tag{2.8}$$

For example, for the quadratic function $f(x) = \frac{1}{2} \sum_{i=1}^n \lambda_i x_i^2$, $x = (x_1, \dots, x_n)$,

$$|\nabla^2 f(x)| \, = \, \frac{\sqrt{\sum_{i=1}^n \lambda_i^4 x_i^2}}{\sqrt{\sum_{i=1}^n \lambda_i^2 x_i^2}} \, \leq \, \max_i |\lambda_i|.$$

The identity (2.7) is easily obtained by the direct differentiation. Thus, in the Euclidean setup Proposition 2.1 may be simplified by using the inequality (2.8) as follows.

Corollary 2.2. *Let a probability measure μ on $\mathbf{R}^n$ satisfy a logarithmic Sobolev inequality with constant σ^2, and let a function f be C^2-smooth in an open neighbourhood of the support of μ. If it has μ-mean zero and $\|f''\| \leq 1$ on the support of μ, then*

$$\int \exp\Big\{\frac{1}{2\sigma^2}\, f\Big\}\, d\mu \leq \exp\Big\{\frac{1}{2\sigma^2} \int |\nabla f|^2 \, d\mu\Big\}.$$

3. **Logarithmic Sobolev Inequality on the Sphere**

An important result due to Mueller and Weissler [M-W] sharpens the Poincaré inequality (1.1) in terms of the logarithmic Sobolev inequality. Namely, the logarithmic Sobolev constant of the unit sphere S^{n-1}, which is equipped with the geodesic metric ρ and the uniform measure σ_{n-1}, coincides with the Poincaré constant $\sigma^2 = \frac{1}{n-1}$. That is, for any C^1-smooth function $f : S^{n-1} \rightarrow \mathbf{R}$,

$$\mathrm{Ent}_{\sigma_{n-1}}(f^2) \leq \frac{2}{n-1} \int |\nabla_S f|^2 \, d\sigma_{n-1}. \tag{3.1}$$

To see the connection of (3.1) with the concentration phenomenon on the sphere in the form (1.2), one may apply (2.6) with $u = f$ and $t = \frac{n-1}{4}$.

We are also in the position to apply the abstract Proposition 2.1 to $(S^{n-1}, \rho, \sigma_{n-1})$ and thus involve the second order modulus of the gradients, $|\nabla_S^2 f|$. On the unit sphere it is defined according to (2.2)-(2.3) with $|\nabla f|$ replaced by

$$|\nabla_S f(\theta)| \, = \, \limsup_{\theta' \rightarrow \theta} \frac{|f(\theta) - f(\theta')|}{\rho(\theta, \theta')} \qquad (\theta, \theta' \in S^{n-1}).$$

Note that both the geodesic and Euclidean metrics on S^{n-1} may equivalently be used for computing the modulus of the gradient of first and second orders.

For example, the Euclidean derivatives of the linear function $f(x) = \langle v, x\rangle$ are just $\nabla f(x) = v$ and $f'' = 0$. As for the first and second order modulus of its spherical gradient, we have

$$|\nabla_S f(\theta)| = \sqrt{|v|^2 - \langle v, \theta\rangle^2} \qquad (|v| = 1),$$

and, by the chain rule,

$$\begin{aligned} \nabla_S |\nabla_S f(\theta)| &= -\frac{1}{2\sqrt{|v|^2 - \langle v, \theta\rangle^2}}\, \nabla_S \left(\langle v, \theta\rangle^2\right) \\ &= -\frac{1}{\sqrt{|v|^2 - \langle v, \theta\rangle^2}}\, \langle v, \theta\rangle\, \nabla_S \langle v, \theta\rangle \qquad (\theta \neq v). \end{aligned}$$

Hence, $|\nabla_S^2 f(\theta)| = |\,\langle v, \theta\rangle\,|$ in contrast with $|\nabla^2 f(\theta)| = 0$.

To simplify the condition $|\nabla_S^2 f| \leq 1$, one may use the following equality which is a full analog of the formula (2.7) mentioned before for the case of open regions in $\mathbf{R}^n$.

Lemma 3.1. *Given a C^2-smooth function f on S^{n-1}, $|\nabla_S f|$ has a finite Lipschitz semi-norm and, for all $\theta \in S^{n-1}$,*

$$|\nabla_S^2 f(\theta)| \, = \, |\nabla_S f(\theta)|^{-1}\, |f_S''(\theta) \nabla_S f(\theta)|,$$

where the right-hand side is understood as $\|f_S''(\theta)\|$ in case $|\nabla_S f(\theta)| = 0$. In particular, $|\nabla_S^2 f(\theta)| \leq \|f_S''(\theta)\|$.

The proof is given in Appendix (Section 10).

Thus, in order to bound exponential moments of f similarly to (2.4), one may require the condition $\|f_S''\| \leq 1$. There is however an alternative way based on the application of Corollary 2.2; the latter would allow us to work with Euclidean derivatives. Let us state both consequences of the logarithmic Sobolev inequality (3.1). Henceforth we shall always understand the mean of functions on the unit sphere to be taken with respect to the measure σ_{n-1}.

Corollary 3.2. *Let f be a C^2-smooth function on S^{n-1} with mean zero. If $\|f_S''\| \leq 1$, then*

$$\log \int \exp\Big\{\frac{n-1}{2}\, f\Big\}\, d\sigma_{n-1} \leq \frac{n-1}{2} \int |\nabla_S f|^2\, d\sigma_{n-1}. \tag{3.2}$$

Moreover, if f is C^2-smooth in an open neighbourhood of the unit sphere with $\|f''\| \leq 1$ on S^{n-1}, then

$$\log \int \exp\Big\{\frac{n-1}{2}\, f\Big\}\, d\sigma_{n-1} \leq \frac{n-1}{2} \int |\nabla f|^2\, d\sigma_{n-1}. \tag{3.3}$$

Applying (3.2) to functions εf with $\varepsilon \to 0$, this inequality returns us to (1.1) with an additional factor 2. The condition $\|\varepsilon f''_S\| \leq 1$ is fulfilled for all ε small enough, so any constraint on the second derivative may be removed from the conclusion. In this sense, Corollary 3.2 provides a sharper form of the Poincaré inequality.

4. **Second Derivative and Laplacian**

In order to estimate the integral appearing on the right-hand side in (3.2), we first derive the formula (1.4), involving the square of the spherical Laplacian, i.e. the operator $\Delta_S^2 f = \Delta_S \Delta_S f$. Given a point $\theta \in S^{n-1}$, it will be convenient to work with the spherical second derivative $f''_S(\theta)$ as a symmetric $n \times n$ matrix, i.e. as a linear operator on $\mathbf{R}^n$, rather than as a linear operator on the tangent space $\theta^\perp$. More precisely, we extend the usual Hessian of f at θ to the whole space by putting $f''_S(\theta)\theta = 0$ (in particular, both the operator norm and the Hilbert-Schmidt norm will not increase for the extended matrix). The extended Hessian $f''_S(\theta)$ may also be defined as the $n \times n$ matrix B with the smallest Hilbert-Schmidt norm, satisfying the Taylor expansion

$$\begin{aligned} f(\theta') &= f(\theta) + \langle \nabla_S f(\theta), \theta' - \theta \rangle \\ &\quad + \frac{1}{2} \langle B(\theta' - \theta), \theta' - \theta \rangle + o\big(|\theta' - \theta|^2\big) \qquad (\theta' \to \theta, \quad \theta' \in S^{n-1}). \end{aligned}$$

When f is C^2-smooth in an open region containing the unit sphere, the spherical second derivative is related to the Euclidean derivatives by

$$f''_S(\theta) = P_{\theta^\perp} B P_{\theta^\perp}, \qquad B = f''(\theta) - \langle \nabla f(\theta), \theta \rangle \, I_n,$$

where $P_{\theta^\perp}$ is the projection operator from $\mathbf{R}^n$ to the space $\theta^\perp$ orthogonal to θ. Also, recall that $\nabla_S f(\theta) = P_{\theta^\perp} \nabla f(\theta)$.

Proposition 4.1. *For any C^4-smooth function f on S^{n-1},*

$$\int \|f''_S\|_{\mathrm{HS}}^2 \, d\sigma_{n-1} = \int f \left(\Delta_S^2 f + (n-2)\, \Delta_S f\right) d\sigma_{n-1}. \tag{4.1}$$

One can give a short proof of (4.1) on the basis of the Bochner-Lichnerowicz formula in Riemannian Geometry (cf. Remark 4.6 below). Nevertheless, for the reader's convenience, we shall provide a direct argument based on partial integration formulas in the multivariate calculus on the sphere which we supply in the Appendix, sections A-G. The first of these formulas connects the spherical second derivative with the iteration of spherical derivatives. The second one is a formula for the commutator of the Laplacian and the gradient.

Lemma 4.2. *Given a C^2-smooth function f on S^{n-1}, for all $\theta \in S^{n-1}$ and $v \in \mathbf{R}^n$,*

$$f_S''(\theta)v \,=\, \nabla_S \left\langle \nabla_S f(\theta), v\right\rangle + \langle v, \theta\rangle \, \nabla_S f(\theta). \tag{4.2}$$

Lemma 4.3. *Given a C^3-smooth function f on S^{n-1}, for all $\theta \in S^{n-1}$ and $v \in \mathbf{R}^n$,*

$$\Delta_S \left\langle \nabla_S f(\theta), v\right\rangle - \left\langle \nabla_S \Delta_S f(\theta), v\right\rangle \,=\, (n-3) \left\langle \nabla_S f(\theta), v\right\rangle - 2\,\langle v, \theta\rangle \, \Delta_S f(\theta).$$

The spherical Laplacian appears when integrating by parts, in particular in the formula

$$\int \langle \nabla_S f, \nabla_S g\rangle \, d\sigma_{n-1} = -\int f \Delta_S g \, d\sigma_{n-1}. \tag{4.3}$$

The following analogous identity involves a linear weight.

Lemma 4.4. *For all C^2-smooth functions f, g on S^{n-1} and for any $v \in \mathbf{R}^n$,*

$$\begin{aligned}\int \langle \nabla_S f(\theta), \nabla_S g(\theta)\rangle \, \langle v, \theta\rangle \, d\sigma_{n-1}(\theta) \ &= \ -\int f(\theta) \Delta_S g(\theta) \, \langle v, \theta\rangle \, d\sigma_{n-1}(\theta) \\ &\quad - \int f(\theta) \, \langle \nabla_S g(\theta), v\rangle \, d\sigma_{n-1}(\theta).\end{aligned}$$

Finally, let us mention how to relate the spherical Laplacian to the Euclidean derivatives. The next representation will be used in Section 6.

Lemma 4.5. *If f is C^2-smooth in an open region containing the unit sphere, then for any $\theta \in S^{n-1}$,*

$$\Delta_S f(\theta) = \Delta f(\theta) - (n-1) \, \langle \nabla f(\theta), \theta\rangle - \langle f''(\theta)\theta, \theta\rangle \, .$$

Proof of Proposition 4.1. Using (4.2), one may write

$$\begin{aligned}\int \|f_S''\|_{\mathrm{HS}}^2 \, d\sigma_{n-1} \ &= \ n \iint |f_S''(\theta)v|^2 \, d\sigma_{n-1}(\theta) d\sigma_{n-1}(v) \\ &= \ n \iint \left|\nabla_S \left\langle \nabla_S f(\theta), v\right\rangle + \langle v, \theta\rangle \, \nabla_S f(\theta)\right|^2 d\sigma_{n-1}(\theta) d\sigma_{n-1}(v) \\ &= \ n\,(I_1 + 2I_2 + I_3),\end{aligned}$$

where

$$\begin{aligned} I_1 &= \iint \left| \nabla_S \left\langle \nabla_S f(\theta), v \right\rangle \right|^2 d\sigma_{n-1}(\theta) d\sigma_{n-1}(v), \\ I_2 &= \iint \left\langle \nabla_S \left\langle \nabla_S f(\theta), v \right\rangle, \nabla_S f(\theta) \right\rangle \left\langle v, \theta \right\rangle d\sigma_{n-1}(\theta) d\sigma_{n-1}(v), \\ I_3 &= \iint |\nabla_S f(\theta)|^2 \left\langle v, \theta \right\rangle^2 d\sigma_{n-1}(\theta) d\sigma_{n-1}(v). \end{aligned}$$

Integration over v immediately gives

$$I_3 = \frac{1}{n} \int |\nabla_S f|^2 \, d\sigma_{n-1},$$

and according to (4.3),

$$I_1 = -\iint \varphi_v(\theta) \Delta_S \varphi_v(\theta) \, d\sigma_{n-1}(\theta) d\sigma_{n-1}(v), \quad \text{where} \quad \varphi_v(\theta) = \left\langle \nabla_S f(\theta), v \right\rangle.$$

To continue, we apply Lemma 4.3, so as to develop $\Delta_S \varphi_v(\theta)$ and represent the above integral in the form

$$I_1 = -\big(I_{11} + (n-3)\, I_{12} - 2 I_{13}\big)$$

with

$$\begin{aligned} I_{11} &= \iint \left\langle \nabla_S f(\theta), v \right\rangle \left\langle \nabla_S \Delta_S f(\theta), v \right\rangle d\sigma_{n-1}(\theta) d\sigma_{n-1}(v), \\ I_{12} &= \iint \left\langle \nabla_S f(\theta), v \right\rangle^2 d\sigma_{n-1}(\theta) d\sigma_{n-1}(v), \\ I_{13} &= \iint \left\langle \nabla_S f(\theta), v \right\rangle \left\langle v, \theta \right\rangle \Delta_S f(\theta) \, d\sigma_{n-1}(\theta) d\sigma_{n-1}(v). \end{aligned}$$

Let us now integrate over v and apply (4.3) with $g = \Delta_S f$ to simplify the first equality as

$$\begin{aligned} I_{11} &= \frac{1}{n} \int \left\langle \nabla_S f(\theta), \nabla_S \Delta_S f(\theta) \right\rangle d\sigma_{n-1}(\theta) \\ &= -\frac{1}{n} \int f \, \Delta_S (\Delta_S f) \, d\sigma_{n-1} \; = \; -\frac{1}{n} \int f \, \Delta_S^2 f \, d\sigma_{n-1}. \end{aligned}$$

We also have

$$I_{12} = \frac{1}{n} \int |\nabla_S f|^2 \, d\sigma_{n-1}, \qquad I_{13} = \frac{1}{n} \int \left\langle \nabla_S f(\theta), \theta \right\rangle \Delta_S f(\theta) \, d\sigma_{n-1}(\theta) = 0.$$

This finally gives

$$I_1 \; = \; \frac{1}{n} \int f \, \Delta_S^2 f \, d\sigma_{n-1} - \frac{n-3}{n} \int |\nabla_S f|^2 \, d\sigma_{n-1}.$$

In order to evaluate the integral I_2, we apply Lemma 4.4 with the function $\langle \nabla_S f(\theta), v\rangle$ in place of f and with f in place of g. After integration over θ, we obtain the integral over the remaining variable v, namely,

$$I_2(v) = -\int \langle \nabla_S f(\theta), v\rangle \, \Delta_S f(\theta) \, \langle v, \theta\rangle \, d\sigma_{n-1}(\theta) - \int \langle \nabla_S f(\theta), v\rangle^2 \, d\sigma_{n-1}(\theta).$$

The subsequent integration over v cancels the first integral, since its integrand will contain the inner product $\langle \nabla_S f(\theta), \theta\rangle = 0$ as a factor. As a result,

$$\begin{aligned} I_2 &= \int I_2(v) \, d\sigma_{n-1}(v) \\ &= -\iint \langle \nabla_S f(\theta), v\rangle^2 \, d\sigma_{n-1}(\theta) d\sigma_{n-1}(v) \ = \ -\frac{1}{n} \int |\nabla_S f(\theta)|^2 \, d\sigma_{n-1}(\theta). \end{aligned}$$

It remains to collect these formulas and conclude that

$$n\,(I_1 + 2I_2 + I_3) \ = \ \int f \, \Delta_S^2 f \, d\sigma_{n-1} - (n-2) \int |\nabla_S f|^2 \, d\sigma_{n-1}.$$

Here the last integral can also be written as $-\int f \Delta_S f \, d\sigma_{n-1}$, cf. (4.3). □

Remark 4.6. According to the Bochner-Lichnerowicz formula (cf. e.g. [B-G-L], p. 509), for any smooth function f on the Riemannian manifold (M,g),

$$\frac{1}{2}\, \Delta_g(|\nabla f|^2) = \langle \nabla f, \nabla(\Delta_g f)\rangle + |\nabla\nabla f|^2 + Ric_g(\nabla f, \nabla f), \tag{4.4}$$

where $Ric_g(\nabla f, \nabla f)$ is the Ricci curvature of (M,g) evaluated at ∇f. The unit sphere $M = S^{n-1}$ in $\mathbf{R}^n$ has a constant curvature, namely, in this case

$$Ric_g(\nabla f, \nabla f) = (n-2)|\nabla_S f|^2.$$

Hence, integrating (4.4) over the sphere, we get

$$\int \Big(\frac{1}{2}\, \Delta_S(|\nabla_S f|^2) - \langle \nabla_S f, \nabla_S(\Delta_S f)\rangle\Big) \, d\sigma_{n-1} = \int (|\nabla_S \nabla_S f|^2 + (n-2)|\nabla_S f|^2) \, d\sigma_{n-1}. \tag{4.5}$$

On the other hand, $\int \Delta_S(|\nabla_S f|^2) \, d\sigma_{n-1} = 0$,

$$\int |\nabla_S f|^2 \, d\sigma_{n-1} = -\int f \Delta_S f \, d\sigma_{n-1},$$

(recall (4.3)), and

$$\int \langle \nabla_S f, \nabla_S(\Delta_S f)\rangle \, d\sigma_{n-1} = -\int f \Delta_S^2 f \, d\sigma_{n-1}.$$

Applying these relations in (4.5), we arrive at (4.1).

5. **Expansions in Spherical Harmonics**

Using Proposition 4.1, one may study relations of the form

$$c\int |\nabla_S f|^2\, d\sigma_{n-1} \leq \int \|f_S''\|_{\rm HS}^2\, d\sigma_{n-1} \qquad (c>0) \tag{5.1}$$

by means of the orthogonal expansion in spherical harmonics,

$$f = \sum_{d=0}^\infty f_d \qquad (f_d \in H_d). \tag{5.2}$$

As is well-known (cf. e.g. [S-W]), the Hilbert space $L^2(S^{n-1})$ can be decomposed into a sum of orthogonal linear subspaces H_d, $d = 0,1,2,\dots$, consisting of all d-homogeneous harmonic polynomials (more precisely - restrictions of such polynomials to the sphere). Any element f_d of H_d represents an eigenfunction of the Laplacian, with the eigenvalue $-d(n+d-2)$. That is,

$$\Delta_S f_d = -d(n+d-2)\, f_d,$$

and hence

$$\Delta_S^2 f_d \ = \ d^2 (n+d-2)^2\, f_d.$$

As a result,

$$\Delta_S f = -\sum_{d=1}^\infty d(n+d-2) f_d, \qquad \Delta_S^2 f = \sum_{d=1}^\infty d^2(n+d-2)^2 f_d$$

which should be understood as equalities in L^2 (Note that both $\Delta_S f$ and $\Delta_S^2 f$ are continuous functions, as long as f is C^4-smooth).

According to the representation (4.1), (5.1) is equivalent to

$$\int f\, \Delta_S^2 f\, d\sigma_{n-1} \ \geq \ -(c+n-2)\int f\, \Delta_S f\, d\sigma_{n-1}. \tag{5.3}$$

Moreover, since the spherical harmonics serve as eigenfunctions both for Δ_S and Δ_S^2, the last inequality need to be verified for elements f_d of H_d only. Here, both integrals are vanishing for constant functions, i.e. for $f \in H_d$ with $d = 0$. If $d \geq 1$, (5.3) becomes

$$c \ \leq \ d^2 + (d-1)(n-2). \tag{5.4}$$

Thus, if we want to involve in (5.1) all C^2-smooth functions f, the optimal value of c is described as the minimum of the right-hand side of (5.4) over all $d \geq 1$. The minimum is achieved for $d = 1$ which leads to the optimal value $c = 1$. However, if we require that f is orthogonal to all linear functions with respect to σ_{n-1}, it means that we only allow the values $d \geq 2$ in (5.4), and then the optimal value is $c = n+2$. As a result, we have proved:

Proposition 5.1. *For any C^2-function f on S^{n-1},*

$$\int |\nabla_S f|^2 \, d\sigma_{n-1} \leq \int \|f''_S\|_{\mathrm{HS}}^2 \, d\sigma_{n-1},$$

where equality is attained for all linear functions. Moreover, if f is orthogonal to all linear functions with respect to σ_{n-1}, then

$$\int |\nabla_S f|^2 \, d\sigma_{n-1} \leq \frac{1}{n+2} \int \|f''_S\|_{\mathrm{HS}}^2 \, d\sigma_{n-1} \tag{5.5}$$

with equality attainable for all quadratic harmonics.

The expansion (5.2) is commonly used to derive Poincaré-type inequalities such as (1.1). If we require additionally that f should be orthogonal to all linear functions, the constant will slightly improve only, since then

$$\int f^2 \, d\sigma_{n-1} \leq \frac{1}{2n} \int |\nabla_S f|^2 \, d\sigma_{n-1}.$$

This bound may be combined with (5.5) to get a second order Poincaré-type inequality which was mentioned in the Introduction. But, one can also apply (5.2) directly in the representation (4.1). Indeed, on spherical harmonics f_d of H_d, the inequality of the form $c \int f^2 \, d\sigma_{n-1} \leq \int \|f''_S\|_{\mathrm{HS}}^2 \, d\sigma_{n-1}$ becomes

$$c \, \leq \, d(n+d-2) \, \big(d(n+d-2) - (n-2)\big).$$

Since the right-hand side is an increasing function of d, we arrive at:

Proposition 5.2. *For any C^2-function f on S^{n-1} with mean zero,*

$$\int f^2 \, d\sigma_{n-1} \leq \frac{1}{n-1} \int \|f''_S\|_{\mathrm{HS}}^2 \, d\sigma_{n-1}, \tag{5.6}$$

where equality is attained for all linear functions. Moreover, if f is orthogonal to all linear functions with respect to σ_{n-1}, then

$$\int f^2 \, d\sigma_{n-1} \leq \frac{1}{2n(n+2)} \int \|f''_S\|_{\mathrm{HS}}^2 \, d\sigma_{n-1} \tag{5.7}$$

with equality attainable for all quadratic harmonics.

An interesting consequence of (5.6) is the statement that the equality $f''_S = 0$ is possible for constant functions, only (in contrast with the Euclidean Hessian).

Remark 5.3. It is much easier to derive (5.7) with suboptimal, although asymptotically correct constants as n tends to infinity, without appealing to Proposition 4.1. The argument is based on the double application of the Poincaré inequality (1.1). Orthogonality of f to all linear functions ensures that the function $\theta \to \langle \nabla_S f(\theta), v \rangle$ has mean zero for any $v \in \mathbf{R}^n$. So, using the identity (4.2), we get

$$\begin{aligned}(n-1)\int \langle \nabla_S f(\theta), v\rangle^2 \, d\sigma_{n-1}(\theta) \ &\leq \ \int |f_S''(\theta)v - \langle v,\theta\rangle \, \nabla_S f(\theta)|^2 \, d\sigma_{n-1}(\theta) \\ &= \ \int |f_S''(\theta)v|^2 \, d\sigma_{n-1}(\theta) + \int \langle v,\theta\rangle^2 \, |\nabla_S f(\theta)|^2 \, d\sigma_{n-1}(\theta) \\ &\quad - 2\int \langle f_S''(\theta)\nabla_S f(\theta), v\rangle \, \langle v,\theta\rangle \, d\sigma_{n-1}(\theta).\end{aligned}$$

The next integration over $d\sigma_{n-1}(v)$ cancels the last integral (due to $f_S''(\theta)\theta = 0$), and we are led to

$$(n-2)\int |\nabla_S f(\theta)|^2 \, d\sigma_{n-1}(\theta) \ \leq \ \int \|f_S''(\theta)\|_{\mathrm{HS}}^2 \, d\sigma_{n-1}(\theta).$$

If f has mean zero, the left integral is estimated from below according to (1.1), and thus

$$\int f^2 \, d\sigma_{n-1} \leq \frac{1}{(n-1)(n-2)} \int \|f_S''\|_{\mathrm{HS}}^2 \, d\sigma_{n-1}, \qquad n \geq 3.$$

The constant in this inequality is slightly worse than (5.7), and we loose information about extremal functions.

The above argument is also applicable in the Euclidean setup when dealing with a probability measure μ on $\mathbf{R}^n$ satisfying a Poincaré-type inequality

$$\int f^2 \, d\mu \leq \sigma^2 \int |\nabla f|^2 \, d\mu \qquad \Big(\int f \, d\mu = 0\Big).$$

For example, the standard Gaussian measure with density $\frac{d\mu(x)}{dx} = (2\pi)^{-n/2} \, e^{-|x|^2/2}$ has the Poincaré constant $\sigma^2 = 1$, which yields a second order Poincaré-type inequality

$$\int f^2 \, d\mu \leq \frac{1}{2} \int \|f_S''\|_{\mathrm{HS}}^2 \, d\mu.$$

It holds true in the class of all C^2-smooth functions f on $\mathbf{R}^n$ that are orthogonal to all affine functions in $L^2(\mu)$. However, in the general case, orthogonality to linear functions should be replaced with the requirement $\int \nabla f \, d\mu = 0$.

We are now prepared to complete the proof of Theorem 1.1.

Proof of Theorem 1.1. Let us return to the bound (3.2) of Corollary 3.2. Using (5.5), we then get

$$\log \int \exp\Big\{\frac{n-1}{2}\, f\Big\} \, d\sigma_{n-1} \ \leq \ \frac{n-1}{2(n+2)} \int \|f_S''\|_{\mathrm{HS}}^2 \, d\sigma_{n-1} \ \leq \ \frac{1}{2}\, b^2,$$

and with a similar inequality for the function $-f$,

$$\int e^{\frac{n-1}{2}\,|f|} \, d\sigma_{n-1} \ \leq \ \int e^{\frac{n-1}{2}\, f} \, d\sigma_{n-1} + \int e^{-\frac{n-1}{2}\, f} \, d\sigma_{n-1} \ \leq \ e^{b^2/2}.$$

It follows that, for any $\lambda \geq 1$,

$$\int e^{\frac{n-1}{2}|f|/\lambda}\,d\sigma_{n-1} \leq \Big(\int e^{\frac{n-1}{2}|f|}\,d\sigma_{n-1}\Big)^{1/\lambda} \leq (2e^{b^2/2})^{1/\lambda}.$$

It remains to note that $(2e^{b^2/2})^{1/\lambda} = 2$ for $\lambda = 1 + \frac{b^2}{\log 4} \leq 1 + b^2$. $\square$

6. Bounds on the L^2-Norm of the Euclidean Gradient

We now turn back to Theorem 1.2 while invoking the second bound of Corollary 3.2. Hence, we need an analog of (5.5) for the modulus of the Euclidean gradient. Assume that a function f is defined and C^2-smooth in some neighbourhood G of S^{n-1}.

Proposition 6.1. *If f is orthogonal to all linear functions with respect to σ_{n-1}, then*

$$\int |\nabla f|^2\,d\sigma_{n-1} \leq \frac{5}{n-1}\int \|f''\|_{\rm HS}^2\,d\sigma_{n-1}. \tag{6.1}$$

At the of the proof it will be apparent that for growing dimensions the constant 5 may be asymptotically improved to 2.

Proof. Since the spherical gradient $\nabla_S f(\theta)$ represents the projection of the usual gradient $\nabla f(\theta)$ to the subspace $\theta^\perp$ of $\mathbf{R}^n$ orthogonal to θ, we have

$$|\nabla f|^2 \,=\, |\nabla_S f(\theta)|^2 + \langle \nabla f(\theta), \theta\rangle^2\,.$$

As a preliminary step, first we show that

$$\int |\nabla_S f|^2\,d\sigma_{n-1} \leq \frac{1}{n-1}\int \|f''\|_{\rm HS}^2\,d\sigma_{n-1}. \tag{6.2}$$

Write

$$\int |\nabla_S f|^2\,d\sigma_{n-1} \,=\, \int |\nabla f(\theta)|^2\,d\sigma_{n-1}(\theta) - \int \langle \nabla f(\theta), \theta\rangle^2\,\,d\sigma_{n-1}(\theta) \tag{6.3}$$

and represent

$$\int |\nabla f|^2\,d\sigma_{n-1} = n\iint \langle \nabla f(\theta), v\rangle^2\,\,d\sigma_{n-1}(\theta)d\sigma_{n-1}(v). \tag{6.4}$$

The assumption that f is orthogonal to all linear functions is equivalent to the property that every function of the form

$$\langle \nabla_S f(\theta), v\rangle = \langle \nabla f(\theta), v\rangle - \langle \nabla f(\theta), \theta\rangle\,\langle v, \theta\rangle$$

has σ_{n-1}-mean zero (by the integration by parts formula). Hence

$$\int \langle \nabla f(\theta), v\rangle\,\,d\sigma_{n-1}(\theta) = \int \langle \nabla f(\theta), \theta\rangle\,\langle v, \theta\rangle\,\,d\sigma_{n-1}(\theta),$$

and, by the Cauchy-Schwarz inequality,

$$\Big(\int \langle \nabla f(\theta), v\rangle \, d\sigma_{n-1}(\theta)\Big)^2 \leq \frac{1}{n} \int \langle \nabla f(\theta), \theta\rangle^2 \, d\sigma_{n-1}(\theta). \tag{6.5}$$

To estimate the L^2-norm of $\langle \nabla f(\theta), v\rangle$, one may apply the Poincaré inequality (1.1). Since $u(x) = \langle \nabla f(x), v\rangle$ has gradient $\nabla u(x) = f''(x)v$, we have, by (6.5),

$$\int \langle \nabla f(\theta), v\rangle^2 \, d\sigma_{n-1}(\theta) \leq \frac{1}{n} \int \langle \nabla f(\theta), \theta\rangle^2 \, d\sigma_{n-1}(\theta) + \frac{1}{n-1} \int |f''(\theta)v|^2 \, d\sigma_{n-1}(\theta).$$

Using this bound in (6.4) and integrating over v, we get

$$\int |\nabla f|^2 \, d\sigma_{n-1} \ \leq \ \int \langle \nabla f(\theta), \theta\rangle^2 \, d\sigma_{n-1}(\theta) + \frac{1}{n-1} \int \|f''(\theta)\|_{\mathrm{HS}}^2 \, d\sigma_{n-1}(\theta).$$

It remains to insert this bound in (6.3) which gives (6.2).

Now, rewrite (6.3) as

$$\int |\nabla f|^2 \, d\sigma_{n-1} = \int |\nabla_S f|^2 \, d\sigma_{n-1} + \int \langle \nabla f(\theta), \theta\rangle^2 \, d\sigma_{n-1}(\theta). \tag{6.6}$$

Here, the first integral on the right-hand side is estimated in terms of $\|f''\|_{\mathrm{HS}}^2$ by (6.2), and our next task will be to derive a suitable bound on the L^2-norm of the function $\langle \nabla f(\theta), \theta\rangle$. To this aim, we employ the representation of Lemma 4.5 for the spherical Laplacian in terms of the Euclidean derivatives. Since in general (by (4.3)),

$$\int \Delta_S f \, d\sigma_{n-1} = -\int \langle \nabla_S 1, \nabla_S f\rangle \, d\sigma_{n-1} = 0,$$

Lemma 4.5 yields

$$(n-1) \int \langle \nabla f(\theta), \theta\rangle \, d\sigma_{n-1}(\theta) = \int \big(\Delta f(\theta) - \langle f''(\theta)\theta, \theta\rangle \big) \, d\sigma_{n-1}(\theta). \tag{6.7}$$

Here the second integrand is equal to

$$I = \sum_{i,j=1}^n \partial_{ij} f(\theta) a_{ij} \quad \text{with} \ \ a_{ij} = \delta_{ij} - \theta_i \theta_j.$$

Note that

$$\sum_{i,j=1}^n a_{ij}^2 \ = \ \sum_{i \neq j}^n \theta_i^2 \theta_j^2 + \sum_{i=1}^n (1-\theta_i^2)^2 \ = \ 1 + \sum_{i=1}^n \big((1-\theta_i^2)^2 - \theta_i^4\big) \ = \ n-1.$$

Hence, by Cauchy's inequality,

$$I^2 \leq \sum_{i,j=1}^n (\partial_{ij} f(\theta))^2 \sum_{i,j=1}^n a_{ij}^2 = (n-1)\, \|f''(\theta)\|_{\mathrm{HS}}^2,$$

and by another application of the Cauchy-Schwarz inequality in (6.7),

$$\Big(\int \langle \nabla f(\theta), \theta\rangle \, d\sigma_{n-1}(\theta)\Big)^2 \, \leq \, \frac{1}{n-1} \int \|f''\|_{\rm HS}^2 \, d\sigma_{n-1}. \tag{6.8}$$

Next, consider the function $u(x) = \langle \nabla f(x), x\rangle$ and restrict its gradient $\nabla u(x) = \nabla f(x) + f''(x)x$ to the unit sphere. Projecting it to $\theta^\perp$, we obtain the spherical gradient

$$\nabla_S u(\theta) = \nabla_S f(\theta) + P_{\theta^\perp}\big(f''(\theta)\theta\big), \qquad \theta \in S^{n-1}.$$

In particular, by the triangle inequality,

$$|\nabla_S u(\theta)| \leq |\nabla_S f(\theta)| + \|f''(\theta)\|.$$

Furthermore, the square of the right-hand side can be estimated by using the elementary inequality $(x+y)^2 \leq \frac{\lambda}{\lambda-1}\, x^2 + \lambda y^2$ ($x,y \geq 0$, $\lambda > 1$), which implies

$$|\nabla_S u(\theta)| \leq \frac{\lambda}{\lambda-1}\, |\nabla_S f(\theta)|^2 + \lambda\, \|f''(\theta)\|^2.$$

Hence, using the Poincaré inequality together with (6.8), and increasing the operator norm to the Hilbert-Schmidt norm, we get

$$\begin{aligned}
\int u^2\, d\sigma_{n-1} \ & \leq \ \Big(\int u\, d\sigma_{n-1}\Big)^2 + \frac{1}{n-1} \int |\nabla_S u|^2\, d\sigma_{n-1} \\
& \leq \ \frac{1}{n-1} \int \|f''\|_{\rm HS}^2\, d\sigma_{n-1} + \frac{1}{n-1} \int \Big(\frac{\lambda}{\lambda-1}\, |\nabla_S f|^2 + \lambda\, \|f''\|_{\rm HS}^2\Big)\, d\sigma_{n-1}.
\end{aligned}$$

Thus,

$$\int \langle \nabla f(\theta), \theta\rangle^2 \, d\sigma_{n-1}(\theta) \, \leq \, \frac{1}{n-1}\, \frac{\lambda}{\lambda-1} \int |\nabla_S f|^2\, d\sigma_{n-1} + \frac{\lambda+1}{n-1} \int \|f''\|_{\rm HS}^2\, d\sigma_{n-1}.$$

It remains to return to (6.6) and combine the above bound with (6.2). Adding and collecting the coefficients, it gives

$$(n-1) \int |\nabla f|^2\, d\sigma_{n-1} \, \leq \, \Big(\frac{1}{n-1}\, \frac{\lambda}{\lambda-1} + \lambda + 1\Big) \int \|f''\|_{\rm HS}^2\, d\sigma_{n-1}.$$

The quantity $\frac{1}{n-1}\, \frac{\lambda}{\lambda-1} + \lambda + 1$ is minimized at $\lambda = 1 + \frac{1}{\sqrt{n-1}}$, which leads to

$$\int |\nabla f|^2\, d\sigma_{n-1} \leq \frac{c_n}{n-1} \int \|f''\|_{\rm HS}^2\, d\sigma_{n-1}, \qquad c_n = 1 + \Big(1 + \frac{1}{\sqrt{n-1}}\Big)^2. \tag{6.9}$$

Clearly, $c_n \leq 5$, thus proving (6.1). $\square$

Note that $c_n \to 2$ as $n \to \infty$. So, the constant 5 in (6.1) may be improved for large values of n.

Combining (6.1) with the Poincaré inequality (1.1), we get a second order Poincaré-type inequality in the Euclidean setup,

$$\int (f-m)^2\,d\sigma_{n-1} \le \frac{5}{(n-1)^2}\int \|f''\|_{\rm HS}^2\,d\sigma_{n-1},$$

assuming that f is orthogonal to all linear functions, and where m is the mean of f with respect to σ_{n-1}. Here the left integral will not change when it is applied to $f_a(x) = f(x) - \frac{a}{2}\,|x|^2$ in place of f, while the right integral will depend on a. More precisely, we get

$$\int (f-m)^2\,d\sigma_{n-1} \le \frac{5}{(n-1)^2}\int \|f'' - aI_n\|_{\rm HS}^2\,d\sigma_{n-1},$$

Hence, we arrive at:

Corollary 6.2. *If f is orthogonal to all affine functions with respect to σ_{n-1}, then for any $a \in \mathbf{R}$,*

$$\int f^2\,d\sigma_{n-1} \le \frac{5}{(n-1)^2}\int \|f'' - aI_n\|_{\rm HS}^2\,d\sigma_{n-1}.$$

7. Proof of Theorems 1.2-1.3

Having proved Proposition 6.1, the proof of Theorem 1.2 is almost identical to the proof of Theorem 1.1.

Proof of Theorem 1.2. Let f be orthogonal to all linear functions with mean m. Applying (6.1) to the function $f - m$ in the bound (3.3) of Corollary 3.2, we get

$$\log \int \exp\Big\{\frac{n-1}{2}\,(f-m)\Big\}\,d\sigma_{n-1} \ \le\ \frac{5}{2}\int \|f''\|_{\rm HS}^2\,d\sigma_{n-1}.$$

Applying it to $f_a(x) = f(x) - \frac{a}{2}\,|x|^2$ in place of f, we get

$$\log \int \exp\Big\{\frac{n-1}{2}\,(f-m)\Big\}\,d\sigma_{n-1} \ \le\ \frac{5}{2}\int \|f'' - aI_n\|_{\rm HS}^2\,d\sigma_{n-1} \ \le\ \frac{5}{2}\,b^2.$$

Assuming that $m = 0$ and applying a similar inequality to the function $-f$, we obtain

$$\int e^{\frac{n-1}{2}\,|f|}\,d\sigma_{n-1} \ \le\ 2e^{5b^2/2}.$$

Hence, for any $\lambda \ge 1$,

$$\int e^{\frac{n-1}{2}\,|f|/\lambda}\,d\sigma_{n-1} \ \le\ \Big(\int e^{\frac{n-1}{2}\,|f|}\,d\sigma_{n-1}\Big)^{1/\lambda} \ \le\ (2e^{5b^2/2})^{1/\lambda}.$$

It remains to note that $(2e^{5b^2/2})^{1/\lambda} = 2$ for $\lambda = 1 + \frac{5b^2}{\log 4} \le 1 + 3.7\,b^2$. $\square$

Proof of Theorem 1.3. Let $l(\theta) = \langle v, \theta \rangle$ be the linear part of f, and recall that

$$|v|^2 = n^2 I, \qquad I = \iint \langle x, y\rangle\, f(x) f(y)\, d\sigma_{n-1}(x) d\sigma_{n-1}(y).$$

To control Gaussian tails of l under σ_{n-1}, we apply an exponential bound

$$\int e^{tl(\theta)}\, d\sigma_{n-1}(\theta) \ \leq \ e^{\frac{t^2}{2(n-1)}\, |v|^2}, \qquad t \in \mathbf{R},$$

which is implied by the logarithmic Sobolev inequality on the sphere, (3.1). Choosing $t = n-1$ and using the assumption $I \leq \frac{b_0}{n^3}$, we get $\int e^{(n-1)|l|}\, d\sigma_{n-1} \leq 2e^{b_0^2/2}$ and hence

$$\int \exp\Big\{\frac{n-1}{1+b_0^2}\, |l|\Big\}\, d\sigma_{n-1} \leq 2.$$

On the other hand, by Theorem 1.2 with the same assumption on the second derivative of f, we have

$$\int \exp\Big\{\frac{n-1}{2(1+4b^2)}\, |Tf|\Big\}\, d\sigma_{n-1} \leq 2.$$

Using $|f| \leq |Tf| + |l|$ and applying the Cauchy-Schwarz inequality, we conclude that

$$\int e^{(n-1)\, |f|/2\lambda}\, d\sigma_{n-1} \ \leq \ \Big(\int e^{(n-1)\, |Tf|/\lambda}\, d\sigma_{n-1}\Big)^{1/2} \Big(\int e^{(n-1)\, |l|/\lambda}\, d\sigma_{n-1}\Big)^{1/2} \ \leq \ 2,$$

provided that $\lambda \geq 2(1+4b^2)$ and $\lambda \geq 1 + b_0^2$. $\square$

8. **Appendix A. Definitions of Spherical Derivatives**

A function f defined on the unit sphere S^{n-1} is C^p-smooth, $p = 1, 2, \dots$, if it can be extended to some open set containing S^{n-1} as a C^p-smooth function (in the usual sense). This is one of the well-known definitions of smoothness on the sphere.

If f is C^1-smooth on S^{n-1}, then at every point $\theta \in S^{n-1}$ it admits the Taylor expansion up to the linear term

$$f(\theta') = f(\theta) + \langle v, \theta' - \theta\rangle + o\big(|\theta' - \theta|\big), \quad \text{as} \ \ \theta' \to \theta, \ \ \theta' \in S^{n-1}, \tag{8.1}$$

with some $v \in \mathbf{R}^n$. If v has the smallest length (Euclidean norm) among all such vectors, it is called the spherical derivative or gradient of f at θ and is denoted $\nabla_S f(\theta)$.

This notion of the derivative of f is independent of the choice of a smooth extension of f in an open neighbourhood of the sphere in $\mathbf{R}^n$. If f is C^1-smooth in a neighbourhood of the unit sphere, then (8.1) holds with the usual (Euclidean) gradient $v = \nabla f(\theta)$, and the spherical gradient may be described as

$$\begin{aligned} \nabla_S f(\theta) \ &= \ P_{\theta^\perp} \nabla f(\theta) \\ &= \ \nabla f(\theta) - \langle \nabla f(\theta), \theta\rangle\, \theta, \end{aligned}$$

where $P_{\theta^\perp}$ is the (orthogonal) projection operator from $\mathbf{R}^n$ to $\theta^\perp$ (the tangent space).

In particular, $\langle \nabla_S f(\theta), \theta\rangle = 0$ and $|\nabla_S f(\theta)| \leq |\nabla f(\theta)|$ for any $\theta \in S^{n-1}$.

The spherical gradient of any C^1-function represents a continuous vector-valued function on S^{n-1}.

Analogously (as was already stressed in Section 4), the second derivative of any C^2-smooth function f on the unit sphere at a given point $\theta \in S^{n-1}$ may be introduced via a Taylor expansion up to the quadratic term

$$f(\theta') \,=\, f(\theta) + \langle \nabla_S f(\theta), \theta' - \theta\rangle + \frac{1}{2}\,\langle B(\theta' - \theta), \theta' - \theta\rangle + o\big(|\theta' - \theta|^2\big), \tag{8.2}$$

where $\theta' \to \theta$, $\theta' \in S^{n-1}$, and B is some $n \times n$ matrix (with real entries).

Recall that the space $\mathbf{M}_n$ of all $n \times n$ matrices is naturally identified with the Euclidean space $\mathbf{R}^{n\times n}$ with its inner product and the Euclidean norm

$$\|B\|_{\rm HS} = \bigg(\sum_{i,j=1}^n B_{ij}^2\bigg)^{1/2}$$

called the Hilbert-Schmidt norm of B. The collection of all B satisfying (8.2) represents an affine subspace of $\mathbf{M}_n$. Therefore, among all of them, there exists a unique matrix which has the smallest Hilbert-Schmidt norm. It can be called the (spherical) second derivative of f at the point θ and will be denoted $f_S''(\theta)$.

If f is C^2-smooth in an open neighborhood of S^{n-1}, then in accordance with the usual Taylor expansion, (8.2) holds with the matrix $B_0 = f''(\theta) - \langle \nabla f(\theta), \theta\rangle\, I_n$. More generally, given $A \in \mathbf{M}_n$, the matrix $B_0 - A$ satisfies (8.2), if and only if

$$\langle A(\theta' - \theta), \theta' - \theta\rangle = o\big(|\theta' - \theta|^2\big) \qquad (\theta' \to \theta, \ \ \theta' \in S^{n-1}).$$

But this is equivalent to saying that $\langle Ax, x\rangle = 0$ for all $x \in \theta^\perp$. This condition defines a linear subspace L of $\mathbf{M}_n$, and the problem

$$\|B_0 - A\|_{\rm HS} \to \min \ \text{ over all } A \in L$$

is then solved uniquely for $B = B_0 - A$ being the orthogonal projection in $\mathbf{M}_n$ of B_0 to the linear space $L^\perp$ of all matrices orthogonal to L. In fact, since B_0 is symmetric, in this minimization problem one may restrict ourselves to symmetric matrices, and by a simple algebra, we arrive at the following description.

Proposition 8.1. *The spherical second derivative of f at each point $\theta \in S^{n-1}$ is a symmetric matrix, which is given by the orthogonal projection*

$$f_S''(\theta) = P_{L_\theta^\perp} B, \qquad B = f''(\theta) - \langle \nabla f(\theta), \theta\rangle\, I_n,$$

to the orthogonal complement of the linear subspace L_θ of all symmetric matrices A in $\mathbf{M}_n$ such that $Ax = 0$ for all $x \in \theta^\perp$. Equivalently,

$$f_S''(\theta) = P_{\theta^\perp} B P_{\theta^\perp}.$$

One immediate consequence of this description is that $f_S''(\theta)\theta = 0$ and hence the vectors $f_S''(\theta)v$ are orthogonal to θ, for all $\theta \in S^{n-1}$ and $v \in \mathbf{R}^n$.

One should also emphasize the contraction property

$$\|f_S''(\theta)\|_{\mathrm{HS}} \, \leq \, \|f''(\theta) - \langle \nabla f(\theta), \theta\rangle \, I_n\|_{\mathrm{HS}}$$

and similarly for the operator norm.

9. **Appendix B. Second Order Gradients**

Let us now turn to Lemma 4.2 with its identity

$$f_S''(\theta)v = \nabla_S \left\langle \nabla_S f(\theta), v\right\rangle + \langle v, \theta\rangle \, \nabla_S f(\theta). \tag{9.1}$$

Note that the usual first and second derivatives are connected by

$$f''v = \nabla \left\langle \nabla f, v\right\rangle \qquad (v \in \mathbf{R}^n). \tag{9.2}$$

As follows from (9.1), we have a similar property for the spherical drivatives – however for v in the tangent space, only.

Proof of Lemma 4.2. We may assume that f is defined and C^2-smooth on an open subset G of $\mathbf{R}^n$ containing the unit sphere. To compute the spherical gradient for the function

$$\psi(\theta) \, = \, \langle \nabla_S f(\theta), v\rangle \, \, = \, \langle \nabla f(\theta), v\rangle - \langle \nabla f(\theta), \theta\rangle \, \langle v, \theta\rangle \, ,$$

let us extend it smoothly to all points $x \in G$ by

$$\psi(x) = \langle \nabla f(x), v\rangle - \langle \nabla f(x), x\rangle \, \langle v, x\rangle \tag{9.3}$$

and write

$$\nabla_S \psi(\theta) = \nabla \psi(\theta) - \langle \nabla \psi(\theta), \theta\rangle \, \theta. \tag{9.4}$$

From (9.3) and (9.2),

$$\begin{aligned}
\nabla \psi(x) \;\; &= \;\; \nabla \left\langle \nabla f(x), v\right\rangle - \nabla\big(\left\langle \nabla f(x), x\right\rangle \langle v, x\rangle \big) \\
&= \;\; f''(x)v - \langle \nabla f(x), x\rangle \, v - \nabla\big(\langle \nabla f(x), x\rangle \big) \, \langle v, x\rangle \, .
\end{aligned}$$

In addition, the function $u(x) = \langle \nabla f(x), x\rangle$ has the gradient $\nabla u(x) = f''(x)x + \nabla f(x)$, so

$$\nabla \psi(x) = f''(x)v - \langle \nabla f(x), x\rangle \, v - \big(f''(x)x + \nabla f(x)\big) \, \langle v, x\rangle \, .$$

Restricting this equality to the unit sphere, we get

$$\nabla \psi(\theta) = f''(\theta)Pv - \langle \nabla f(\theta), \theta\rangle \, v - \langle v, \theta\rangle \, \nabla f(\theta), \tag{9.5}$$

which also implies

$$\langle \nabla \psi(\theta), \theta\rangle \, \theta = \langle f''(\theta)Pv, \theta\rangle \, \theta - \langle \nabla f(\theta), \theta\rangle \, \langle v, \theta\rangle \, \theta - \langle v, \theta\rangle \, \langle \nabla f(\theta), \theta\rangle \, \theta. \tag{9.6}$$

where for short $P = P_{\theta^\perp}$.

Following (9.4), it remains to subtract (9.6) from (9.5). First note that

$$f''(\theta)Pv - \langle f''(\theta)Pv, \theta\rangle\, \theta = Pf''(\theta)Pv$$

which is deduced from the general formula $Pw = w - \langle w, \theta\rangle\, \theta$ with $w = f''(\theta)Pv$. The equality $v - \langle v, \theta\rangle\, \theta = Pv$ can be used for the second terms on the right of (9.5)-(9.6). Finally, for the third terms we have

$$\nabla f(\theta) - \langle \nabla f(\theta), \theta\rangle\, \theta = P\, \nabla f(\theta) = \nabla_S f(\theta).$$

Therefore, using the matrix B from Proposition 7.1, the difference between (9.5) and (9.6) is exactly

$$\begin{aligned} Pf''(\theta)Pv - \langle \nabla f(\theta), \theta\rangle\, Pv - \langle v, \theta\rangle\, \nabla_S f(\theta) &= PBP - \langle v, \theta\rangle\, \nabla_S f(\theta) \\ &= f''_S(\theta)v - \langle v, \theta\rangle\, \nabla_S f(\theta). \end{aligned}$$

Thus,

$$\nabla_S \langle \nabla_S f(\theta), v\rangle = f''_S(\theta)v - \langle v, \theta\rangle\, \nabla_S f(\theta), \tag{9.7}$$

which is the desired equality (9.1). □

10. Appendix C. Second Order Modulus of Gradients

Let us give more details explaining Lemma 3.1. Recall that, by the very definition of the second order modulus of the gradient,

$$\begin{aligned} |\nabla_S^2 f(\theta)| &= |\nabla_S\, |\nabla_S f(\theta)|\,| \\ &= \limsup_{\theta'\to\theta} \frac{|\,|\nabla_S f(\theta)| - |\nabla_S f(\theta')|\,|}{|\theta - \theta'|}, \qquad \theta \in S^{n-1}. \end{aligned}$$

Proof of Lemma 3.1. First let us show that the function $|\nabla_S f|$ has a finite Lipschitz semi-norm. Since the first two spherical derivatives of f are continuous and therefore bounded on the unit sphere, it follows from (9.7) that

$$|\nabla_S \langle \nabla_S f(\theta), v\rangle\,| \le C, \qquad |v| = 1,$$

with some constant C (independent of θ and v). Hence, the function $\theta \to \langle \nabla_S f(\theta), v\rangle$ has Lipschitz semi-norm at most C, so that

$$|\,\langle \nabla_S f(\theta'), v\rangle - \langle \nabla_S f(\theta), v\rangle\,| \le C\rho(\theta', \theta)$$

for all $\theta, \theta' \in S^{n-1}$. Taking here the supremun over all unit vectors v and applying the triangle inequality, we get

$$\Big|\, |\nabla_S f(\theta')| - |\nabla_S f(\theta)|\, \Big| \le |\nabla_S f(\theta') - \nabla_S f(\theta)| \le C\rho(\theta', \theta),$$

which is the Lipschitz property (with constant C).

Next, to derive the required identity for the second order modulus of the gradient, we fix $\theta \in S^{n-1}$ and apply the identity (9.7) once more. By the definition of the spherical gradient, it yields the Taylor expansion up to the linear term,

$$\langle \nabla_S f(\theta'), v\rangle = \langle \nabla_S f(\theta), v\rangle + \langle V, \theta' - \theta\rangle + o(|\theta' - \theta|) \tag{10.1}$$

as $\theta' \to \theta$, $\theta' \in S^{n-1}$, where

$$V = f_S''(\theta)v - \langle v, \theta\rangle\, \nabla_S f(\theta).$$

Moreover, by the Taylor formula in the integral form, and since any continuous function on a compact metric space is uniformly continuous, the o-term in (10.1) can be bounded by a quantity which is independent of $v \in S^{n-1}$. That is,

$$\sup_{v \in S^{n-1}} |\, \langle \nabla_S f(\theta') - \nabla_S f(\theta), v\rangle - \langle V, \theta' - \theta\rangle \leq \varepsilon(|\theta' - \theta|)$$

with some function $\varepsilon(t)$ such that $\varepsilon(t) \to 0$ as $t \to 0$.

Now, let us rewrite (10.1) as

$$\langle \nabla_S f(\theta'), v\rangle = \langle \nabla_S f(\theta) + L, v\rangle + o(|\theta' - \theta|), \tag{10.2}$$

where

$$\langle L, v\rangle = \langle V, \theta' - \theta\rangle = \langle f_S''(\theta)v - \langle v, \theta\rangle\, \nabla_S f(\theta), \theta' - \theta\rangle\,,$$

that is, with

$$L = f_S''(\theta)(\theta' - \theta) - \langle \nabla_S f(\theta), \theta' - \theta\rangle\, \theta. \tag{10.3}$$

Taking an absolute value of both sides in (10.2) and turning to the supremum over all $v \in S^{n-1}$, we obtain that

$$|\nabla_S f(\theta')| = |\nabla_S f(\theta) + L| + o(|\theta' - \theta|). \tag{10.4}$$

Next, write

$$|\nabla_S f(\theta) + L|^2 = |\nabla_S f(\theta)|^2 + 2\, \langle \nabla_S f(\theta), L\rangle + |L|^2. \tag{10.5}$$

Since $\nabla_S f(\theta)$ is orthogonal to the vector θ, we have from (10.3) that

$$\langle \nabla_S f(\theta), L\rangle = \langle \nabla_S f(\theta), f_S''(\theta)(\theta' - \theta)\rangle = \langle w, \theta' - \theta\rangle\,,$$

where

$$w = f_S''(\theta)\nabla_S f(\theta).$$

Since also $|L|^2 = O(|\theta' - \theta|^2)$, (10.5) yields

$$|\nabla_S f(\theta) + L|^2 = |\nabla_S f(\theta)|^2 + 2\, \langle w, \theta' - \theta\rangle + o(|\theta' - \theta|),$$

and therefore in case $|\nabla_S f(\theta)| > 0$,

$$|\nabla_S f(\theta) + L| = |\nabla_S f(\theta)| + |\nabla_S f(\theta)|^{-1}\, \langle w, \theta' - \theta\rangle + o(|\theta' - \theta|).$$

Using this in (10.4), we find that

$$|\nabla_S f(\theta')| - |\nabla_S f(\theta)| = |\nabla_S f(\theta)|^{-1}\, \langle w, \theta' - \theta\rangle + o(|\theta' - \theta|)$$

and hence

$$\limsup_{\theta'\to\theta} \frac{|\,|\nabla_S f(\theta')| - |\nabla_S f(\theta)|\,|}{|\theta'-\theta|} = |\nabla_S f(\theta)|^{-1} \limsup_{\theta'\to\theta} \frac{|\,\langle w, \theta'-\theta\rangle\,|}{|\theta'-\theta|} = |\nabla_S f(\theta)|^{-1} P_{\theta^\perp} w.$$

Thus, by the definition,

$$|\nabla_S^2 f(\theta)| = |\nabla_S f(\theta)|^{-1} P_{\theta^\perp} w.$$

But, as was noted before, the vector w is always orthogonal to θ. Therefore, $P_{\theta^\perp} w = w$, and we arrive at the required identity

$$|\nabla_S^2 f(\theta)| = |\nabla_S f(\theta)|^{-1}\, w \quad \big(|\nabla_S f(\theta)| > 0\big).$$

Finally, consider the remaining case $|\nabla_S f(\theta)| = 0$. Then $L = f_S''(\theta)(\theta'-\theta)$, and (10.4) is simplified to

$$|\nabla_S f(\theta')| = |L| + o(|\theta'-\theta|).$$

Again, by the very definition, and using orthogonality of $f_S''(\theta)h$ to θ,

$$\begin{aligned} |\nabla_S^2 f(\theta)| &= \limsup_{\theta'\to\theta} \frac{|\nabla_S f(\theta')|}{|\theta'-\theta|} \\ &= \limsup_{\theta'\to\theta} \frac{|f_S''(\theta)(\theta'-\theta)|}{|\theta'-\theta|} = \limsup_{h\to 0,\ h\in\theta^\perp} \frac{|f_S''(\theta)h|}{|h|} = \|f_S''(\theta)\|. \end{aligned}$$

□

11. **Appendix D. Laplacian**

The Laplacian operator $\Delta_S f = \mathrm{Tr}\, f_S''$, acting in the class of all C^2-smooth function f on S^{n-1}, can be related to the "spherical partial derivatives" $D_i f(\theta) = \langle \nabla_S f(\theta), e_i\rangle$, where $e_1, \dots, e_n$ is the canonical basis in $\mathbf{R}^n$. Thus,

$$\nabla_S f(\theta) = \sum_{i=1}^n D_i f(\theta)\, e_i.$$

As the next partial derivatives, consider "second order" differential operators

$$D_{ij} f = D_i(D_j f) = \langle \nabla_S \langle \nabla_S f, e_j\rangle, e_i\rangle, \qquad i,j = 1,\dots,n.$$

Proposition 11.1. $\Delta_S = \sum_{i=1}^n D_{ii}$.

In fact, any basis could be used in place of e_i's in the definition of D_{ii}, and the above statement will continue to hold.

Proof. By (9.1), for all $v \in \mathbf{R}^n$,

$$\nabla_S \langle \nabla_S f(\theta), v\rangle = f''_S(\theta)v - \langle v, \theta\rangle \nabla_S f(\theta).$$

Hence

$$D_{ii} f(\theta) = \langle f''_S(\theta) e_i, e_i\rangle - \langle \theta, e_i\rangle \langle \nabla_S f(\theta), e_i\rangle$$

and thus

$$\sum_{i=1}^n D_{ii} f(\theta) \, = \, \mathrm{Tr}\, f''_S(\theta) - \langle \nabla_S f(\theta), \theta\rangle \, = \, \mathrm{Tr}\, f''_S(\theta).$$

$\square$

The following identity was used in the proof of Theorem 1.2.

Proposition 11.2. *If f is defined and C^2-smooth in an open neighborhood of S^{n-1}, then, for all $\theta \in S^{n-1}$,*

$$\Delta_S f(\theta) = \Delta f(\theta) - (n-1) \langle \nabla f(\theta), \theta\rangle - \langle f''(\theta)\theta, \theta\rangle . \tag{11.1}$$

In turn, it is obtained from the following explicit formula for the derivatives D_{ij}.

Lemma 11.3. *If f is C^2-smooth in an open neighborhood of S^{n-1}, then for all $\theta \in S^{n-1}$ and all $i, j = 1, \dots, n$,*

$$\begin{aligned} D_{ij} f(\theta) \ &= \ \partial_{ij} f(\theta) - \theta_j \, \partial_i f(\theta) - \delta_{ij} \langle \nabla f(\theta), \theta\rangle + 2\theta_i\theta_j \langle \nabla f(\theta), \theta\rangle \\ &\quad -\theta_j \langle f''(\theta)\theta, e_i\rangle - \theta_i \langle f''(\theta)\theta, e_j\rangle + \theta_i\theta_j \langle f''(\theta)\theta, \theta\rangle . \end{aligned} \tag{11.2}$$

In particular,

$$\begin{aligned} D_{ii} f(\theta) \ &= \ \partial_{ii} f(\theta) - \langle \nabla f(\theta), \theta\rangle - \theta_i \, \partial_i f(\theta) + 2\theta_i^2 \langle \nabla f(\theta), \theta\rangle \\ &\quad - \, 2\theta_i \langle f''(\theta)\theta, e_i\rangle + \theta_i^2 \langle f''(\theta)\theta, \theta\rangle . \end{aligned}$$

Summing the latter equality over all $i \le n$, we arrive at (11.1).

Note that the operators D_i and D_j are not commutative, i.e. we do not have the identity $D_{ij} f = D_{ji} f$ in the entire class C^2. Indeed, by (11.2), $D_{ij} f(\theta) = D_{ji} f(\theta)$, if and only if $\theta_i \, \partial_j f(\theta) = \theta_j \, \partial_i f(\theta)$.

Proof of Lemma 11.3. Assume f is C^2-smooth in the open region G. Fix an index $j \le n$ and consider the smooth function in n real variables

$$\begin{aligned} u(x) \ &= \ \langle \nabla f(x), e_j\rangle - \langle \nabla f(x), x\rangle \langle x, e_j\rangle \\ &= \ \partial_j f(x) - x_j \sum_{k=1}^n x_k \, \partial_k f(x), \qquad x = (x_1, \dots, x_n) \in G. \end{aligned} \tag{11.3}$$

In particular, $u(\theta) = D_j\, f(\theta)$ for $\theta \in S^{n-1}$ and therefore $D_{ij} f = D_i u$. Again following the definition of D_i, we have

$$D_i\, u(x) = \partial_i u(x) - x_i \sum_{l=1}^n x_l\, \partial_l u(x). \tag{11.4}$$

By (11.3),

$$\begin{aligned}
\partial_i u(x) &= \partial_{ij} f(x) - \delta_{ij} \sum_{k=1}^n x_k\, \partial_k f(x) - x_j\, \partial_i f(x) - x_j \sum_{k=1}^n x_k\, \partial_{ik} f(x), \\
\partial_l u(x) &= \partial_{lj} f(x) - \delta_{lj} \sum_{k=1}^n x_k\, \partial_k f(x) - x_j\, \partial_l f(x) - x_j \sum_{k=1}^n x_k\, \partial_{lk} f(x).
\end{aligned}$$

Plugging these equalities in (11.4), we get

$$\begin{aligned}
D_i\, u(x) &= \partial_{ij} f(x) - \delta_{ij} \sum_{k=1}^n x_k\, \partial_k f(x) - x_j\, \partial_i f(x) - x_j \sum_{k=1}^n x_k\, \partial_{ik} f(x) \\
&\quad - x_i \sum_{l=1}^n x_l \left[\partial_{lj} f(x) - \delta_{lj} \sum_{k=1}^n x_k\, \partial_k f(x) - x_j\, \partial_l f(x) - x_j \sum_{k=1}^n x_k\, \partial_{lk} f(x)\right] \\
&= \partial_{ij} f(x) - \delta_{ij} \sum_{k=1}^n x_k\, \partial_k f(x) - x_j\, \partial_i f(x) - x_j \sum_{k=1}^n x_k\, \partial_{ik} f(x) \\
&\quad - x_i \sum_{l=1}^n x_l\, \partial_{lj} f(x) + 2 x_i x_j \sum_{k=1}^n x_k\, \partial_k f(x) + x_i x_j \sum_{l=1}^n \sum_{k=1}^n x_l x_k\, \partial_{lk} f(x).
\end{aligned}$$

In a bit more compact form,

$$\begin{aligned}
D_i\, u(x) &= \partial_{ij} f(x) - x_j\, \partial_i f(x) - \delta_{ij} \left\langle \nabla f(x), x \right\rangle + 2 x_i x_j \left\langle \nabla f(x), x \right\rangle \\
&\quad - x_j \left\langle f''(x) x, e_i \right\rangle - x_i \left\langle f''(x) x, e_j \right\rangle + x_i x_j \left\langle f''(x) x, x \right\rangle .
\end{aligned}$$

It remains to restrict this function to the sphere. $\square$

12. Appendix E. Homogeneous functions

A function $F : \mathbf{R}^n \setminus \{0\} \to \mathbf{R}$ is called homogeneous of order d (where d is a real number), or d-homogeneous, if it satisfies the relation

$$F(\lambda x) = \lambda^d\, F(x), \qquad x \neq 0,\ \lambda > 0.$$

Any such function is completely determined by its values on the unit sphere. Alternatively, starting from a function f on S^{n-1}, one may define its unique d-homogeneous extension by putting

$$F(x) = r^d F(r^{-1}x), \qquad r = |x|, \quad x \neq 0.$$

For example, if $f = 1$, then $F(x) = |x|^d$.

In this section, we collect several formulas for the derivatives of d-homogeneous functions. We will use the notations

$$r = |x|, \qquad \theta = r^{-1}x = \frac{x}{|x|} \qquad (x \neq 0).$$

Proposition 12.1. *For the d-homogeneous extension $F(x) = r^d f(\theta)$ of a C^1-smooth function f on S^{n-1}, we have that, for all $x \neq 0$,*

$$\nabla F(x) = r^{d-1}\big[d\, f(\theta)\theta + \nabla_S f(\theta)\big]. \tag{12.1}$$

This formula can be easily verified by the direct differentiation (assuming that f is defined and C^1-smooth in a neighborhood of the sphere), so we omit the proofs.

For example, for the 1-homogeneous extension $F(x) = r f(\theta)$, we have

$$\nabla F(x) = f(\theta)\theta + \nabla_S f(\theta), \qquad |\nabla F(x)|^2 = f(\theta)^2 + |\nabla_S f(\theta)|^2.$$

In this particular case, such functions may be used, for example, to recover the Poincaré inequality on the sphere on the basis of the Poincaré-type inequality for the Gaussian measure (which in turn has many elementary proof).

For the 0-homogeneous extension $F(x) = f(\theta)$, we have

$$\nabla F(x) = r^{-1}\, \nabla_S f(\theta),$$

and thus the usual (Euclidean) and spherical gradients coincide on the unit sphere: $\nabla F = \nabla_S f$ on S^{n-1}.

It is therefore interesting to know whether a similar identity holds for the second derivative as well. The answer is negative, although some relationship does exist.

Proposition 12.2. *For the d-homogeneous extension $F(x) = r^d f(\theta)$ of a C^2-smooth function f on S^{n-1}, we have for all $x \neq 0$ and $v \in \mathbf{R}^n$,*

$$\begin{aligned} F''(x)v \ &= \ r^{d-2}\Big[d(d-1) f(\theta)\, \langle v, \theta\rangle\, \theta + df(\theta) P_{\theta^\perp} v \\ &\qquad + \ (d-1)\, \langle \nabla_S f(\theta), v\rangle\, \theta + (d-1)\, \langle v, \theta\rangle\, \nabla_S f(\theta) + f_S''(\theta) v\Big]. \end{aligned} \tag{12.2}$$

In interesting particular cases $d = 0, 1$, (12.2) is simplified. For the 0-homogeneous extension, we have

$$F''(x)v = r^{-2}\,\big[-\langle \nabla_S f(\theta), v\rangle\, \theta - \langle v, \theta\rangle\, \nabla_S f(\theta) + f_S''(\theta) v\big],$$

while for the 1-homogeneous extension,

$$F''(x) = r^{-1}\left[f(\theta)P_{\theta^\perp} + f_S''(\theta)\right].$$

Proof. From (12.1),

$$\langle \nabla F(x), v\rangle = r^{d-1}\big[df(\theta)\,\langle v,\theta\rangle + \langle \nabla_S f(\theta), v\rangle\,\big].$$

We are in position to apply (12.1) once more, now with $d-1$ in place of d and with

$$\psi(\theta) = df(\theta)\,\langle v,\theta\rangle + \langle \nabla_S f(\theta), v\rangle$$

in place of f. It gives

$$\begin{aligned} F''(x)v &= \nabla\,\langle \nabla F(x), v\rangle \\ &= r^{d-2}\big[(d-1)\psi(\theta)\theta + \nabla_S\psi(\theta)\big] \\ &= r^{d-2}\big[\,d(d-1)f(\theta)\,\langle v,\theta\rangle\,\theta + (d-1)\,\langle \nabla_S f(\theta), v\rangle\,\theta + \nabla_S\psi(\theta)\big]. \end{aligned} \tag{12.3}$$

To develop the last gradient, using $\nabla_S\,\langle v,\theta\rangle = Pv$ ($P = P_{\theta^\perp}$), first write

$$\nabla_S\psi(\theta) = d\,\langle v,\theta\rangle\,\nabla_S f(\theta) + df(\theta)Pv + \nabla_S\,\langle \nabla_S f(\theta), v\rangle\,.$$

In order to evaluate the last gradient, we apply the identity (9.1), which gives

$$\nabla_S\psi(\theta) = (d-1)\,\langle v,\theta\rangle\,\nabla_S f(\theta) + df(\theta)Pv + f_S''(\theta)v.$$

Inserting this expression in (12.3), we arrive at the formula (12.2). $\square$

Corollary 12.3. *For the d-homogeneous extension $F(x) = r^d f(\theta)$ of a C^2-smooth function f on S^{n-1}, we have for all $x \neq 0$,*

$$\Delta F(x) \;=\; r^{d-2}\big[\,d(n+d-2)f(\theta) + \Delta_S f(\theta)\big]. \tag{12.4}$$

In particular, $\Delta F(x) = r^{d-2}\Delta_S f(\theta)$ for the 0-homogeneous extension $F(x) = f(\theta)$, so the Euclidean and spherical Laplacians coincide on the unit sphere. The same conclusion is also true when $d = 2-n$.

The identity (12.4) is well-known. It implies that, for any spherical harmonic f on S^{n-1} of degree d (so that $\Delta F = 0$), we necessarily have $\Delta_S f = -d(n+d-2)f$, cf. e.g. [S-W].

Proof. Applying (12.2) with $v = e_i$, we get

$$\begin{aligned} \langle F''(x)e_i.e_i\rangle \;&=\; r^{d-2}\big[d(d-1)f(\theta)\,\langle \theta, e_i\rangle^2 + df(\theta)\,\langle P_{\theta^\perp}e_i, e_i\rangle \\ &\quad + 2(d-1)\,\langle \nabla_S f(\theta), e_i\rangle\,\langle \theta, e_i\rangle + \langle f_S''(\theta)e_i, e_i\rangle\,\big]. \end{aligned}$$

Here $P_{\theta^\perp}e_i = e_i - \langle \theta, e_i\rangle\,\theta$, so

$$\sum_{i=1}^n \langle P_{\theta^\perp}e_i, e_i\rangle = \sum_{i=1}^n \big(1 - \langle \theta, e_i\rangle^2\,\big) = n-1.$$

In addition,

$$\sum_{i=1}^n \langle \nabla_S f(\theta), e_i \rangle \langle \theta, e_i \rangle = \langle \nabla_S f(\theta), \theta \rangle = 0.$$

Hence,

$$\Delta F(x) = \sum_{i=1}^n \langle F''(x)e_i.e_i \rangle = r^{d-2}\big[d(d-1)f(\theta) + d(n-1)\, f(\theta) + \Delta_S f_S''(\theta)\big].$$

□

13. Appendix F. Commutator of Laplacian and Gradient

If a function f is defined and C^3-smooth in an open region of $\mathbf{R}^n$, then, for any $v \in \mathbf{R}^n$,

$$\Delta \langle \nabla f(x), v \rangle = \langle \nabla \Delta f(x), v \rangle \tag{13.1}$$

throughout the region. In a more compact form, $\Delta\nabla = \nabla\Delta$, that is, these two operators – the Euclidean Laplacian and the Euclidean gradient – commute. However, due to curvature of S^{n-1}, this is no longer true for the spherical Laplacian and the spherical gradient which may be seen from the formula for the commutator given in Lemma 4.3. In a more compact vector form, this formula may be written as

$$\Delta_S \nabla_S f(\theta) - \nabla_S \Delta_S f(\theta) \,=\, (n-3)\,\nabla_S f(\theta) - 2\Delta_S f(\theta)\,\theta.$$

This identity may also be rewritten component-wise in terms of the operators D_i as

$$\Delta_S D_i f(\theta) - D_i\,\Delta_S f(\theta) \,=\, (n-3) D_i f(\theta) - 2\,\langle \theta, e_i \rangle\, \Delta_S f(\theta).$$

Proof of Lemma 4.3. By Proposition 12.1 and Corollary 12.3, for any C^3-smooth d-homogeneous function u on $\mathbf{R}^n \setminus \{0\}$, for all $x \neq 0$,

$$\nabla u(x) \;=\; r^{d-1}\big[\,d\,u(\theta)\theta + \nabla_S u(\theta)\big], \tag{13.2}$$

$$\Delta u(x) \;=\; r^{d-2}\big[\,d(n+d-2)\,u(\theta) + \Delta_S u(\theta)\big], \tag{13.3}$$

where $r = |x|$ and $\theta = r^{-1}x$.

The identity (13.1) will be used with the 0-homogeneous extension $F(x) = f(\theta)$, $x \neq 0$. Being restricted to the points lying on the unit sphere, it becomes

$$\Delta \langle \nabla F(\theta), v \rangle = \langle \nabla \Delta F(\theta), v \rangle\,. \tag{13.4}$$

In that case, $\nabla_S f(\theta) = \nabla F(\theta)$, so

$$\Delta_S \langle \nabla_S f(\theta), v \rangle = \Delta_S \langle \nabla F(\theta), v \rangle\,.$$

Moreover, the function $u(x) = \langle \nabla F(x), v \rangle$ is (-1)-homogeneous, and we may apply (13.3) with $d = -1$. Again, being restricted to the unit sphere, this identity becomes

$$\Delta u(\theta) = -(n-3)\,u(\theta) + \Delta_S u(\theta),$$

so

$$\Delta_S \langle \nabla_S f(\theta), v\rangle \;=\; \Delta_S u(\theta) \;=\; \Delta u(\theta) + (n-3)\, u(\theta)$$
$$= \Delta \langle \nabla F(\theta), v\rangle + (n-3)\, \langle \nabla F(\theta), v\rangle\,. \qquad (13.5)$$

On the other hand, the function $u(x) = \Delta F(x)$ is (-2)-homogeneous, and we may apply (13.2) with $d = -2$. It gives

$$\nabla u(\theta) = -2u(\theta)\,\theta + \nabla_S u(\theta).$$

Since ΔF coincides with $\Delta_S f$ on S^{n-1}, we get that

$$\begin{aligned}\langle \nabla_S \Delta_S f(\theta), v\rangle &= \langle \nabla_S\, u(\theta), v\rangle \\ &= \langle \nabla u(\theta), v\rangle + 2u(\theta)\,\langle \theta, v\rangle \\ &= \langle \nabla \Delta F(\theta), v\rangle + 2\Delta F(\theta)\,\langle \theta, v\rangle\,. \qquad (13.6)\end{aligned}$$

It remains to subtract (13.6) from (13.5) and apply (13.4), which leads to

$$\Delta_S \langle \nabla_S f(\theta), v\rangle - \langle \nabla_S \Delta_S f(\theta), v\rangle \;=\; (n-3)\,\langle \nabla F(\theta), v\rangle - 2\,\langle \theta, v\rangle\, \Delta F(\theta).$$

But $\nabla F(\theta) = \nabla_S f(\theta)$ and $\Delta F(\theta) = \Delta_S f(\theta)$. $\square$

14. **Appendix G. Integrals Involving Laplacian**

Integrals involving the spherical Laplacian may be evaluated using integration by parts (in general based on Gauss divergence or Stokes formula for arbitrary manifolds without boundaries) which on S^{n-1} may be based on the equality

$$\int f\ \nabla_S g\, d\sigma_{n-1} = -\int \nabla_S f\ g\, d\sigma_{n-1}.$$

In particular, it yields

$$\int \langle \nabla_S f, \nabla_S g\rangle\ d\sigma_{n-1} = -\int f \Delta_S g\, d\sigma_{n-1}. \qquad (14.1)$$

In fact, this identity and the one of Lemma 4.4 may be extended to a more general Green-type formula.

Proposition 14.1. *For all C^2-smooth functions f, g and any C^1-smooth function w on S^{n-1},*

$$\int \langle \nabla_S f, \nabla_S g\rangle\ w\, d\sigma_{n-1} = -\int f \Delta_S g\, w\, d\sigma_{n-1} - \int f\, \langle \nabla_S g, \nabla_S w\rangle\, d\sigma_{n-1}. \qquad (14.2)$$

When $w = 1$, we return to the formula (14.1). In the case of the linear weight $w(\theta) = \langle v, \theta\rangle$, let us recall that $\nabla_S w(\theta) = P_{\theta^\perp} v = v - \langle v, \theta\rangle\, \theta$ and that $\nabla_S g(\theta)$ is orthogonal to θ. Hence, (14.2) is simplified to

$$\begin{aligned}\int \langle \nabla_S f(\theta), \nabla_S g(\theta)\rangle \, \langle v,\theta\rangle \, d\sigma_{n-1}(\theta) \ &= \ -\int f(\theta)\Delta_S g(\theta)\, \langle v,\theta\rangle \, d\sigma_{n-1}(\theta) \\ &\quad - \int f(\theta)\, \langle \nabla_S g(\theta), v\rangle \ d\sigma_{n-1}(\theta),\end{aligned}$$

which is the statement of Lemma 4.4.

Proof of Proposition 14.1. Using the canonical basis in $\mathbf{R}^n$, first write

$$\begin{aligned}\int \langle \nabla_S f, \nabla_S g\rangle \, w \, d\sigma_{n-1} \ &= \ \sum_{i=1}^n \int \langle \nabla_S f, e_i\rangle \, \langle \nabla_S g, e_i\rangle \, w \, d\sigma_{n-1} \\ &= \ \sum_{i=1}^n \Big\langle \int \nabla_S f \, \langle \nabla_S g, e_i\rangle \, w \, d\sigma_{n-1}, e_i \Big\rangle. \qquad (14.3)\end{aligned}$$

Integrating by parts and applying the formula for the spherical gradient of the product of two functions, $\nabla_S(\varphi\psi) = \psi\nabla_S\varphi + \varphi\nabla_S\psi$, we see that the last vector integral in (14.3) is equal to

$$\begin{aligned}-\int f \, \nabla_S\big(\langle \nabla_S g, e_i\rangle \, w\big) \, d\sigma_{n-1} \ &= \ -\int f\Big(w \, \nabla_S \, \langle \nabla_S g, e_i\rangle + \langle \nabla_S g, e_i\rangle \, \nabla_S w\Big) \, d\sigma_{n-1} \\ &= \ -\int f \, w \, D_{ii} g \, d\sigma_{n-1} \ e_i - \int f \, \langle \nabla_S g, e_i\rangle \, \nabla_S w \, d\sigma_{n-1}.\end{aligned}$$

Hence, the sum in (14.3) is equal to

$$\begin{aligned}&-\int fw \, \Delta_S g \, d\sigma_{n-1} - \sum_{i=1}^n \int f \, \langle \nabla_S g, e_i\rangle \, \langle \nabla_S w, e_i\rangle \ d\sigma_{n-1} \\ &\quad = -\int fw \, \Delta_S g \, d\sigma_{n-1} - \int f \, \langle \nabla_S g, \nabla_S w\rangle \ d\sigma_{n-1}.\end{aligned}$$

$\square$

Finally, let us emphasize one useful consequence of the formula (14.1) which we state as a relation between integrals of vector-valued functions.

Corollary 14.2. *For any smooth function f on S^{n-1},*

$$\int f(\theta)\theta \, d\sigma_{n-1}(\theta) = \frac{1}{n-1} \int \nabla_S f(\theta) \, d\sigma_{n-1}(\theta).$$

In particular, f is orthogonal to all linear functions in $L^2(S^{n-1})$, if and only if all linear forms $\langle \nabla_S f(\theta), v\rangle$ have σ_{n-1}-mean zero.

Proof. The linear function $g(\theta) = \langle v, \theta \rangle$ has the spherical gradient and respectively the spherical Laplacian

$$\nabla_S g(\theta) = P_{\theta^\perp} v, \qquad \Delta_S g(\theta) = -(n-1)\,\langle v, \theta \rangle\,.$$

In this case, (13.1) becomes

$$\int \langle \nabla_S f(\theta), P_{\theta^\perp} v \rangle \; d\sigma_{n-1}(\theta) = (n-1) \int f(\theta)\,\langle v, \theta \rangle \; d\sigma_{n-1}(\theta).$$

But $\nabla_S f(\theta)$ is orthogonal to θ, so, $\langle \nabla_S f(\theta), P_{\theta^\perp} v \rangle = \langle \nabla_S f(\theta), v \rangle$. It follows that

$$\int \langle \nabla_S f, v \rangle \; d\sigma_{n-1} = (n-1) \int f(\theta)\,\langle v, \theta \rangle \; d\sigma_{n-1}(\theta)$$

which is an equivalent form of the required identity. □

References

[A-M-S] S. Aida, T. Masuda, I. Shigekawa. Logarithmic Sobolev inequalities and exponential integrability. J. Func. Anal. 126 (1994), 83–101.

[B-G-L] D. Bakry, I. Gentil, M. Ledoux. Analysis and Geometry of Markov Diffusion Operators. Springer, Berlin Heidelberg New York, 2014.

[B-G] S. G. Bobkov, F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. J. Func. Anal. 1 (1999), 1–28.

[K] B. Klartag. A Berry-Esseen type inequality for convex bodies with an unconditional basis. Probab. Theory Related Fields 145 (2009), no. 1-2, 1-33.

[K-S] B. Klartag, S. Sodin. Variations on the Berry-Esseen theorem. Teor. Veroyatn. Primen. 56 (2011), no. 3, 514–533; translation in Theory Probab. Appl. 56 (2012), no. 3, 403-419.

[L1] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. Séminaire de Probabilités XXXIII. Lect. Notes in Math. 1709 (1999), 120–216, Springer.

[L2] M. Ledoux. The concentration of measure phenomenon. Math. Surveys and monographs, vol. 89, AMS, 2001.

[M-S] V. D. Milman, G. Schechtman. Asymptotic theory of finite-dimensional normed spaces. With an appendix by M. Gromov. Lecture Notes in Mathematics, 1200, Springer-Verlag, Berlin, 1986, viii+156 pp.

[M-W] C. E. Mueller, F. B. Weissler. Hypercontractivity for the heat semigroup for ultraspherical polynomials and on the n-sphere. J. Funct. Anal. 48 (1992), 252-283.

[S-W] E. M. Stein, G. Weiss. Introduction to Fourier analysis on Euclidean spaces. Princeton Mathematical Series, No. 32. Princeton University Press, Princeton, N.J., 1971. x+297 pp.

[S] V. N. Sudakov. Typical distributions of linear functionals in finite-dimensional spaces of high dimension. (Russian) Soviet Math. Dokl. 19 (1978), 1578–1582; translation in: Dokl. Akad. Nauk SSSR, 243 (1978), no. 6, 1402-1405.

[T1] M. Talagrand. A new isoperimetric inequality for product measure and the tails of sums of independent random variables. Geom. Funct. Anal. 1 (1991), no. 2, 211-223.

[T2] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. Publications Mathématiques de l'I.H.E.S. 81 (1995), 73–205.

Sergey G. Bobkov
School of Mathematics, University of Minnesota
127 Vincent Hall, 206 Church St. S.E., Minneapolis, MN 55455 USA
E-mail address: `bobkov@math.umn.edu`

Gennadiy P. Chistyakov
Fakultät für Mathematik, Universität Bielefeld
Postfach 100131, 33501 Bielefeld, Germany
E-mail address: `chistyak@math.uni-bielefeld.de`

Friedrich Götze
Fakultät für Mathematik, Universität Bielefeld
Postfach 100131, 33501 Bielefeld, Germany
E-mail address: `goetze@mathematik.uni-bielefeld.de`